



Station Biologique  
Roscoff



Thèse de Doctorat de l'Université Pierre & Marie Curie  
Spécialité DIVERSITE DU VIVANT

Présentée par

**Nathalie KOWALCZYK**

Pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITE PIERRE ET MARIE CURIE

Sujet de la thèse

**Le stress abiotique  
chez l'algue rouge *Chondrus crispus*  
développement d'un modèle d'études  
physiologiques et moléculaires**

Soutenue le 11 Janvier 2013  
devant le jury composé de :

**Rapporteurs**

Giovanni FINAZZI

DR CNRS CEA Grenoble

Claire LURIN

DR INRA URGV Evry

**Examineurs**

Jonas COLLEN

MC UPMC Roscoff

Directeur de thèse

Patrick CORMIER

PR UPMC Roscoff

Président du jury

Ivan COUEE

PR Université de Rennes 1

Angela FALCIATORE

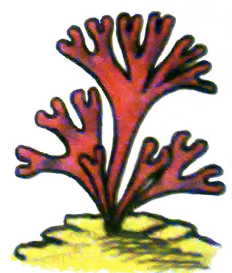
CR CNRS Paris

Pierre JOLIOT

PR UPMC IBPC Paris



à mon Pap's...







## Une thèse c'est un peu...

- de biologie,
- de rencontres scientifiques à but professionnel,
- un combat contre son cerveau surchargé,
- de remises en questions,
- de victoires aussi, parfois,
- épuisant et inhumain, notamment le vendredi lorsqu'il est précédé d'un jeudi.

Mais c'est surtout :

- des rencontres scientifiques (ou pas) à but non lucratif,
- des jeudis,
- de la procrastination et des trolls,
- de la biographie,
- de la musique. Beaucoup de musique.

Bref, une thèse c'est loin d'être que de la science...

Vous ne l'aviez pas remarqué ?

Maintenant, vous ne verrez plus que ça !

Oui, je suis machiavélique.



## Remerciements

Merci (aux grands) :

- Mon jury, Claire Lurin, Giovanni Finazzi, Pierre Joliot, Patrick Cormier, Angela Falciatore, Ivan Couée et Jonas Collén d’avoir accepté de lire et écouter mes digressions autour des aventures de *Chondrus*, l’algue la plus funky de toute la phycologie.
- Bernard Kloareg, *Trugarez dit Bernez evit bezañ degemeret ac’hanon e familh bras Savlec’h Biologel Rosko*.
- Catherine Boyen et Jonas Collén pour votre encadrement durant ces trois années (et demi), pour votre confiance, pour votre soutien, pour votre patience, votre gentillesse et surtout de m’avoir permise de participer à cette aventure complètement inhumaine.
- Pierre Joliot, Fabrice Rappaport et Francis-André Wollman d’avoir su rendre la biophysique non seulement moins obscure mais tout à fait fascinante à mes yeux.
- Les membres de mon comité de thèse, Ivan, Angela et Christophe Six, pour leur gentillesse et leurs encouragements.
- La région Bretagne, pour la “Bretagne” et aussi accessoirement de m’avoir nourrie, logée et blanchie.

Merci (les collègues et collaborateurs) :

- Brigitte, pour les causettes à l’accueil.
- Maryvonne, pour ta bonne humeur et ta patience envers toutes nos paperasseries.
- Sylvie d’avoir RT-qPCRisé pour moi avec brio et de ne pas m’avoir maudite lorsque j’oubliais des plaques dans la glace.
- Wilfrid d’avoir patiemment répondu à mes questions bioinformatiques et d’avoir subi aussi les reads du génoscope.
- Erwan, Gildas Lecorguillé (alors ce Diablo III ?) et les autres ABIMSiens, d’avoir gardé éveillé le démon geek qui sommeille en moi.
- Laurent (tiens!), pour la microscopie, pour tes idées du tonnerre, mais surtout pour ta bonne humeur et ton énergie qui sont hautement contagieuses. Tu es un véritable rayon de soleil : le remède idéal contre la baisse de motivation (aka. la pseudo-dépression du thésard en deuxième année) !
- Simon, pour ta gentillesse et ta précieuse aide statistique.
- Le bureau 336, Laurence, parce que tu m’as foutue dans ce train (mais pas que, hein), Sophie, et mon trouple.
- Lionel, son gros zoom, mon chocolat et le pommeau de douche.
- Bernard, la sérénité incarnée et Bénédicte la tornade.
- Et toute l’UMR7139.

Merci (à mon trouple) :

- Zofia (Renée), d’avoir instillé toutes sortes de breizhoneries dans mon petit cerveau, pour m’avoir initiée aussi bien aux danses bretonnes qu’à l’accordéon, mais aussi aux panneaux et aux autocollants, et à la meilleure coloc’ qui soit. Sans toi, la vie en Breizh aurait été ’achement moins funky. *Ur banne chistr, mar plij!*
- Jean-Baptou (Kozi Serek) de m’avoir acceptée au sein de ton harem, et parce que rien ne vaut un bon Circassien ou une Scottish Yiddish avec toi. Ceci est un message subliminal.

Pour tous ceux qui ne liront peut-être jamais ce manuscrit, néanmoins, je prends mes dispositions... sait-on jamais, vous avez illuminé mes week-ends et m’avez sortie la tête de l’eau (salée),

*mersi bras deoc'h :*

- Pouch, notre Patron préféré. Le grand oiseau en trois restera le meilleur chez moi du monde.
- Olivier-boudin-gris, pour les nuits blanches, le swing, et parce que t'es aussi un super coloc' (ex-eaquo, hein).
- Clem, mon punk-sans-chien préféré. Goodman Youssef, lovely, lovely, come back for the craic !
- Hervé, d'avoir bien voulu être mon cavalier récurrent, même en période de tachycardie.
- Alan, pour tes formidables festoù-noz à l'île de Batz.
- Glen et les Dispac'h'iens, parce que vous êtes indétrônablement le meilleur groupe à danser de toute la Bretagne mondiale.
- Julie et Kabac'h d'avoir accepté que je fasse partie de votre groupe et de m'avoir laissée le temps de progresser.
- Tous les circassiens du Léon, j'aurais aimé m'investir davantage.

Merci (les Roskovites) :

- Pauline, pour les litres de tisanes à midi et les baignades d'hiver (toute l'année!).
- Kevin, stagiaire présidentiel (je suis machiavélique).
- Mahdi, parce que tu es l'âme de la station.
- Noan (Voisinou!) d'avoir toléré avec le sourire mes siècles (cumulés) de répétitions accordéonistiques (et violonistiques) même/surtout à 4h du mat'.
- Alex et Julie, pour l'Auvergne... Mmmm, ça gueume... (mais pas que).
- Alexia et Mick, pour les sessions grimpe, slack et forêt.
- Ronan, Dr Jekyll le jour & Mr Hyde après quelques bières.

Merci les non-roskovites-anymore (ou les rescapés) :

- Gael, mon vermouthologue préféré. Toi, tu as l'âme pure du vrai chercheur.
- L'ami René, ami-des-amies-des-René(e)s, pour le formidable mois de mai et d'avoir patiemment lu et commenté mes folles aventures bretonnes.
- Etienne, je ne sais que dire, car si je commence, cette thèse ne sera pas scientifique mais littéraire.
- Mathieu, parce qu'on est un peu trop pareil, sauf que toi tu t'en sors mieux.
- Caroline, pour les sauvetages de poulpes nocturnes.
- Carina, pour une première super coloc' et les correspondances à l'étranger (un peu partout).

Les vraiment-jamais-roscovites (ou presque) :

- Manu SG aka. l'Ours, oui tu es parfait (enfin à peu près).
- Fanny, mon Boudinou irremplaçable!
- Laura, ma correspondante à l'étranger (c'est à dire en France).
- Dimitri, pour la hotline du nouveau monde et le "bon de colère".

La famille, d'ici et d'ailleurs :

- mon Pap's, tu as toujours crû en moi et tu es parti avant même de voir le résultat de ton investissement... C'est pas juste.
- ma môman, ma soeur, mon beau-frère, ma nièce, vous qui avez supporté mon absence sans broncher.
- mon oncle, tu étais là quand il faut.
- et toute la famille de Pologne et du monde entier, promis, je viendrai!

# Table des matières

<b>Abréviations</b>	<b>11</b>
<b>Avant-propos</b>	<b>12</b>
<b>1 Introduction</b>	<b>15</b>
1.1 Le modèle <i>Chondrus crispus</i> . . . . .	16
1.1.1 Un cycle de vie trigénétique . . . . .	19
1.1.2 Les carraghénanes . . . . .	21
1.1.3 <i>Chondrus crispus</i> en culture . . . . .	22
1.2 Introduction à la génomique . . . . .	24
1.2.1 Un peu d'histoire... . . . .	25
1.2.2 La transcriptomique quantitative . . . . .	28
1.2.3 La génomique en milieu naturel et la métagénomique . . . . .	28
1.3 Le stress abiotique . . . . .	29
1.3.1 Le stress abiotique chez les plantes terrestres . . . . .	30
1.3.2 La marée : source de stress . . . . .	33
1.3.3 Le stress oxydatif : résultat de facteurs multiples . . . . .	36
1.3.4 Le stress chez les algues . . . . .	40
1.3.5 La lumière : à consommer avec modération . . . . .	42
1.4 La photosynthèse comparée . . . . .	43
1.4.1 Une obscure clarté... . . . .	43
1.4.2 Fonctionnement de l'appareil photosynthétique . . . . .	44
1.4.3 Les phycobilisomes . . . . .	47
1.4.4 Régulation de la lumière . . . . .	49
1.5 L'étude de la photosynthèse <i>in vivo</i> . . . . .	52
1.5.1 La fluorescence . . . . .	54
1.5.2 Spectrométrie P <sub>700</sub> et b <sub>6</sub> f . . . . .	54
1.5.3 Phénomènes de photoprotection . . . . .	55
1.6 Avant d'attaquer... . . . .	58
<b>2 Matériels &amp; Méthodes</b>	<b>59</b>
2.1 Les algues . . . . .	59

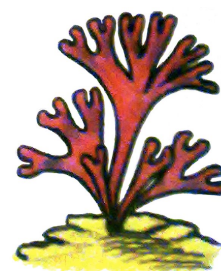
2.2	Biologie moléculaire & bioinformatique . . . . .	60
2.2.1	Séquençage du génome et annotation automatique . . . . .	60
2.2.2	Transcriptomique - stades de vie . . . . .	61
2.2.3	Transcriptomique - stress . . . . .	61
2.2.4	RT-PCR quantitative . . . . .	62
2.3	Biophysique . . . . .	64
<b>3</b>	<b>Résultats &amp; Discussion</b>	<b>66</b>
3.1	Le génome de <i>Chondrus crispus</i> . . . . .	66
3.1.1	Le génome en chiffres . . . . .	67
3.1.2	L'annotation . . . . .	68
3.2	Le transcriptome de <i>Chondrus crispus</i> . . . . .	76
3.2.1	Caractéristiques générales . . . . .	77
3.2.2	Les stades de vie de <i>Chondrus crispus</i> . . . . .	80
3.2.3	Lorsque <i>Chondrus crispus</i> stress... . . . .	100
3.3	À la recherche des gènes de normalisation de RT-qPCR . . . . .	108
	Préambule . . . . .	108
	RT-qPCR normalization genes (pré-article) . . . . .	110
3.4	La photosynthèse en rouge . . . . .	133
	Préambule . . . . .	133
	Modèle . . . . .	134
	Photosynthesis in red (article) . . . . .	135
<b>4</b>	<b>Conclusion &amp; Perspectives</b>	<b>145</b>
	<b>Références</b>	<b>150</b>
	<b>Annexes</b>	<b>163</b>
	<b>Communications</b>	<b>177</b>





## Abréviations

ADN : acide désoxyribonucléique  
ADP : adénosine diphosphate  
APC : allophycocyanine  
ARN : acide ribonucléique  
ATP : adénosine triphosphate  
CAT : catalase  
Da : Dalton  
dd-NTP : didésoxyribonucléotide  
EST : expressed sequence tag (étiquette de séquence transcrite)  
 $F_0$  : fluorescence basale  
 $F_M$  : fluorescence maximale  
 $F_V$  : fluorescence variable  
FW : fresh weight (masse fraîche)  
HSP : heat shock protein  
LHC : light harvesting complex (antenne collectrice de lumière)  
Mbp : million de paires de bases  
NADP : nicotinamide adénine dinucléotide phosphate (oxydé)  
NADPH : nicotinamide adénine dinucléotide phosphate (réduit)  
NPQ : non photochemical quenching (extinction de fluorescence non photochimique)  
NSW : natural sea water (eau de mer naturelle)  
OCP : orange carotenoid protein (protéine caroténoïde orange)  
PAM : pulse amplitude modulated (pulse modulé en amplitude)  
pb : paire de bases  
PBS : phycobilisome  
PC : phycocyanine  
PCR : polymerase chain reaction (amplification par polymérisation en chaîne)  
PE : phycoérythrine  
PQ : photochemical quenching (extinction de fluorescence photochimique)  
PS : photosystème  
ROS : reactive oxygen species (espèce réactive de l'oxygène)  
ROSE : ROS scavenging enzyme (enzyme de capture des ROS)  
RT-qPCR : reverse transcription quantitative PCR (retrotranscription et PCR quantitative)  
SOD : superoxide dismutase



## Avant-propos

### De l'étude du stress chez *Chondrus crispus*

Le stress environnemental - ou interactions abiotiques - est une thématique importante aussi bien sur le plan économique qu'écologique, et a longuement été étudié chez les plantes terrestres, (*i.e.*[1]). D'une part sur un plan fondamental à l'aide de modèles comme *Arabidopsis thaliana*, et de manière plus appliquée chez des plantes d'intérêt agronomique (riz, blé, vigne). En revanche, les connaissances concernant d'autres plantes "sauvages" ou encore le milieu marin et ses occupants sont encore limitées.

La zone intertidale, dans laquelle vivent une partie des végétaux marins, offre un univers tout à fait original, avec des contraintes différentes de celles rencontrées en milieu terrestre. Alors que les plantes peuvent subir des changements de température et des épisodes de sécheresse sur des longues durées (plusieurs jours), les algues de la zone intertidale sont exposées à des variations de salinité, température et lumière courtes, brutales et cycliques. Les algues forment la majeure partie de la végétation de l'estran et sont adaptées à ces contraintes, grâce à des mécanismes pouvant être différents de ceux des plantes terrestres.

L'objectif de ma thèse est d'essayer de comprendre quelles sont les stratégies mises en place par les macroalgues rouges pour vivre dans un milieu aussi dynamique que la zone de balancement des marées.

Pour répondre à cette question, j'étudie l'algue rouge *Chondrus crispus*, en analysant les réponses de l'algue à des facteurs de stress *in situ*, en me concentrant sur les effets de la lumière et de la salinité. Mon travail est basé sur l'hypothèse que la réponse au stress est fortement liée au métabolisme oxydatif. Ce projet de recherche est mené en combinant plusieurs approches :

- au niveau physiologique, en étudiant les capacités de l'appareil photosynthétique à réguler l'acquisition de la lumière
- au niveau moléculaire, en analysant l'expression globales des gènes lors d'un stress lumineux et/ou salin.

En terme de génomique fonctionnelle, *C. crispus* est un modèle émergent et ne possède pas tous les outils et techniques d'études existant chez d'autres modèles plus anciens, telle qu'*A. thaliana* ou l'algue verte unicellulaire *Chlamydomonas reinhardtii*. Il est devenu nécessaire d'adapter, optimiser ou créer des nouveaux protocoles pour *C. crispus*, afin d'obtenir des résultats pertinents et exploitables. Ainsi, une part conséquente du travail qui suit consiste en la mise au



point de techniques qui seront par la suite appliquées à la thématique principale de ma thèse.

### **Du déroulement de ce manuscrit**

L'essentiel du texte est rédigé en français, à part l'article "photosynthèse en rouge" et pré-manuscrit "gènes de normalisation de RT-qPCR" , qui eux sont écrits en anglais. Plutôt que de réécrire le contenu de ces derniers en français, j'ai préféré les introduire au moyen de préambules en français et de garder le reste du texte en anglais. Les matériels et méthodes de chaque article ont été en revanche réécrits en français et se trouvent naturellement inclus dans la section [2](#). La numérotation des figures suivra celle du manuscrit total, sauf dans le cas de l'article soumis.





# 1 Introduction

Les algues, contrairement aux plantes terrestres, ne forment pas un groupe monophylétique avec un ancêtre commun, mais davantage un groupe fonctionnel. On pourrait les définir ainsi : "tout organisme faisant de la photosynthèse oxygénique, à l'exception des embryophytes, est une algue". En réponse à cette définition, on trouve une multitude d'organismes, appartenant à des groupes divers et qui peuvent être évolutivement très distants. Les algues peuvent être des procaryotes (les cyanobactéries), des eukaryotes unicellulaires (diatomées, dinoflagellés, glaucophytes, microalgues vertes et rouges, etc.), des eukaryotes pluricellulaires (les macroalgues). Ce dernier groupe contient trois types d'organismes, cohabitant dans le même habitat, depuis la plus haute zone de l'estran jusqu'à la zone subtidale. Ils sont historiquement caractérisés par leur couleur et cette classification est supportée par les études moléculaires plus récentes. On distingue :

- les algues vertes (Chlorophyta), par exemple du genre *Ulva* et *Caulerpa*,
- les algues brunes (Phaeophyceae), telles que *Fucus*, *Laminaria* et *Ectocarpus*,
- les algues rouges (Rhodophyceae), notamment *Porphyra*, *Gracilaria* et *Chondrus*.

Malgré une certaine ressemblance morphologique, ces trois groupes appartiennent à des radiations bien distinctes et ont ainsi des histoires évolutives divergentes (pour s'orienter, voir la figure 1 d'après Adl [2]).

Mes travaux ont porté sur une algue rouge, dont la position phylogénétique en fait un organisme intéressant du point de vue évolutif. Si l'on s'accorde sur l'hypothèse d'une endosymbiose primaire unique [3], tous les archaeplastides sont issus d'un même événement ayant lieu il y a environ 1,6 milliard d'années [4], au cours duquel une cellule de protiste eukaryote aurait phagocyté un ancêtre des cyanobactéries actuelles, devenant alors le premier organisme à posséder un chloroplaste et acquérant ainsi la capacité d'effectuer la photosynthèse. Les Rhodophyceae sont le groupe frère des glaucophytes, des algues unicellulaires, et de la lignée verte, chloroplastida, qui contient les algues vertes (Chlorophyta) et les végétaux terrestres. Les algues rouges ont divergé des autres archaeplastides il y a 1,5 million d'années [5], pour créer une lignée indépendante et sont les premiers organismes multicellulaires complexes pour lesquels des fossiles ont été retrouvés [6]. Cette position, de lignée ancienne, leur confère des propriétés uniques qui seront décrites dans la suite du manuscrit. Les algues rouges sont aussi considérées comme une des sources d'une endosymbiose secondaire [3] où elles auraient à leur tour été phagocytées

par une cellule eukaryote, pour donner plusieurs lignées, dont celle des straménopiles à laquelle appartiennent les algues brunes.

## 1.1 Le modèle *Chondrus crispus*

Dans les travaux présentés dans ce manuscrit, j'ai utilisé comme modèle *Chondrus crispus* (Stackhouse, J., 1797), communément appelé Mousse d'Irlande ou Pioka. C'est une macroalgue rouge, de l'ordre des Gigartinales, appartenant à la sous-classe des Rhodymeniophyceae (voir la phylogénie des algues rouges en figure 2). C'est un sujet d'étude historique de part son importance écologique, cette espèce étant abondamment représentée dans les zones intertidales et subtidales des côtes rocheuses de l'Atlantique Nord [7]. Elle a longtemps été une de principales sources de carraghénanes, des polysaccharides sulfatés utilisés fréquemment dans l'industrie alimentaire et cosmétique comme gélifiant et épaississant.

*Chondrus crispus* est depuis près de 20 ans un modèle dans le laboratoire à la Station Biologique de Roscoff et a été étudié sous de nombreux aspects. Sa paroi était et est toujours un sujet d'étude majeur, avec d'une part la dissection des voies de biosynthèse des carraghénanes [8], d'autre part, la structure et les propriétés rhéologiques des carraghénanes, étudiées en analysant les mécanismes de dégradation de ces polysaccharides, par des enzymes issues de bactéries dont la paroi des algues est le substrat majoritaire [9].

Des efforts de recherche importants ont aussi porté sur l'étude des stress biotiques et abiotiques, où l'interaction de *C. crispus* avec une algue endophyte a révélé des mécanismes de défense impliquant une production d'espèces réactives de l'oxygène par une enzyme membranaire, la NADPH oxydase [12]. Ceci a mené à approfondir la caractérisation de cette enzyme [13], avec le clonage et le séquençage du gène *Ccrboh* codant pour la NADPH oxydase et des mesures d'expression de *Ccrboh* montrant l'induction de la transcription du gène en réponse à des traitements au méthyl jasmonate (MeJa). Cette molécule est également impliquée dans d'autres réponses au stress, telles que l'induction de l'expression de glutathion-S-transférases et la synthèse d'oxylipines [15]. L'importance du MeJa a conduit à l'étude à plus grande échelle du transcriptome de *C. crispus*, révélant que 6 % des gènes sont régulés en réponse à un traitement au MeJa, dont de nombreux gènes relatifs au stress [14], suggérant que le MeJa peut être une hormone de stress chez *C. crispus*.

Enfin, des analyses transcriptomiques de la réponse de *C. crispus* à une plus grande palette de stress naturels et simulés en laboratoire a suggéré l'importance du métabolisme oxydatif pour

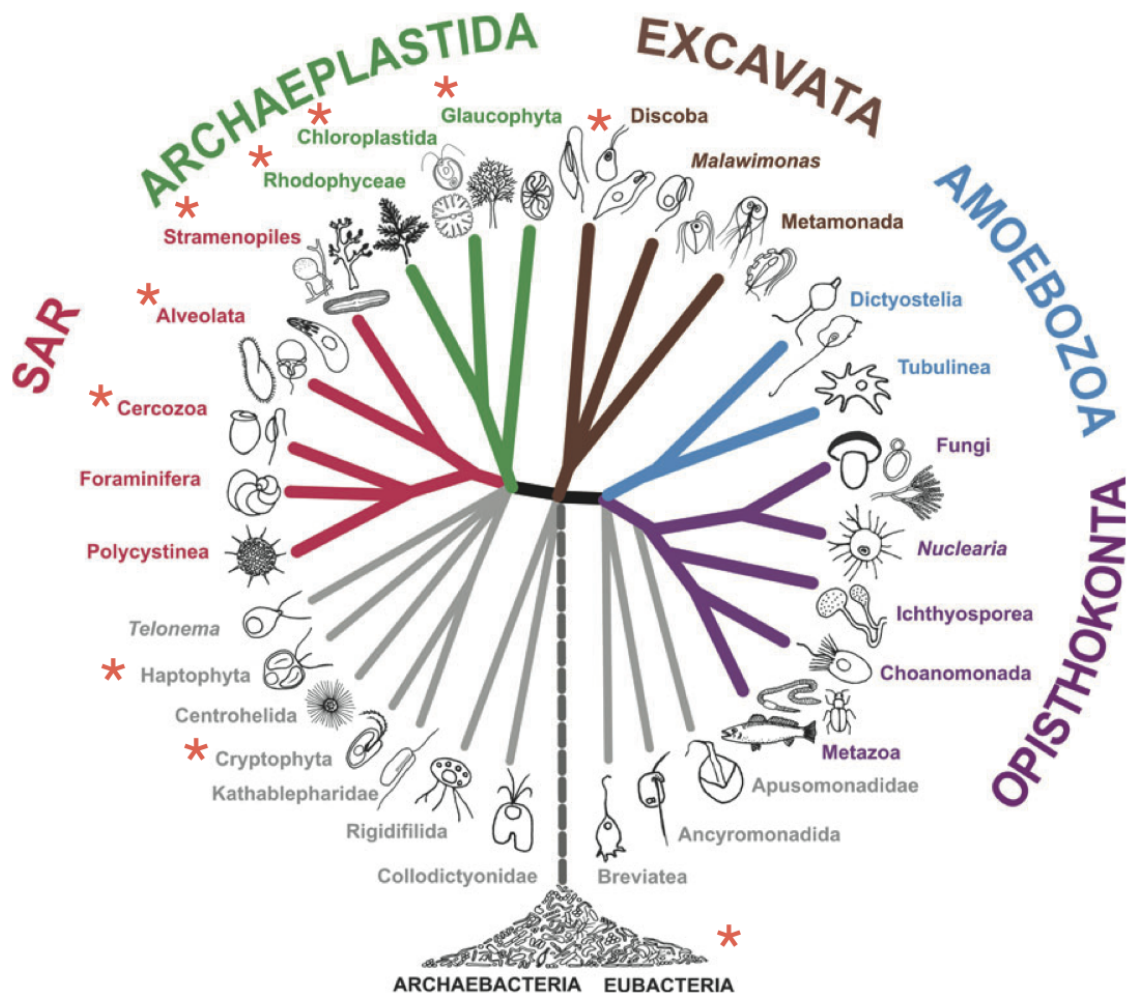


FIGURE 1: Arbre consensus des Eukaryotes (d'après Adl *et al.*, 2012). Les étoiles indiquent les taxa contenant des algues. Seuls trois d'entre eux contiennent des algues multicellulaires : les chloroplastides, rhodophycées et straménopiles. Les algues rouges (rhodophytes) sont le groupe frère des algues vertes (chlorophytes) et des plantes terrestres, au sein du royaume des Archaeplastides. Les algues brunes (straménopiles), haptophytes, cryptophytes, alvéolates, etc., organismes également photosynthétiques, en sont plus éloignées.

A consensus tree of Eukaryotes (adapted from Adl *et al.*, 2012). Stars show taxa containing algae. Only three phyla contain multicellular algae : chloroplastida, rhodophyceae and stramenopiles. Red algae (rhodophyta) are the sister group to green algae and land plants, and all belong to Archaeplastida. Brown algae (stramenopiles), haptophytes, cryptophytes, alveolates, etc., also photosynthetic organisms, are more distantly related.

la tolérance et l'acclimatation à ces stress [16].

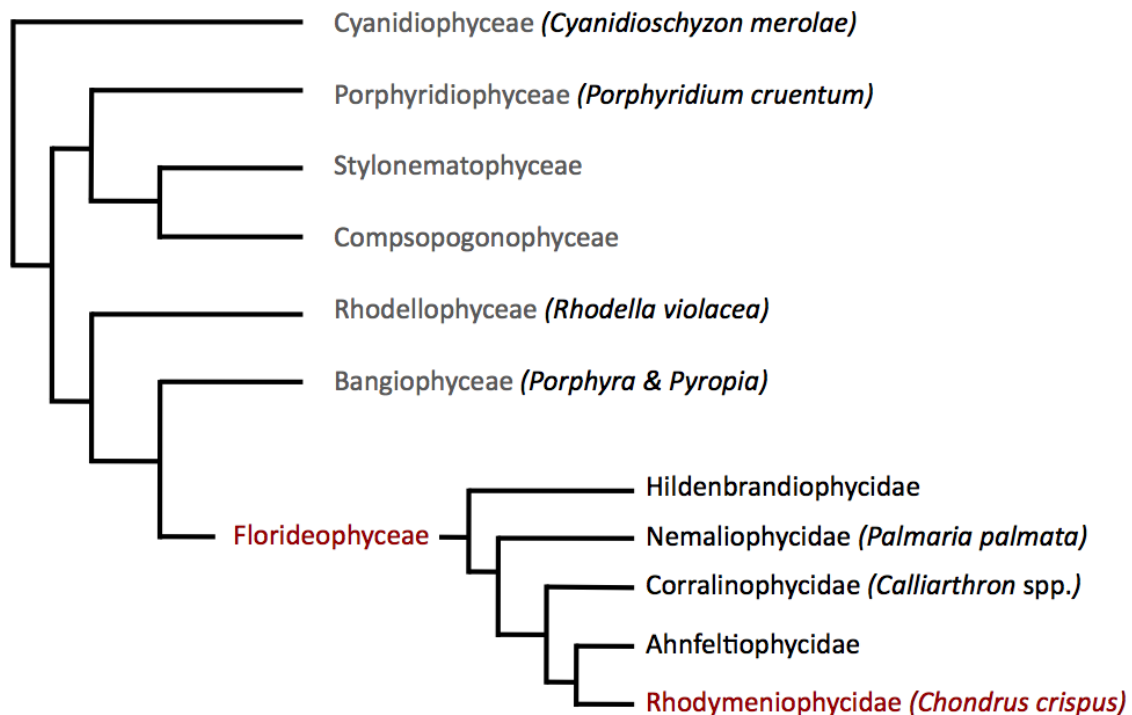


FIGURE 2: Phylogénie simplifiée des algues rouges (Rhodophyta), d'après Verbruggen *et al.*, 2010 [18]. Le nom des genres abordés dans le manuscrit sont précisés entre parenthèses.

Simplified tree of red algae (Rhodophyta) adapted from Verbruggen *et al.*, 2010 [18]. The name of some genera discussed in the manuscript are precised in brackets.

Ces études ont cependant été limitées par le manque d'informations sur les gènes de *C. crispus*. Afin de palier à cette lacune, deux banques de 2000 EST chacune (étiquettes de séquences transcrites) ont été générées à partir de protoplastes et de thalles de *C. crispus* [17] couvrant environ 500 gènes. Et aujourd'hui, en plus d'être un modèle pour les études physiologiques et biochimiques, *C. crispus* est également un modèle en génomique, amorcé avec la thèse de C. Leblanc qui était consacrée au séquençage, à l'annotation et à l'analyse du génome mitochondrial [10], poursuivi plus récemment par le séquençage et l'annotation du génome nucléaire et plastidial et donnera lieu bientôt à une publication. Une partie de la section résultats 3.1 y sera consacrée.

N'existant pas encore de modèle reconnu de macroalgue rouge, le choix de *C. crispus* a été motivé par de nombreux critères. Tout d'abord, c'est une florideophyceae, un groupe monophylétique qui comprend 95 % des algues rouges, contrairement au groupe des bangiophyceae qui contient certes les espèces les plus importantes économiquement, *Porphyra* et *Pyropia*, mais qui

représente moins de 5 % des algues rouges [11].

- c’est une algue rouge typique, permettant d’envisager de transférer les informations obtenues vers d’autres algues rouges,
- l’espèce est bien définie et relativement bien documentée,
- elle est facile à maintenir en laboratoire,
- c’est un organisme multicellulaire complexe,
- son génome est relativement petit (105 Mbp),
- elle est relativement proche d’espèce à importance économique, comme *Euchema* et *Kappaphycus* (deux gigartinales, sources de gélifiant), et de manière plus distante de *Porphyra* et *Pyropia* (nori).

### 1.1.1 Un cycle de vie trigénétique

*Chondrus crispus* est une algue relativement petite, sa taille adulte se situant généralement entre 10 et 15 cm. Elle vit fixée à son substrat rocheux par un petit crampon en forme de disque et son stipe fin est surmonté d’un thalle de texture cartilagineuse, ramifié de manière dichotomique avec des lobes plus ou moins arrondies, donnant un aspect frisé (figure 3). Sa couleur est variable et dépend de son état physiologique [19] : généralement pourpre, il peut aussi être vert, jaune ou brun.

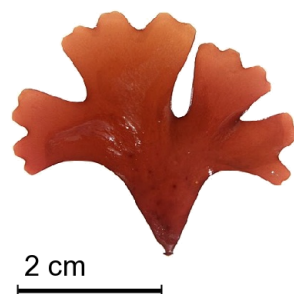


FIGURE 3: Jeune gamétophyte de *Chondrus crispus* (Stackhouse), Gigartinales. Photo par J. Collén.  
Young gametophyte of *Chondrus crispus* (Stackhouse), Gigartinales. Photo by J. Collén.

*C. crispus* a un cycle de vie trigénétique, une spécificité de certaines algues rouges et unique au sein des eukaryotes [20] [21], où se succèdent trois phases distinctes, représentées dans la figure 4 : le tetrasporophyte, diploïde, génère par méiose des tetraspores haploïdes. Celles-ci germent pour donner d’une part des gamétophytes femelles et d’autre part des gamétophytes mâles. Ces derniers produisent des gamètes, haploïdes, qui vont féconder les gamètes femelles

portés par le gamétophyte femelle. Le carposporophyte - structure diploïde d'environ 3 mm, issue de la fécondation, qui se développe sur le gamétophyte femelle - peut à son tour libérer des carpospores qui germeront pour redonner un tetrasporophyte. Les spores et les gamètes des algues rouges ne possèdent pas de flagelles et par conséquent ne peuvent pas nager.

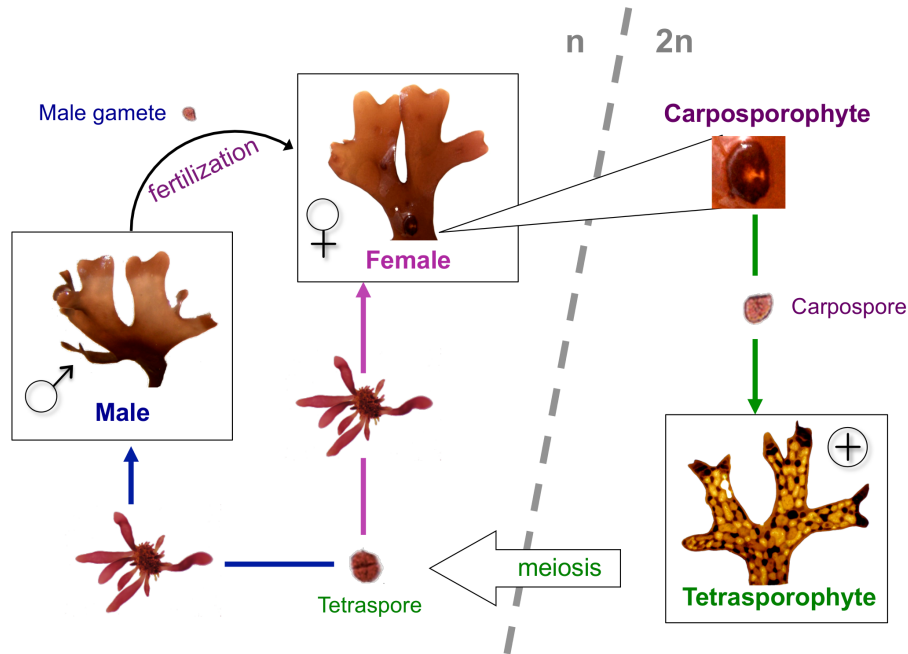


FIGURE 4: Cycle de vie trigénétique de *C. crispus*. Trois générations se succèdent : le tetrasporophyte, diploïde, produit des tetraspores haploïdes par méiose. Les tetraspores germent en gamétophytes mâle ou femelle, qui à leur tour produisent des gamètes. La fusion des gamètes mâle et femelle produit la troisième génération, le carposporophyte, porté par la femelle.

Trigenetic life cycle of *C. crispus*. Three generations follow one another : the diploid tetrasporophyte produces haploid tetraspores by meiosis. Tetraspores germinate into male and female gametophytes. The fusion of the gametes made by gametophytes produces the third generation, the carposporophyte, on the female thallus.

Les deux générations principales, gamétophytes et tetrasporophytes, sont quasiment identiques du point de vue morphologique, lorsqu'ils sont non-fertiles. Il est cependant parfois possible de les distinguer grâce à une iridescence bleue visible (au soleil) sur les gamétophytes [22]. Le tetrasporophyte ne possède pas cette propriété, et lorsqu'il est fertile il est reconnaissable grâce à ses structures reproductives semblables à des tâches de léopard. Les gamétophytes femelles sont également reconnaissables lorsqu'ils sont fertiles, grâce aux carposporophytes qu'ils portent sur le thalle. Les gamétophytes mâles sont plus délicats à identifier, ils ont parfois une bande plus claire aux apex de chaque lobe et sont moins charnus que les femelles. Ces critères sont des pistes d'identification rapides, mais ne sont pas fiables à 100 %. De plus, à l'heure



actuelle, nous ne possédons pas de critère permettant de distinguer un gamétophyte mâle d'un gamétophyte femelle lorsqu'ils ne sont pas fertiles.

Néanmoins, malgré cette forte ressemblance macroscopique, un certain nombre de différences ont été identifiées à l'échelle cellulaire et moléculaire. Par exemple, les deux générations ne possèdent pas les mêmes polysaccharides pariétaux majoritaires. Il a été montré que les gamétophytes contiennent surtout des carraghénanes de type  $\kappa$  et  $\iota$  alors que les tetrasporophytes ont une paroi constituée de carraghénanes de type  $\lambda$  et possiblement  $\theta$  [23].

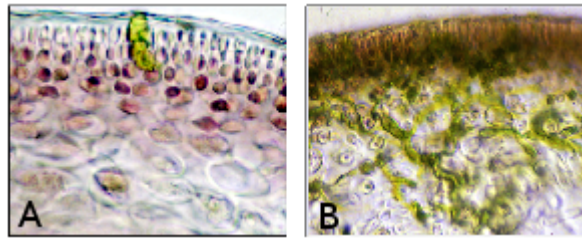


FIGURE 5: Coupe transversale de *C. crispus* infecté par l'endophyte vert *Acrochaete operculata* (J.A. Correa & R. Nielsen). A : le gamétophyte est résistant. B : le tetrasporophyte, par contre, est sensible [12].

Cross-section of *C. crispus* inoculated by the green algal endophyte *Acrochaete operculata* (J.A. Correa & R. Nielsen). A : the gametophyte is resistant. B : the tetrasporophyte is sensitive [12].

Les deux générations présentent une sensibilité différente à la virulence d'une algue verte pathogène, *Acrochaete operculata*. Le tetrasporophyte peut être entièrement envahi (photo 5B) par l'endophyte alors que le gamétophyte peut résister (figure 5A) en empêchant le parasite de se développer. Cette résistance est corrélée à l'émission d'un burst oxydatif de  $H_2O_2$  par une enzyme de type NADPH oxidase chez le gamétophyte. Ce burst étant quasiment nul chez le tetrasporophyte [12].

### 1.1.2 Les carraghénanes

Les carraghénanes sont les composants les plus abondants de la paroi de certaines algues rouges, pouvant représenter jusqu'à 50 % de leur masse sèche. Ce sont des polysaccharides complexes composés d'assemblages de sous-unités galactanes — des dimères de galactose et de 3,6-anhydrogalactose (figure 6) — plus ou moins sulfatés selon le type de carraghénane considéré. Dans l'industrie alimentaire, les types  $\kappa$ ,  $\iota$  et  $\lambda$  sont les plus utilisés, et présentent des propriétés rhéologiques spécifiques selon leur degré de sulfatation : en présence de chlorure de potassium, le  $\kappa$ -carraghénane, mono-sulfaté, permet de former des gels rigides. Le  $\iota$ -carraghénane, bi-sulfaté,

additionné de chlorure de calcium, forme des gels plus souples. Quand au carraghénane de type  $\lambda$ , portant trois fonctions sulfate et ne possédant pas de pont anhydro, ne forme pas de gel, mais permet d'obtenir des solutions hautement visqueuses et ainsi est utilisé comme épaississant [24], [26], [25], [27]. Malgré leurs intérêts biologique et économique, les connaissances concernant les voies de biosynthèse des carraghénanes sont encore limitées.

En Bretagne mais aussi au Canada, *C. crispus* a longtemps été récolté dans le but d'extraire ses carraghénanes. Il a, par la suite, été remplacé par d'autres espèces, telles que les algues tropicales *Kappaphycus alvarezii* et *Eucheuma denticulatum*, cultivées massivement dans des "fermes", notamment à Hawaii et aux Philippines [28]. L'aquaculture a permis d'améliorer le coût de main-d'oeuvre et le rendement, ces algues sont ainsi industriellement plus rentables.

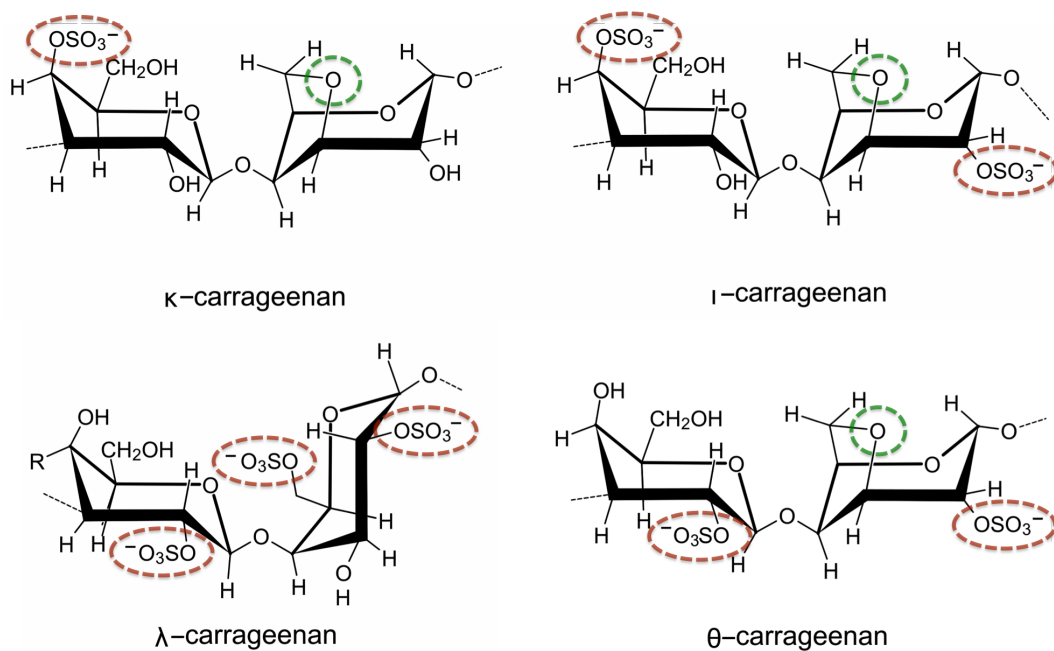


FIGURE 6: Structures de quelques sous-unités des différents types de polysaccharides pariétaux présents chez *C. crispus*. En rouge, sont entourés les groupements sulfate, en vert, le pont 3-6-anhydro. Les pointillés représentent les liaisons entre les disaccharides.

Structures of some of the different types of cell-wall polysaccharide sub-units in *C. crispus*. In red : sulfate groups, in green : 3,6-anhydro bridge. The dotted lines represent the links between disaccharides.

### 1.1.3 *Chondrus crispus* en culture

La culture de *C. crispus* au laboratoire est généralement effectuée dans des bidons d'eau de mer naturelle, sous bullage. Son cycle de vie est long, excédant une année, ce qui rend sa

maîtrise délicate, le maintien des algues se fait par reproduction végétative (bouturage) plutôt que par reproduction sexuée. Néanmoins, un effort de recherche a été axé sur les croisements dans l'équipe de M. Valéro (UMR7144, Station Biologique de Roscoff) dans le but de pouvoir mener des études génétique sur *C. crispus*.

Les études concernant la culture de *C. crispus* sont assez rares, compte tenu du caractère industriel et donc confidentiel de ce type de recherches. Il est cependant possible de trouver quelques publications sur la culture en pleine mer [29], avec des mesures sur l'année du taux de croissance et du contenu en carraghénanes. Le taux de croissance journalier de *C. crispus* a été estimé entre 2 et 4 % et la richesse des algues en carraghénane est la plus forte en été lorsque l'eau de mer est appauvrie en phosphore et azote. Ce dernier point avait également été observé pour des cultures en bac [30].

Mathieson & Prince (1973) [7] ont rapporté que *C. crispus* présente une grande plasticité et est tolérant à une vaste gamme de température et de conditions lumineuses. Sa croissance serait optimale en été, jusqu'à 0,4 mm par jour, et serait d'abord corrélée à la température de l'eau, puis à la luminosité.

D'autres études de culture en laboratoire portait sur l'optimisation de la croissance en fonction de l'intensité lumineuse. Dans un premier cas, la croissance de *C. crispus* était maximale pour une densité de flux de photon de  $20 \mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  avec une photopériode de 16h/8h jour/nuit à 16°C, et l'algue semblait intolérante aux intensités lumineuses plus élevées [31] et une forte mortalité était constatée après un an de culture à plus de  $30 \mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ . Une autre étude a cependant montré que la croissance des algues était optimale entre 60 et 140  $\mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  [32].

Ces valeurs semblent toutes très faibles comparées à l'intensité lumineuse naturelle, qui peut excéder les 2000  $\mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  lorsque le soleil est à son zénith [33]. Il a été d'ailleurs intéressant de constater, que lors de mes expériences portant sur la photosynthèse, les réponses relatives à la régulation lumineuse étaient bien plus prononcées et efficaces lorsque les algues étaient fraîchement récoltées sur le terrain, plutôt que celles mises en culture depuis plusieurs jours (voire plusieurs années dans le cas de la souche Peggy's Cove) à 100  $\mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ .

## 1.2 Introduction à la génomique

Un génome peut être considéré comme un plan contenant toutes les instructions de montage d'un être vivant. On y trouve les informations essentielles sur le contenu moléculaire des cellules, permettant parfois de reconstruire des voies de biosynthèse théoriques, d'identifier des éléments spécifiques d'un cycle de vie, repérer une résistance ou au contraire une sensibilité à des facteurs pathogènes. Etablir un inventaire des gènes présents dans un organisme permet de faciliter la compréhension de son fonctionnement. Il est ainsi possible de comparer des génomes entre eux, mais aussi d'effectuer des approches ciblées plus aisément, comme surexprimer des protéines à partir de la séquence des gènes correspondant dans le but d'une caractérisation biochimique, et de gagner du temps grâce aux approches bioinformatiques en anticipant plus facilement les réponses physiologiques par de la modélisation [34]. Néanmoins, la génomique a ses limites. L'annotation des gènes est effectuée par homologie de séquences et des fonctions leurs seront assignées en fonction des informations déjà présentes dans les bases de données. Si un gène ne possède pas d'homologue connu, il sera impossible de le caractériser juste à partir de sa séquence. De plus, si la bioinformatique permet de prédire des modèles théoriques, ils devront par la suite être vérifiés par des expériences.

Ces quelques dernières années, les techniques de séquençage ont évolué de manière extrêmement rapide avec en parallèle, une baisse drastique des coûts de séquençage et de stockage des données. Cette progression ouvre de nombreuses possibilités pour des études de physiologie et d'écologie, qui avant étaient impossibles pour des raisons de contraintes économiques et techniques. Actuellement, s'il est facile de produire les données, les analyser devient un problème croissant, par manque de moyens humains.

En octobre 2012, plus de 3 700 génomes d'organismes avaient été séquencés, dont 182 eukaryotes, et près de 15 000 projets de séquençage étaient en cours, ces nombres évoluant quotidiennement ([www.genomesonline.org](http://www.genomesonline.org)).

**Des génomes et des algues.** Jusqu'à présent, l'effort de séquençage de génome d'algues était axé sur les unicellulaires. Ainsi, parmi les algues vertes, on peut citer *Chlamydomonas reinhardtii* [36] [37], la "levure" de la biologie végétale, ou *Ostreococcus tauri* [38] qui est le plus petit organisme eukaryote libre décrit et possédant le plus petit génome connu pour un eukaryote photosynthétique. Un seul génome d'algue rouge a été séquencé, celui de *Cyanidioschyzon merolae* [39], une unicellulaire extrémophile, distante des algues rouges florideophytes. Un seul

génom de glaucophyte, celui de *Cyanophora paradoxa*, a été publié à ce jour [40]. Du côté des straménopiles, deux génomes de diatomées ont également été séquencés, celui de *Thalassiosira pseudonana* et *Phaeodactylum tricornutum*.

Les séquences de génome d'algues multicellulaires sont encore rares et seule l'algue brune *Ectocarpus siliculosus* a son génome publié actuellement (en 2011 [35]). Les génomes d'algues rouges *Porphyra umbilicalis* [41] et *P. yezoensis*, sont à notre connaissance encore au stade d'assemblage. La table 1 résume les génomes d'algues eukaryotes publiés et quelques projets en cours.

La lignée des rhodophytes est particulièrement sous représentée, et il manque notamment des connaissances génomiques sur les algues rouges pluricellulaires. Ceci a conduit au séquençage du génome de *C. crispus*. Le projet a été coordonné par notre équipe, en collaboration avec le centre national de séquençage Génomscope (CEA, Évry) ainsi qu'un consortium international d'annotateurs. À ce jour, l'analyse et l'annotation experte du génome sont terminées et la publication est imminente [42]. Les données concernant le génome de *C. crispus* seront détaillées dans la section Résultats 3.1.

### 1.2.1 Un peu d'histoire...

**Le séquençage première génération.** Frederick Sanger a été un des premiers à développer le séquençage d'ADN [44]. Celui-ci est réalisé par synthèse de brin complémentaire, à l'aide d'amorces, d'une DNA polymérase, par l'introduction de nucléotides classiques et de didésoxyribonucléotides (dd-NTP) marqués, permettant d'interrompre la synthèse d'ADN. L'introduction aléatoire de ces dd-NTP "terminateurs" permet d'obtenir des fragments d'ADN plus ou moins longs, qu'il est possible de faire migrer par électrophorèse sur gel afin de déterminer leur taille. La réaction de séquence est effectuée quatre fois, pour chaque base. La séquence d'ADN est lue dans l'ordre des fragments, des plus courts aux plus longs. Plus tard, la migration des fragments d'ADN est effectuée par électrophorèse capillaire. Ceci permet l'automatisation du séquençage et la possibilité de lire des fragments d'ADN de plus en plus longs, jusqu'à quelques 800-900 paires de bases (pb). Cependant, cette technique étant coûteuse aussi bien financièrement qu'en temps, d'autres techniques ont ensuite été développées.

Parallèlement, Maxam et Gilbert [45] ont conçu une technique de séquençage où de courts fragments d'ADN étiquetés sont clivés chimiquement avant d'être séparés sur gel de polyacrylamide. Le principe de lecture est le même que pour Sanger.

TABLE 1: Les génomes d'algues eukaryotes publiés et une sélection de projets en cours (il y en avait plus de 600 en octobre 2012). Species : espèce, Status : statut, Published : publié, Ongoing : en cours. Les étoiles représentent les algues multicellulaires. The genomes of eukaryotic algae that have been published and a selection of ongoing projects (more than 600 in october 2012). Stars show multicellular algae.

Species	Group	Status
<i>Chlamydomonas reinhardtii</i>	Chlorophyta	Published
<i>Micromonas pusilla</i>	Chlorophyta	Published
<i>Ostreococcus 'lucimarinus'</i>	Chlorophyta	Published
<i>Ostreococcus tauri</i>	Chlorophyta	Published
<i>Volvox carteri</i> (*)	Chlorophyta	Published
<i>Chlorella</i> sp. NC64A	Chlorophyta	Published
<i>Cyanophora paradoxa</i>	Glaucophyta	Published
<i>Emiliania huxleyi</i>	Haptophyta	Published
<i>Cyanidioschyzon merolae</i>	Rhodophyta	Published
<i>Thalassiosira pseudonana</i>	Stramenopiles	Published
<i>Phaeodactylum tricornutum</i>	Stramenopiles	Published
<i>Ectocarpus siliculosus</i> *	Stramenopiles	Published
<i>Symbiodinium</i> sp	Alveolata	Ongoing
<i>Chlorarachnion reptans</i>	Cercozoa	Ongoing
<i>Bigelowiella natans</i>	Cercozoa	Ongoing
<i>Botryococcus braunii</i>	Chlorophyta	Ongoing
<i>Coccomyxa</i> sp.	Chlorophyta	Ongoing
<i>Bathycoccus prasinos</i>	Chlorophyta	Ongoing
<i>Chlorella vulgaris</i>	Chlorophyta	Ongoing
<i>Dunaliella salina</i>	Chlorophyta	Ongoing
<i>Chara vulgaris</i> *	Chlorophyta	Ongoing
<i>Ulva rigida</i> *	Chlorophyta	Ongoing
<i>Ulva compressa</i> *	Chlorophyta	Ongoing
<i>Guillardia theta</i>	Cryptophyta	Ongoing
<i>Phaeocystis antarctica</i>	Haptophyta	Ongoing
<i>Phaeocystis globosa</i>	Haptophyta	Ongoing
<i>Porphyra yezoensis</i> *	Rhodophyta	Ongoing
<i>Calliarthron cheilosporioides</i> *	Rhodophyta	Ongoing
<i>Chondrus crispus</i> *	Rhodophyta	Ongoing
<i>Gelidium elegans</i> *	Rhodophyta	Ongoing
<i>Galdieria sulphuraria</i>	Rhodophyta	Ongoing
<i>Porphyra umbilicalis</i> *	Rhodophyta	Ongoing
<i>Bolidomonas pacifica</i>	Stramenopiles	Ongoing
<i>Aureococcus anophagefferens</i>	Stramenopiles	Ongoing
<i>Ochromonas</i>	Stramenopiles	Ongoing
<i>Pseudo-nitzschia</i>	Stramenopiles	Ongoing
<i>Thalassiosira rotula</i>	Stramenopiles	Ongoing
<i>Cyclotella meneghiniana</i>	Stramenopiles	Ongoing
<i>Fragilariopsis cylindrus</i>	Stramenopiles	Ongoing
<i>Chattonella subsalsa</i>	Stramenopiles	Ongoing

**Le séquençage massif haut-débit parallèle.** Les besoins en séquençage ont considérablement augmenté au cours des dernières années, et afin de diminuer les frais et accélérer les réactions, l'effort de recherche au sein des entreprises de biotechnologie a explosé.

Le pyroséquençage est une technique développée en 1996 [46] [47] utilisée par la technologie 454 (Roche Life Science, USA), reposant également sur le principe de "séquençage par synthèse". La première différence avec Sanger se trouve au niveau de la lecture, puisqu'elle est effectuée en temps réel, par détection de la lumière émise par une réaction secondaire initiée par le clivage du pyrophosphate lors de l'incorporation séquentielle de chaque nucléotide. L'intensité du signal est ainsi proportionnelle au nombre de bases consécutives identiques, jusqu'à saturation. Les homopolymères peuvent ainsi être source d'erreur.

L'amplification des fragments d'ADN à séquencer est réalisée par PCR dans des micro-réacteurs — des gouttelettes d'eau au sein d'un bain d'huile — permettant d'effectuer des milliers de réactions et de lectures en parallèle, réduisant drastiquement la durée du séquençage, ainsi que le coût. En revanche, les fragments séquencés sont plus courts que par la méthode de Sanger (400-500 pb).

Actuellement, une autre technique développée par Solexa/Illumina, est très largement utilisée. L'amplification enzymatique, se fait sur des fragments fixés à une lame de verre, dont la densité permet davantage de réactions parallèles qu'en pyroséquençage. La détection se fait par lecture du signal fluorescent de chaque nucléotide incorporé. Les nucléotides sont tous bloqués chimiquement et doivent être débloqués entre chaque réaction par clivage chimique du fluorochrome et du groupe bloquant en 3'. L'incorporation est donc d'une seule base à la fois, ce qui permet de réduire le taux d'erreur de séquençage. Si la rapidité, la fidélité et le faible coût de cette technologie est un avantage considérable, en revanche la longueur des fragments générés est beaucoup plus courtes (32 pb pour les anciennes versions jusqu'à 150 pb actuellement).

Selon la taille et la complexité du génome à séquencer, il est possible d'utiliser une ou plusieurs techniques de manière complémentaire. Les méthodes de Sanger et de pyroséquençage ont ainsi été combinées pour l'étude métagénomique des fèces humain [48], et l'assemblage du génome d'un champignon filamenteux a été achevé à l'aide de trois technologies : Sanger, 454 et Illumina [49]. Pour le génome de *C. crispus*, ces trois technologies ont également été utilisées : Sanger pour le séquençage proprement dit, 454 et Illumina pour l'annotation.

### 1.2.2 La transcriptomique quantitative

L'utilisation du séquençage n'est pas limitée aux seuls génomes, il peut aussi être appliqué aux ARNs (en particulier les ARN messagers). Il est alors nécessaire de rétrotranscrire les ARNs en ADN complémentaires (ADNc) et ce sont ces derniers qui seront séquencés. Les objectifs sont variés :

- générer des ESTs (étiquettes de séquences transcrites) pour identifier les gènes transcrits,
- identifier des transcrits à épissage alternatif,
- quantifier l'expression des gènes.

La méthode la plus couramment utilisée jusqu'à présent, pour quantifier à haut débit l'expression des gènes, était l'hybridation sur puces à ADN. Celle-ci est progressivement remplacée par le *RNAseq*, séquençage massif des ARNs, plus simple à mettre en oeuvre, compétitif et de plus en plus exhaustif [50]. La transcriptomique permet d'avoir une idée de l'expression des gènes au moment de l'échantillonnage, à un stade de culture particulier, dans un tissu ou dans une condition physiologique particulière. Elle peut être qualitative ou quantitative et reflète alors l'équilibre existant entre transcription et dégradation de chaque ARN [51]. Ce n'est pas une estimation absolue de la synthèse d'ARN. Il est important de préciser également que si la transcriptomique quantitative reflète effectivement le contenu en ARN des cellules, il existe des niveaux de régulation supplémentaires notamment au niveau traductionnel et post-traductionnel, ce qui ne rend pas forcément compte de la proportion en protéines, ni des activités enzymatiques. Pour une étude intégrative complète et comparative, il peut être intéressant d'aborder des approches protéomiques voire métabolomiques.

### 1.2.3 La génomique en milieu naturel et la métagénomique

Les méthodes précédentes visaient à étudier une espèce particulière, bien identifiée et de préférence mise en culture. Le *RNAseq* a été utilisé pour suivre des effets physiologiques dans des organismes, par exemple étudier les effets de l'augmentation du niveau de CO<sub>2</sub> sur les coraux [52], mais aussi le développement des baies chez la vigne [53]. Cette méthode permet d'identifier des gènes importants pour des processus physiologiques spécifiques, afin d'étudier ces gènes par la suite.

Il est possible aussi de s'intéresser à des communautés plus larges, par exemple la microfaune et flore d'un sol [54] ou d'un échantillon d'eau de mer [55], dont le contenu exhaustif en organismes est généralement inconnu ou largement sous-estimé et où l'on ignore les interactions entre les



êtres vivants qui le composent. Il est souvent difficile de mettre en culture certaines bactéries ou protistes, dont le cycle de vie peut être long ou lorsque ces organismes ont des exigences en matière de milieu que l'on ne maîtrise pas. Il est alors possible de faire du séquençage direct des acides nucléiques contenus dans un échantillon complexe, puis de récupérer toutes les séquences et tenter d'y repérer des marqueurs. La puissance des techniques actuelles permettent de retrouver des séquences d'organismes rares, que l'on n'aurait probablement jamais réussi à isoler en culture. La métagénomique, consistant à séquencer la totalité des génomes d'un échantillon, permet d'identifier, de manière la plus exhaustive qui soit, les individus en présence. La metatranscriptomique, où l'on séquence l'ensemble des transcriptomes, apporte une couche supplémentaire d'informations sur le fonctionnement de la communauté, permettant des études d'écophysiologie à grande échelle (voir Handelsman, 2004, pour revue [56]).

Durant ma thèse j'ai souhaité utiliser le RNAseq pour étudier un niveau se trouvant à l'interface entre l'étude physiologique pure d'un organisme unique cultivé dans un environnement artificiel au laboratoire et les études environnementales de métagénomique/transcriptomique visant à étudier de vastes complexes de (micro)organismes. Le concept a donc été de suivre l'expression des gènes dans la nature, sur une espèce fraîchement collectée sur le terrain, avec toute leur diversité génétique, des différentes histoires de vie, un rythme diurnal normal et qui a vécu avec tous ses organismes associés. A ma connaissance, ce genre d'approche a été assez peu abordée en utilisant des méthodes de grande échelle comme le RNAseq. Les études semblables les plus récentes, portant sur la moule [57] et le saumon [58], utilisent des puces à ADN, mais renseignent sur les contraintes générées par l'utilisation d'échantillons environnementaux.

### 1.3 Le stress abiotique

Le "stress" dans le langage commun représente la conséquence d'une contrainte sur l'organisme qui la subit, sa réponse à cette contrainte ainsi que la contrainte elle-même, c'est à dire le facteur de stress. Davison & Pearson [59] précisent d'ailleurs la différence entre un stress "limitant", par exemple un apport insuffisant de ressources, et un stress "perturbateur" conduisant à une modification de l'homéostasie cellulaire voire des dommages sur l'organisme et nécessitant un réajustement (à court terme) ou une adaptation (à plus long terme). Dans ce manuscrit, j'utiliserai le terme "stress" essentiellement pour parler du facteur extérieur contraignant causant une modification de l'homéostasie de la cellule voire des dommages potentiels : le stress perturbateur. Dans le cas du stress "oxydatif", il s'agira donc aussi de la contrainte imposée par

la production d'espèces réactives de l'oxygène.

### 1.3.1 Le stress abiotique chez les plantes terrestres

Le stress abiotique est une thématique importante et largement étudiée chez les plantes d'intérêt agronomique car les conditions environnementales influencent le rendement des récoltes et ont par conséquent des enjeux économiques. Les études fondamentales sur les plantes terrestres sont également nombreuses, dans le but de mieux caractériser les capacités d'adaptation des végétaux à leur environnement et entrent dans le cas d'études plus globales de prédiction de scénarii face aux changements climatiques. Il existe ainsi des revues traitant d'analyses globales des réponses au stress [61], ainsi que des études plus ciblées comme l'identification de marqueurs de tolérance au stress salin chez le pois [62], ou encore grâce aux outils disponibles chez *Arabidopsis thaliana* : les connaissances génomiques, la possibilité de faire de nombreux mutants, permettent de déchiffrer de manière très précises les voies de signalisation de réponse aux stress, par exemple la caractérisation d'un nouvel acteur dans les voies de signalisation communes au stress froid et au stress osmotique [63]. Enfin, des études au niveau écologique permettent de relier l'invasivité d'une plante à sa résistance au stress, avec l'exemple de la tolérance au stress hydrique de l'acacia [64].

**Perception, signalisation, réponse.** Les facteurs de stress majeurs chez les plantes, responsables des plus grandes pertes de rendement, sont le froid, la déshydratation et la salinité. L'adaptation des plantes à ces stress est contrôlée par des cascades de réseaux moléculaires qui activent des mécanismes de réponse permettant de ré-établir l'homéostasie et de protéger et de réparer les protéines et membranes endommagées [65].

L'impact du stress peut être découpé en trois grandes catégories, schématisées en figure 7 :

- La **perception**. Elle se situe généralement au niveau de la membrane plasmique. Une perte de turgescence de la cellule causée par un stress hydrique ou salin a pour conséquence une modification de la tension membranaire et un changement de température induit une modification de la fluidité des lipides membranaires. Ces stress sont également perçus par des récepteurs-kinases ou des canaux calciques stretch-dépendants.
- La **transduction du signal**. Après perception d'un stress, des cascades de signalisation sont déclenchées de manière à conduire le signal perçu au niveau de la matrice extracellulaire jusqu'au noyau. Les acteurs de la signalisation impliquent des messagers secon-

daïres, incluant les espèces réactives de l'oxygène (ROS), les ions calcium, l'inositol phosphate, et l'hormone de stress par excellence : l'acide abscissique (ABA). Le signal est également transmis le long de cascades de phosphorylation par des protéines kinases activées par des mitogènes (MAPK) et des kinases calcium dépendantes (CDPK).

- L'**activation de gènes**. De nombreux facteurs de transcriptions, incluant des facteurs à doigt de zinc et des leucine-zipper, régulent les éléments de réponse au stress.

Il existe une voie spécifique du stress froid, comme la cascade ACE-CBF activant des gènes de réponse au froid (COR : cold responsive genes) [66]. Une autre voie est déclenchée cas de stress salin, impliquant les gènes SOS (salt overly sensitive) [67].

Une majorité de cascades de signalisation sont cependant communes à plusieurs types de stress, notamment les voies ABA-dépendantes, impliquant d'autres seconds messagers comme les ROS ou l'inositol-phosphate, activant des facteurs de transcription de type Myc/Myb, mais aussi des voies ABA-indépendantes où interviennent les CDPK ou l'élément de réponse à la sécheresse (DRE : drought response element) [68].

**Les éléments de réponse au stress.** la réponse au stress est complexe et implique plusieurs mécanismes, qui sont souvent communs à plusieurs types de stress. Selon Vinocur & Altman, 2005 [65], elle se décompose en quatre parties :

- L'activation de **transporteurs** d'ions et d'eau. Dans le cas d'un stress salin, des pompes  $H^+$ -ATPases créent un gradient de proton permettant de faire fonctionner des antiports  $Na^+$ -  $H^+$  afin d'abaisser la concentration cellulaire en  $Na^+$ , toxiques pour une majorité de plantes.
- la synthèse d'**osmolytes** compatibles. Afin de rééquilibrer la pression osmotique, en cas de sécheresse, de stress salin ou de stress froid, la cellule végétale produit de la glycine-bétaine et de la proline [71] ainsi que du mannitol et sorbitol [70]. Ces molécules ont également un rôle d'osmoprotectant en aidant à maintenir les structure membranaires.
- la synthèse de protéines **chaperones**. Elles contribuent au bon repliement tridimensionnel des protéines, et jouent ainsi un rôle majeur en cas de stress résultant dans la déstructuration des protéines, tel que le stress salin, la chaleur ou encore la dessiccation [72] (développé dans la section 3.1.2).
- l'activation de systèmes de **détoxication**. La production de ROS étant souvent la conséquence d'un stress abiotique, la cellule végétale possède de nombreux mécanismes de détoxi-

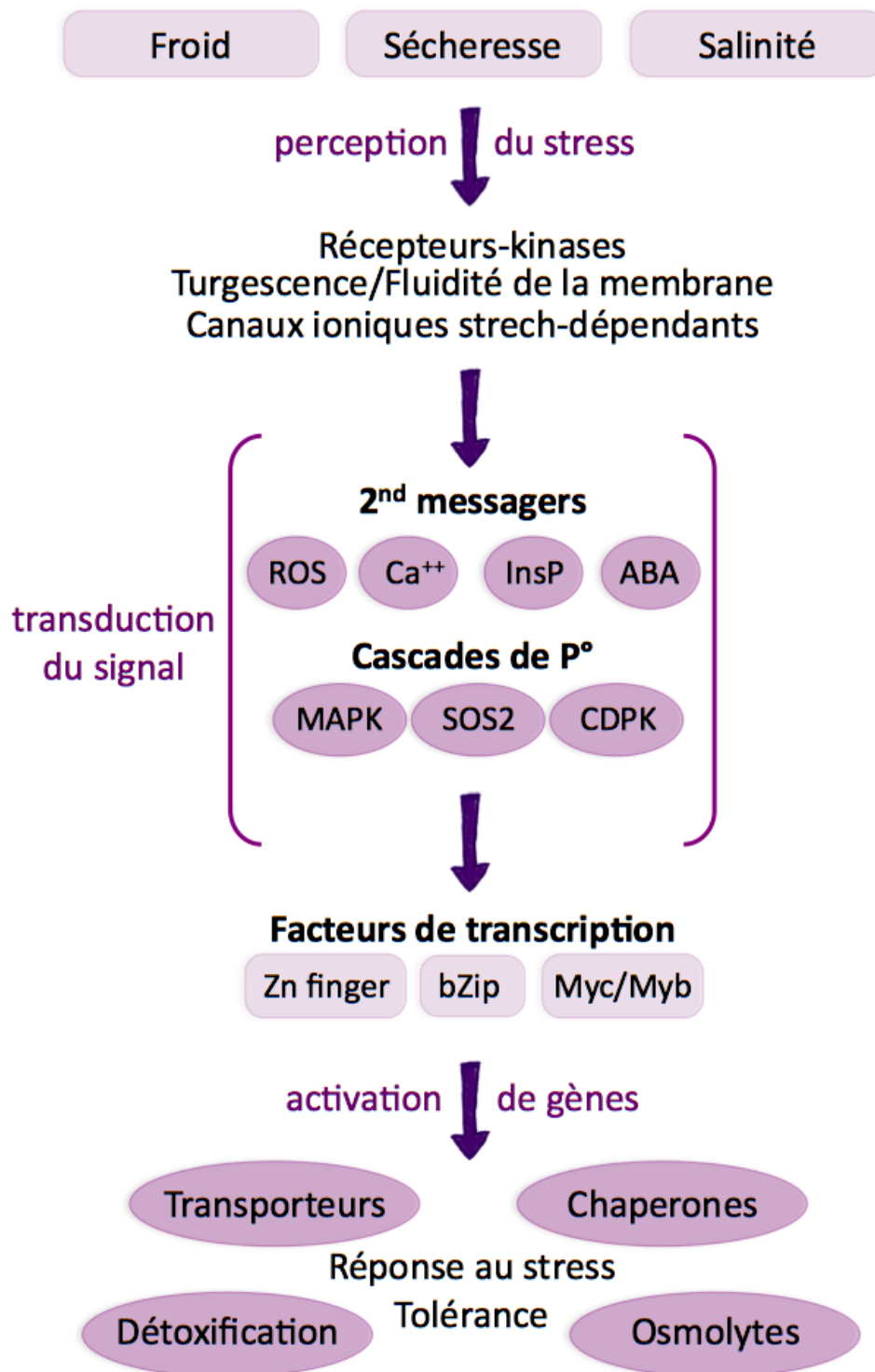


FIGURE 7: Exemple de mécanismes de réponse aux stress abiotiques chez les plantes terrestres. Adapté d'après Vinocur & Altman, 2005 et Huang *et al.*, 2012 [65][66]. ROS : espèces réactives de l'oxygène, Ins5P : inositol-5-phosphate, ABA : acide abscissique, P° : phosphorylation, MAPK : mitogen-activated protein kinase, SOS : salt overly sensitive, CDPK : calcium-dependant protein kinase, Zn finger : doigt de zinc, bZip : basic leucine zipper. Example of abiotic stress response mechanisms in land plants. Adapted from Vinocur & Altman, 2005 and Huang *et al.*, 2012 [65][66]. ROS : reactive oxygen species, ABA : abscisic acid, Zn : zinc.

cation impliquant des enzymes de capture de ROS et des antioxydants, qui seront décrits plus en détails dans la section [1.3.3](#).

### 1.3.2 La marée : source de stress

En considérant uniquement les organismes photosynthétiques, la lignée verte est quasiment le seul taxon ayant colonisé le milieu terrestre. Alors que le milieu marin, dont sont issus les premiers végétaux, présente une beaucoup plus grande diversité. Il est peuplé aussi bien d'organismes de la lignée verte (embryophytes et chlorophytes), mais aussi des lignées bien plus distantes, incluant les algues rouges, les straménopiles, et de nombreux autres organismes unicellulaires procaryotes (cyanobactéries) et eukaryotes (haptophytes, alveolates, cryptophytes) dont beaucoup forment le phytoplancton.

Les algues intertidales sont des organismes aquatiques, mais leur milieu de vie, l'estran, leur impose de vivre en alternance, deux fois par jour, totalement immergée et plus ou moins émergée. Selon la hauteur à laquelle elles vivent, le rapport entre le temps passé à l'état immergé et émergé est variable et définit en partie les niches écologiques [\[59\]](#) [\[73\]](#). Ainsi certaines algues, telle que l'algue brune *Pelvetia caniculata* colonise le haut de l'estran (étage supralittoral) et n'est immergée qu'en pleine mer, à l'inverse, les laminaires (d'autres algues brunes), qui peuplent l'étage infra-littoral, sont immergées la majorité du temps et ne sont exposées à l'air qu'en période de basse mer de vives-eaux. *C. crispus* vit entre les deux, dans l'étage medio-littoral et subit l'alternance des marée quelque soit le coefficient de marée.

La figure [8](#) montre une marée basse et une marée haute de vives-eaux à Rosko, avec un marnage (amplitude de la hauteur d'eau) de 8 m pour le jour où ont été prises les photos.

L'estran représente donc un environnement complexe qui comporte les avantages et les inconvénients à la fois du milieu terrestre et du milieu marin, et offre un terrain pertinent pour étudier les acclimations et les adaptations au stress abiotique. Les variations imposées par l'alternance des marée additionnées aux conditions météorologiques font de la zone intertidale un environnement très dynamique notamment en matière de régime lumineux, de température, de salinité et d'humidité [\[74\]](#). Pour vivre dans cet environnement complexe, les algues ont mis en place des mécanismes pour vivre et prospérer aussi bien à l'air que dans l'eau, à tous les niveaux de l'estran. Connaître ces mécanismes peut ainsi permettre de mieux comprendre l'écosystème intertidal. Les parties suivantes seront consacrées au détail de ces facteurs de stress.



FIGURE 8: Un jour comme les autres à Rosko. A. marée basse, 14h. B. marée haute, 20h.  
A normal day in Rosko. A. low tide, 2pm. B. high tide, 8pm.

**La lumière** L'eau de mer se comporte comme un filtre sélectif de longueurs d'onde [74]. La transmission de la lumière est variable en fonction de la turbidité et la hauteur de la colonne d'eau, le rouge est absorbé en surface, pour ne laisser passer que le vert et le jaune plus en profondeur. Dans un environnement sujet aux marées, les variations lumineuses atteignent des extrêmes : à marée haute, sous un ciel couvert et une forte turbidité, les algues recevront très peu de lumière, à l'inverse à marée basse, en plein soleil, l'intensité lumineuse sera maximale (plus de  $2000 \mu\text{mol} \cdot \text{photons} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  [33]). Ces variations de qualité et de quantité lumineuse influencent directement les processus photosynthétiques.

Le rayonnement ultraviolet (UV) est nocif et est responsable de nombreux dégâts sur les macromolécules biologiques, notamment des lésions et mutations de l'ADN et dégradation de protéines [84][83] et endommage l'appareil photosynthétique [85]. L'eau est un filtre à UV efficace, en particulier dans les zones côtières de l'océan Nord Atlantique [86], où une eau turbide est capable d'absorber 90 % du rayonnement UV-A (mesures à 380 nm) à partir de 20 cm de profondeur, et les rayons UV-B (mesures à 310 nm) dès 6 cm. Le stress UV à marée haute est donc minimal. Mais lorsque la hauteur d'eau est faible, les algues possèdent différentes stratégies pour se protéger des rayonnements UV. Les algues brunes produisent des phlorotanins, les algues vertes possèdent des phenylpropanoïdes, certaines microalgues et algues rouges synthétisent des acides aminés similaires à la mycosporine [88] et de nombreuses cyanobactéries synthétisent de la scytonémine [87].

Le soleil a également une influence directe sur l'augmentation de la température de l'air et de l'eau.

**La température** Au large, l'océan a une température relativement stable au cours de la journée, et les variations saisonnières moyennes sont estimées à 5 °C dans les latitudes moyennes et sont encore moindre (2 °C) aux pôles et entre les tropiques [74]. À l'inverse, dans les cuvettes de basse mer, voire lors de l'émersion totale, la dynamique de température est bien plus importante et est conditionnée par la température de l'air. Dans certaines régions, les algues peuvent geler en hiver et à l'inverse, en été, elles seront exposées à des fortes chaleurs [59]. L'élévation de la température a un impact direct sur la vitesse des réactions métaboliques qui seront alors plus rapides, mais également sur la solubilité des gaz. Ainsi une conséquence de l'augmentation de température est l'accélération de la respiration combinée à une moins bonne solubilité de l'O<sub>2</sub> dans l'eau, rendant le milieu potentiellement anoxique. Le froid, à l'inverse, ralentit les réactions métaboliques et le gel, bien plus destructeur, altère la structure des membranes lipidiques [89] et la formation de cristaux de glace et peut induire la lyse cellulaire [90]. Afin de prévenir les dommages causés par ces variations de température, certains organismes synthétisent des cryoprotectants [91] en réponse au froid (par exemple le raffinose chez *A. thaliana* [92], et le mannitol chez l'algue rouge *Mastocarpus stellatus* [93]), et à l'inverse des protéines chaperones de type *heat shock protein* pour stabiliser les autres protéines risquant de se dénaturer [94].

**La salinité** La salinité de l'eau de mer océanique est en moyenne de 35 g.L<sup>-1</sup>. Dans les cuvettes de marée basse, elle peut varier de manière importante, résultant de la combinaison des facteurs climatiques, incluant le vent, la température, la pluie et l'humidité de l'air. En cas de pluie, la concentration en sel diminuera et à l'inverse, en cas de forte chaleur, l'évaporation augmentera la salinité de l'eau de mer. Les conséquences de ces fluctuations seront d'abord d'ordre physique. Lorsque la pression osmotique du milieu extérieur varie, des mouvements d'eau entre l'intérieur et l'extérieur de la cellule vont modifier la turgescence de cette dernière. La cellule peut notamment plasmolyser dans des conditions de forte hypersalinité. La létalité de la plasmolyse chez les algues rouges a été décrite chez Hoefler en 1930 [75] et revue par Reed en 1990 [76]. Additionné à cela, la concentration en ions de l'eau de mer est également modifiée, provoquant des désordres de nature chimique. La solubilité de l'oxygène et du dioxyde de carbone change, ce qui influence le pH. L'équilibre ionique de la cellule, les gradients, les conformations des protéines et transporteurs sont perturbés, provoquant des dysfonctionnements métaboliques notamment au niveau de la vitesse de photosynthèse. Le maintien de l'homéostasie a un coût énergétique, et en réponse aux changements de salinité, les cellules peuvent compenser le déséquilibre par



le transport actif des ions (accumulation de NaCl chez *E. siliculosus* [145] ou par la synthèse d'osmolytes : polyols (mannitol chez certaines algues brunes [77] et rouges [78]) ou acides aminés (proline chez *E. siliculosus* [145], dans les algues vertes [79] et plantes [80]).

**La dessiccation** Lorsque l'eau se retire à marée basse, certaines algues restent émergées pendant plusieurs heures et peuvent alors se déshydrater très rapidement. La vitesse d'évaporation de l'eau est d'autant plus rapide lorsque la température de l'air est élevée et selon la force du vent, d'autant plus que contrairement aux plantes terrestres, les algues n'ont pas de cuticule imperméable. Les désordres métaboliques résultant de la dessiccation sont nombreux, l'eau est en effet nécessaire à la majorité des réactions biochimiques, ainsi qu'au maintien de l'intégrité des structures cellulaires et impacte fortement la croissance des végétaux [81]. La perte d'eau intracellulaire cause aussi des effets similaires à ceux d'une augmentation de salinité, puisqu'elle résulte en l'augmentation de la concentration des ions dans les cellules.

Une autre conséquence de la dessiccation est le ralentissement brutal de l'activité photosynthétique causé par la diminution de la disponibilité en carbone inorganique. Dans l'air, le carbone inorganique existe sous forme de CO<sub>2</sub>, alors que dans l'eau il est présent également sous forme de HCO<sub>3</sub><sup>-</sup> (d'ailleurs majoritaire dans l'eau de mer, dont le pH se situe autour de 8,1) que les algues peuvent utiliser en tant que source de carbone [74], de manière préférentielle chez certaines, notamment *Chondrus crispus* [96][97]. En absence d'eau, les algues ne peuvent utiliser que le CO<sub>2</sub> dont la concentration dans l'air est beaucoup plus faible que la concentration totale de carbone inorganique dans l'eau.

### 1.3.3 Le stress oxydatif : résultat de facteurs multiples

Dans la partie précédente, nous avons vu que les variations des conditions environnementales peuvent affecter le fonctionnement général de la cellule. La conséquence de tous ces stress est un dysfonctionnement du métabolisme de l'algue et plus particulièrement de la photosynthèse. La fixation du CO<sub>2</sub> et le fonctionnement du photosystème II (complexe protéique de l'appareil photosynthétique) étant très sensibles aux situations de stress [98], tous ces événements conduisent à un ralentissement du transport d'électrons, dont l'excès participerait à réduire l'oxygène moléculaire et serait la cause d'une production accrue d'espèces réactives de l'oxygène, les **ROS** [99][100]. Ainsi, un stress environnemental conduirait à un stress oxydatif.



**Nature des ROS.** Lors du fonctionnement normal des cellules, les ROS sont produites en continu comme sous-produits des métabolismes producteurs d'énergie : la respiration [102] et la photosynthèse [103], mais également par d'autres compartiments cellulaires incluant les péroxysomes et le reticulum endoplasmique (pour revue [101]). Il existe plusieurs types de ROS, par ordre d'état de réduction [104] :

- l'oxygène singulet,  $^1\text{O}_2$ , qui résulte notamment du transfert d'un exciton issu de la chlorophylle lors de réactions photochimiques,
- l'ion radical superoxide  $\text{O}_2^-$ , correspondant à la première forme réduite de l'oxygène moléculaire,
- le peroxyde d'hydrogène  $\text{H}_2\text{O}_2$ , qui peut résulter de la dismutation spontanée ou enzymatique de l'ion superoxide ou directement de la réduction divalente de l'oxygène. Il est plus stable et moins réactif que  $\text{O}_2^-$ .
- le radical hydroxyl  $\text{HO}\cdot$ , extrêmement réactif et capable d'initier des réactions radicalaires en chaîne. Il peut être issu directement des formes précédentes par réduction supplémentaire, mais il est surtout produit en présence de  $\text{H}_2\text{O}_2$  et  $\text{O}_2^-$  dans une réaction catalysée par des métaux (réaction de Fenton).

Enfin, l'état le plus réduit de l'oxygène moléculaire  $^3\text{O}_2$  est l'eau  $\text{H}_2\text{O}$ .

**Les producteurs de ROS.** La production mitochondriale des ROS résulte de la fuite d'électrons de la chaîne respiratoire et leur fixation sur l'oxygène moléculaire  $\text{O}_2$ , conduisant à sa réduction incomplète et la production d'ions superoxide  $\text{O}_2^-$  et de peroxyde d'hydrogène  $\text{H}_2\text{O}_2$  [102]. Cependant, dans les organismes photosynthétiques, la mitochondrie est une source de ROS minoritaire comparée au chloroplaste [103]. Une première source de ROS se situe au niveau du centre réactionnel du photosystème II (PSII) : lorsque la chlorophylle absorbe un photon et se retrouve à l'état de triplet excité,  $\text{Chl}^*$ , elle peut réagir avec l' $\text{O}_2$  pour former de l'oxygène singulet,  $^1\text{O}_2$ , à fort pouvoir oxydant. Le photosystème I (PSI) est également un producteur important de ROS. Lorsque le flux d'électrons est trop important par rapport à la quantité d'accepteur  $\text{NADP}^+$ , il est redirigé vers la ferrédoxine, qui utilise l'oxygène moléculaire comme accepteur d'électrons, pour former des ions superoxide,  $\text{O}_2^-$  (voir figure 10). En cas de carence en  $\text{CO}_2$ , la RubisCO est susceptible de fixer l'oxygène, et mener à la production de  $\text{H}_2\text{O}_2$  par photorespiration.

En observant au microscope un thalle de *Chondrus crispus* préilluminé pendant 10 min et

incubé dans du DCF (2',7'-dichlorofluoresceine) — une sonde qui devient fluorescente lorsqu'elle réagit avec le  $H_2O_2$  — j'ai pu constater la colocalisation de la fluorescence du DCF avec celle de la chlorophylle, montrant que la production des ROS était située principalement au niveau des chloroplastes dans *Chondrus crispus* (figure 9).

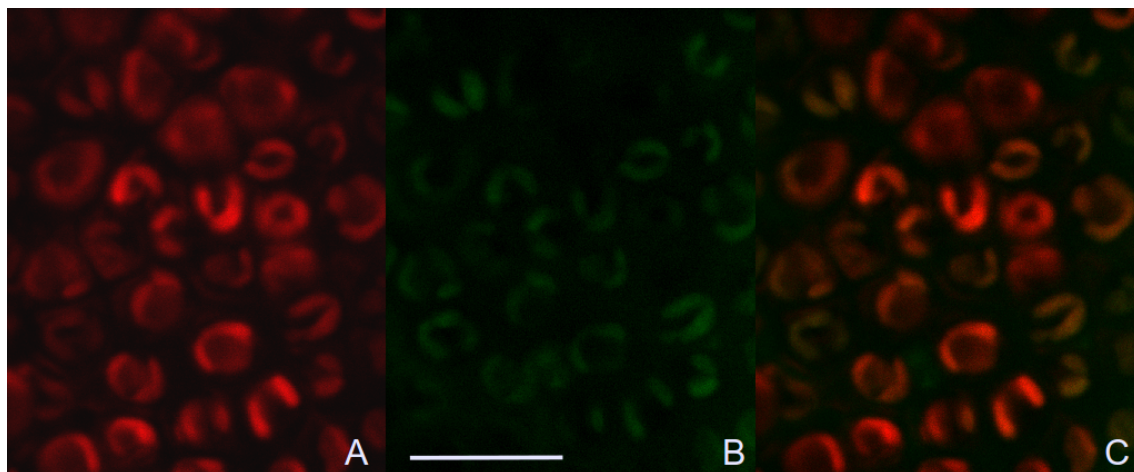


FIGURE 9: Thalle de *C. crispus* après 10 minutes d'exposition à une lumière de forte intensité. A. autofluorescence de la chlorophylle, B. fluorescence du DCF, C. superposition de A et B. Longueur d'onde d'excitation 488 nm. Echelle : 10  $\mu$ m. Le DCF colocalise avec les chloroplastes.

Thallus of *C. crispus* after 10 minutes of exposition under strong intensity light. A. chlorophyll autofluorescence, B. DCF fluorescence, C. merging of A and B. Excitation wavelength 488 nm. Scale : 10  $\mu$ m. The DCF colocalizes with the chloroplasts.

**Les ROS en tant que signal.** Le rôle des ROS peut être utile pour le fonctionnement cellulaire [101], notamment en tant qu'acteur direct ou indirect dans la signalisation où ils ont un rôle activateur de voies de MAP kinases et de facteurs de transcriptions, soit en oxydant directement les composants de ces voies de transduction et modifiant ainsi leur activité, ou en étant perçus d'abord par des récepteurs, qui déclencheront en aval des cascades de signalisation [105]. Au final, les ROS sont susceptibles d'avoir un impact régulateur sur l'expression des gènes [105]. Certaines familles de gènes sont connus pour être activées par les ROS chez les plantes. Ils sont également des composants majeurs de la réponse à des attaques pathogènes aussi bien chez les plantes [106] que chez les algues rouges [13]. Un burst oxydatif est généré par des enzymes membranaires de type NADPH oxydase et certaines peroxydases en cas de stress d'origine biotique. Les ions superoxide et le peroxyde d'hydrogène libérés agissent à la fois comme molécules toxiques pour le pathogène et signal de l'attaque pour la cellule et les cellules voisines.

**Toxicité des ROS.** Par leur nature radicalaire ou énergisée, les ROS sont instables et réactives. Elles sont ainsi à l'origine de nombreux dommages sur les composants moléculaires de la cellule, aboutissant à un dysfonctionnement de l'organisme. Selon l'espèce de ROS considérée, les dommages ne sont d'ailleurs pas les mêmes [104] :

- en oxydant l'ADN, elles occasionnent des modifications des bases azotées et sont par conséquent facteurs de mutations, notamment au niveau de l'ADN mitochondrial et chloroplastique, proches des sources de ROS. Le radical hydroxyl est le plus réactif, l'oxygène singulet attaque surtout la guanine, quant au peroxyde d'hydrogène et ion superoxide, ils ne réagissent pas directement avec l'ADN.
- en carbonylant les acides aminées, elles conduisent à la déstabilisation des structures secondaires et tertiaires des protéines et l'inactivation des enzymes. Les protéines altérées sont alors condamnées à être dégradées par des protéases.  $\text{HO}\cdot$  et  $^1\text{O}_2$  sont aussi capables d'oxyder les groupements thiols des cystéines et méthiones.
- en oxydant les sucres,  $\text{HO}\cdot$  est susceptible d'altérer la structure des polysaccharides de paroi.
- en provoquant une réaction en chaîne de peroxydation des lipides, notamment par le radical  $\text{HO}\cdot$ , les structures membranaires sont désorganisées et fragilisées, impactant le fonctionnement des protéines et enzymes transmembranaires.

Les nombreuses conséquences de la toxicité des ROS nécessitent de réguler efficacement le contenu cellulaire en ROS afin d'éviter leur cytoléthalité [107] tout en leur permettant de participer à la signalisation cellulaire.

**Détoxication.** Pour se protéger de ces effets délétères des ROS, les cellules possèdent des systèmes de réparation des composants cellulaires (correction de l'ADN, réparation des nucléotides, resynthèse et réparation des protéines et lipides). Afin d'éviter que la production de ROS soit excessive par rapport aux capacités de réparation, la cellule met en œuvre une batterie de systèmes de détoxication [105] comprenant des enzymes, les ROSEs (*reactive oxygen scavenging enzymes*), capables de piéger les ROS, ainsi que des molécules anti-oxydantes, toutes plus ou moins spécifiques de compartiments cellulaires [104][101] :

- les superoxide dismutases, **SOD**, catalysent la réaction de dismutation de l'ion  $\text{O}_2^-$  en peroxyde d'hydrogène ( $\text{H}_2\text{O}_2$ ) et  $\text{O}_2$  et limitent ainsi la formation de radicaux hydroxyl issue de l'ion superoxide. Les SOD sont présentes dans de nombreux compartiments (plaste, mitochondrie, peroxysome et cytosol) et jouent ainsi un rôle de première ligne de défense

contre les ROS.

- les catalases, **CAT**, sont capable de dismuter très rapidement le  $\text{H}_2\text{O}_2$  en  $\text{H}_2\text{O}$  et  $\text{O}_2$ , sans consommer de cofacteur réducteur. Elles sont présentes dans le cytosol et les peroxysomes (et donc absentes du plaste).
- la glutathion peroxidase, **GPX**, permet également d'éliminer le  $\text{H}_2\text{O}_2$  et surtout les acides gras peroxydés et nécessite la présence d'un co-facteur réducteur, le glutathion (GSH). On la retrouve dans les plastes, les mitochondries, le cytosol et le reticulum endoplasmique.
- l'ascorbate peroxidase, transforme le  $\text{H}_2\text{O}_2$  en eau, en utilisant l'ascorbate comme substrat.

D'autres enzymes et anti-oxydants agissent de concert, comme dans le système de détoxification présenté en figure 10. Le cycle "eau-à-eau" [108] permet de dégrader une partie des ROS produites en excès par les photosystèmes et de dissiper un flux d'électrons excédentaire dans le chloroplaste. Dans une situation de stress, lorsque la fixation du carbone est réduite ou qu'il y a une carence en  $\text{NADP}^+$ , accepteur final d'électrons, les électrons sont détournés par la ferrédoxine vers l'oxygène moléculaire  $\text{O}_2$  qui est alors réduit en anion superoxide  $\text{O}_2^-$ . Ces molécules sont dismutées en peroxyde d'hydrogène  $\text{H}_2\text{O}_2$  par la superoxide dismutase (SOD). Deux voies sont possibles pour transformer le  $\text{H}_2\text{O}_2$  en eau : directement par la catalase (CAT) ou en impliquant davantage d'acteurs tels que l'acide ascorbique, le glutathion et des enzymes.

Les enzymes telles que SOD, CAT et APX sont aussi des acteurs importants dans la décomposition des ROS produites dans les autres compartiments cellulaires [105].

#### 1.3.4 Le stress chez les algues

Le stress abiotique chez les macroalgues a fait l'objet de recherches principalement du point de vue physiologique (voir Davison & Pearson, 1996 pour revue [59]). Il a ainsi été montré que la croissance d'une laminaire est affectée par la chaleur, les faibles lumières et la limitation en nutriments [109]. La zonation verticale des fucales dans l'estran a été corrélée à l'activité des enzymes de détoxification des ROS, mais pas à la quantité absolue d'antioxydants et d'enzymes ([60]). L'effort de recherche est d'ailleurs le plus important sur le métabolisme oxydatif. Chez l'algue verte *Ulva fasciata*, le cycle "eau-eau" est induit par un stress UV [110]. Chez l'algue rouge encroûtante *Corallina officinalis*, les UV induisent aussi les activités des enzymes de détoxification [111].

Chez *C. crispus*, quelques études ciblées ont également porté sur le stress. *C. crispus* est plus sensible à la lumière que d'autres algues intertidales [112], lorsqu'il est placé en culture exté-

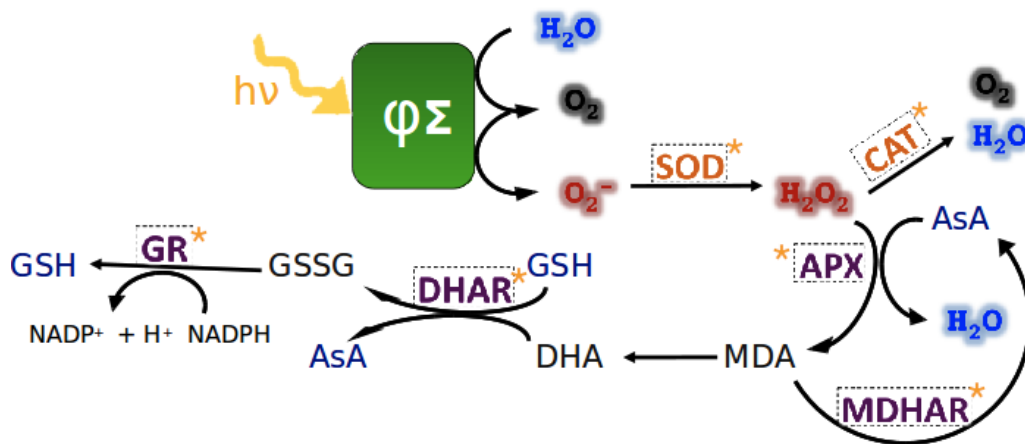


FIGURE 10: Cycle “eau-à-eau”. SOD : superoxyde dismutase, CAT : catalase, APX : ascorbate peroxidase, (M)DHAR : (mono)dehydroascorbate réductase, GR : glutathion réductase, GSH : glutathion réduit, GSSG : glutathion oxydé, AsA : acide ascorbique, MDA : monodehydroascorbate, DHA : dehydroascorbate,  $\Phi\Sigma$  : éléments de l’appareil photosynthétique, en rouge : formes actives de l’oxygène (d’après Asada, 2000 [108]). Les étoiles indiquent les gènes présent dans le génome de *C. crispus*.

The “water-water” cycle. SOD : superoxide dismutase, CAT : catalase, APX : ascorbate peroxidase, (M)DHAR : (mono)dehydroascorbate reductase, GR : glutathione reductase, GSH : reduced glutathione, GSSG : oxidized glutathione, AsA : ascorbic acid, MDA : monodehydroascorbate, DHA : dehydroascorbate,  $\Phi\Sigma$  : photosynthetic apparatus element, in red : reactive oxygen species (adapted from Asada, 2000 [108]). Stars indicate the genes found in the genome of *C. crispus*.

rieure à différentes intensités lumineuses, le rendement photosynthétique diminue, reflétant des phénomènes de photoinhibition, dès 8 % de lumière incidente. Des mesures de croissance ont montré que les jeunes gamétophytes sont significativement moins sensibles aux rayonnements UV que les tetrasporophytes [83]. Une autre étude [93], portant sur le stress oxydatif au cours de l'année chez *C. crispus* et *Mastocarpus stellatus*, a montré que les activités enzymatiques de SOD, CAT et GR sont plus élevées chez ces deux algues au cours des mois d'hiver qu'au cours des mois d'été, indiquant que la production de ROS est plus importante lorsque le métabolisme (notamment photosynthétique) est ralenti par le froid. Ceci semble indiquer que le rapport lumière/vitesse de photosynthèse est plus élevé en hiver qu'en été. Dans d'autres études [114][113], il a été montré que la tolérance au gel peut expliquer en partie la distribution des deux algues dans l'estran, *Mastocarpus stellatus*, qui vit plus haut que *C. crispus*, ayant une capacité à capturer les ROS plus élevée que ce dernier. Enfin, une dernière étude [100] indique que le métabolisme oxydatif de *C. crispus* ne varie pas avec le cycle des marées (contrairement à celui *Mastocarpus stellatus*). Des expériences à plus grande échelle ont également été menées, dans le but d'étudier la réponse plus globale de *C. crispus*, à des stress contrôlés ou naturels [14][16] en utilisant des puces à ADN. Les résultats ont montré qu'en conditions de laboratoire les changements osmotiques semblaient être un facteur de stress important et qu'une forte intensité lumineuse, ainsi que des conditions naturelles considérées comme fortement stressantes, induisaient toutes deux l'expression de gènes relatifs au métabolisme oxydatif. Cependant ces études avaient été menées avant le séquençage du génome et ne prenaient en considération que 15 % des gènes.

Si l'ensemble de ces études met l'accent sur l'importance du métabolisme oxydatif dans la réponse aux stress abiotiques chez les algues et particulièrement chez *C. crispus*, il manque cependant de réelles études globales relatives au stress dans des conditions naturelles. La position phylogénétique des algues rouges en tant que lignée indépendante, bien que groupe frère des plantes terrestres, leur confère sans doute des mécanismes communs, mais à la vue de la quantité de gènes aux fonctions totalement inconnues, le potentiel de découverte de nouveaux mécanismes chez ces algues est considérable.

### 1.3.5 La lumière : à consommer avec modération

La photosynthèse est un processus essentiel permettant aux végétaux et autres organismes autotrophes de convertir l'énergie lumineuse solaire en énergie chimique (ATP et NADPH), tout

en fixant le carbone minéral ( $\text{CO}_2$ ) pour ainsi produire de la matière organique,  $(\text{CH}_2\text{O})_n$ .

En faible éclaircissement, le rendement de la photosynthèse est proportionnel à l'intensité lumineuse. Mais lorsque cette dernière dépasse un seuil, variable selon les organismes, le rendement devient limité par la vitesse du flux d'électrons et par la capacité de fixation du dioxyde de carbone dans le cycle de Calvin [95]. La chaîne d'électrons ne fonctionnant pas assez vite, la lumière excédentaire génère des ROS, directement destructrices pour les protéines du PSII. Celles-ci (particulièrement la protéine D1) sont resynthétisées en continu, néanmoins si la photodestruction excède la vitesse de synthèse protéique, la quantité de protéines D1 diminue, il y a alors photoinhibition [98]. Les végétaux doivent être capables de réguler la capture de lumière surtout en cas d'énergie excédentaire, mais également en fonction des besoins cellulaires en ATP et NADPH. Il existe plusieurs mécanismes permettant aux organismes photosynthétiques de se protéger et s'adapter au régime lumineux, qui seront décrits dans la section 1.4.4.

## 1.4 La photosynthèse comparée

Il existe deux types principaux de photosynthèse. La première, pratiquée par certaines bactéries vertes, pourpres ou héliobactéries, est dite anoxygénique. Elle consomme du  $\text{H}_2\text{S}$  et produit du soufre  $\text{S}_2$ . Elle implique une chaîne d'électrons comportant un seul photosystème, des pigments particuliers : les bactériochlorophylles et, comme son nom l'indique, elle ne produit pas d'oxygène. Le deuxième type de photosynthèse, dite oxygénique, est la plus répandue dans le vivant et concerne les phototrophes eukaryotes et les cyanobactéries. Elle implique deux photosystèmes, des chlorophylles, elle consomme de l'eau,  $\text{H}_2\text{O}$  et produit du dioxygène. *C. crispus* faisant partie des eukaryotes photosynthétiques, c'est sur la photosynthèse oxygénique que sera focalisée la suite de ce manuscrit.

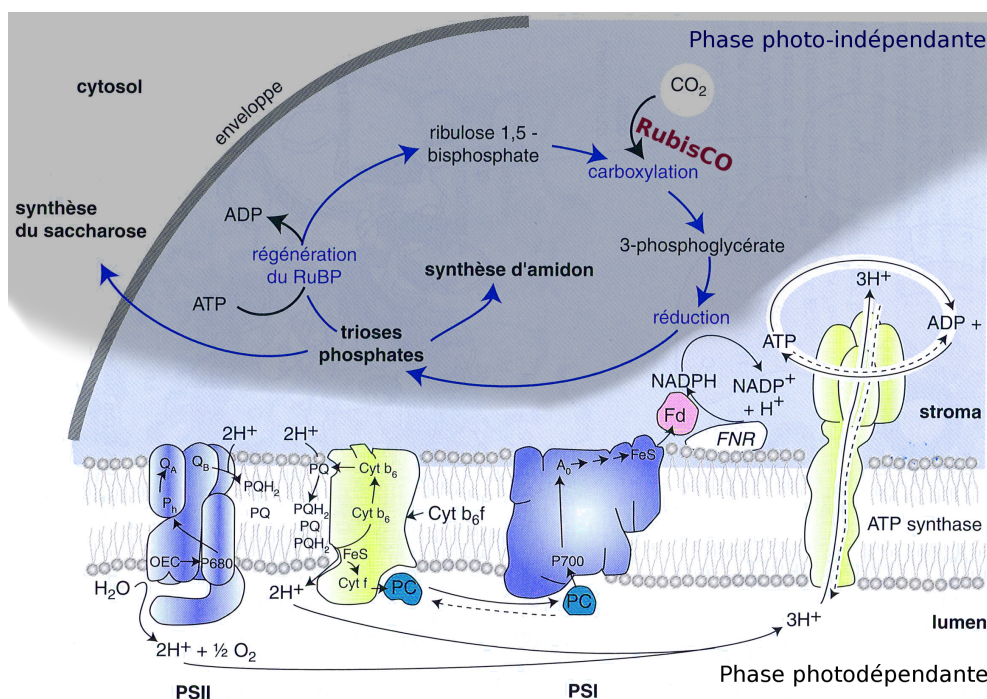
### 1.4.1 Une obscure clarté...

Le processus photosynthétique se déroule selon deux phases distinctes, inter-dépendantes, schématisées en figure 11 :

- la phase dite “claire” ou photochimique, qui est photo-dépendante. Elle débute par la séparation de charge ( $\text{P}_{680} \longrightarrow \text{P}_{680}^* + \text{e}^-$ ) et par la photolyse de l'eau qui, en produisant de l'oxygène moléculaire et des protons, va céder ses électrons dans la chaîne de transporteurs qui sera décrite dans la section suivante. Cette phase produit directement de l'énergie réductrice sous forme de NADPH, et de l'ATP grâce une enzyme transmembranaire thyla-



- la phase dite “sombre”, photo-indépendante. C’est au cours de cette phase que se déroule le cycle de Calvin [115]. L’ATP et le NADPH produit lors de la phase “claire” sont utilisés conjointement pour fixer le  $\text{CO}_2$ , via une enzyme clef, la RubisCO (Ribulose-1,5-bisphosphate Carboxylase/Oxygénase). Le produit de cette phase est le 3-phosphoglycérate qui est par la suite transformé en glucose  $\text{C}_6\text{H}_{12}\text{O}_6$ .



Physical and metabolic reactions of photosynthesis. Adapted from Farineau & Moreau-Gaudry, 2011 [33].

**La photosynthèse en vert.** L'appareil photosynthétique le mieux décrit jusqu'à présent est celui des végétaux terrestres et des algues vertes (voir Eberhard *et al.*, 2008 [116], pour revue). Il est schématisé en figure 12A. Les membranes thylakoïdales des végétaux verts sont organisés en piles de grana dans les zones dites appressées et entre chaque granum se trouve une zone intergranaire. Ceci permet de définir des compartiments séparant les photosystèmes : les PSII se situant dans les zones appressées alors que les PSI seraient majoritairement au niveau des



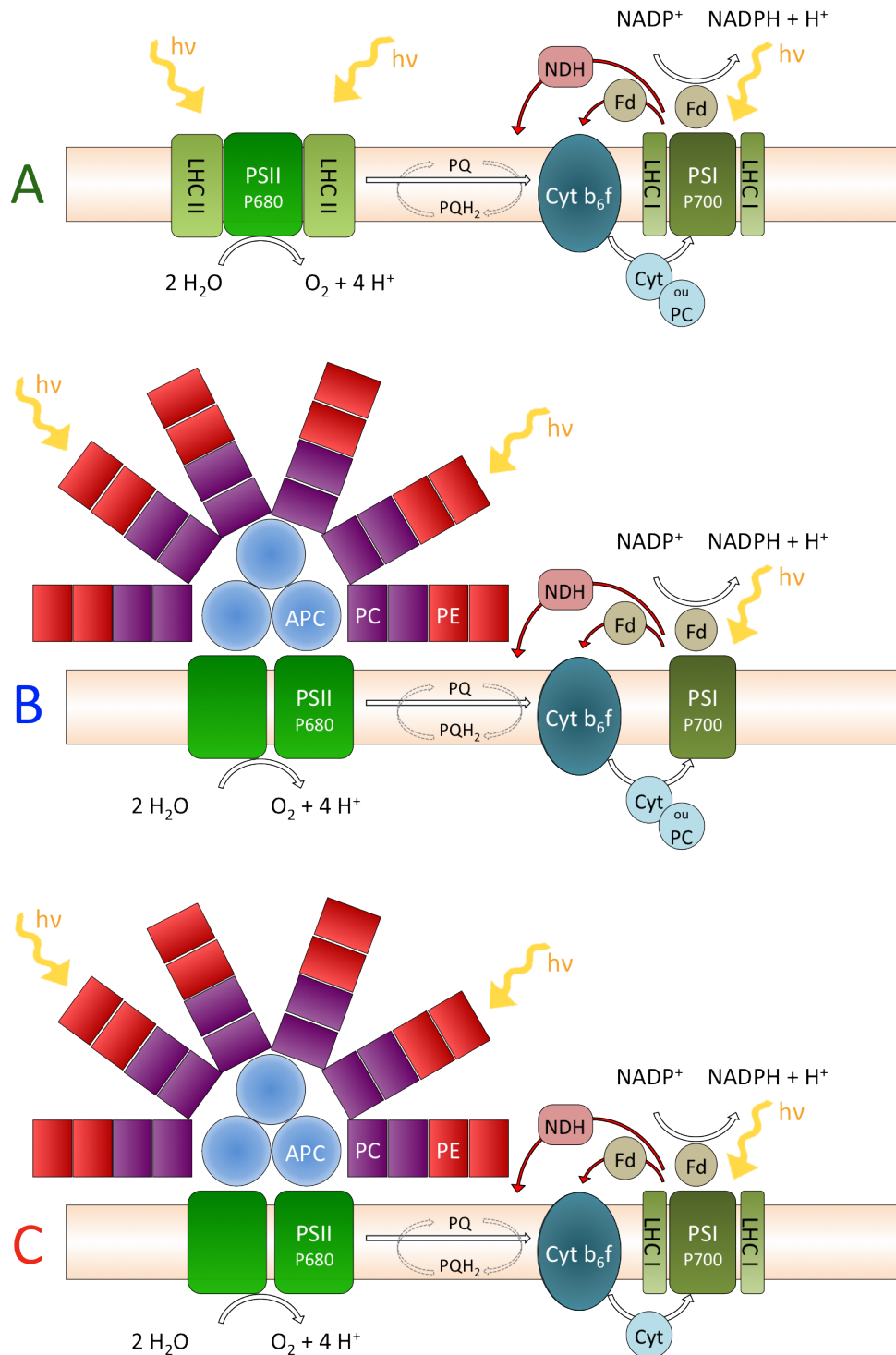


FIGURE 12: Schéma de trois types d'appareils photosynthétiques. A : chez les végétaux verts, B : chez les cyanobactéries, C : chez les rhodophytes. APC : allophycocyanine, PC : phycocyanine, PE : phycoérythrine, PS : photosystème, LHC : complexe collecteur de lumière, PQ : pool de quinones, Fd : ferrédoxine, NDH : NADPH deshydrogénase, Cyt : cytochrome,  $h\nu$  : lumière.

Schematic representation of three types of photosynthetic apparatus. A : in plants and green algae, B : in cyanobacteria, C : in rhodophyta. APC : allophycocyanin, PC : phycocyanin, PE : phycoerythrin, PS : photosystem, LHC : light harvesting complex, PQ : quinone pool, Fd : ferredoxin, NDH : NADPH dehydrogenase, Cyt : cytochrome,  $h\nu$  : light.

zones intergranaires :

- Le PSII, possède une paire spéciale de chlorophylles,  $P_{680}$ , et une antenne collectrice de lumière transmembranaire (LHCII) comportant des chlorophylles a et b,
- le PSI, possède un autre paire spéciale de chlorophylles,  $P_{700}$ , et une antenne collectrice de lumière chlorophyllienne transmembranaire (LHCI).

Entre les deux, une chaîne de transporteurs permet d'assurer le transfert des électrons :

- le pool de plastoquinones : un cortège de molécules solubles dans la membrane et mobiles,
- le complexe de cytochromes  $b_6f$ , transmembranaire,
- les plastocyanines et/ou cytochromes mobiles.

Pendant la photosynthèse, parallèlement au transport d'électrons, des protons  $H^+$  sont transloqués du stroma vers la lumière des thylakoïdes (lumen), entraînant leur accumulation. Le gradient électrochimique ainsi formé permet le fonctionnement de l'ATP synthase, complexe enzymatique également enchâssé dans la membrane thylakoïdale, assurant la production d'ATP à partir de l'ADP, dans le stroma [117].

**La photosynthèse en brun/doré.** Chez les algues brunes et diatomées, la configuration des éléments de l'appareil photosynthétique est semblable à celle des végétaux verts, ils possèdent des antennes transmembranaires de structure proche des LHC mais dont la pigmentation est sensiblement différente [118] : la chlorophylle b y est remplacée par de la chlorophylle c et ces organismes sont riches en caroténoïdes de type fucoxanthine, ce qui leur vaut l'appellation FCP pour *Fucoxanthin-Chlorophyll binding Proteins*, et donnent leur couleur brune ou dorée à ces algues. Il semblerait que les antennes ne puissent pas être attribuées spécifiquement au PSI ou PSII [119].

**La photosynthèse en bleu.** Chez les cyanobactéries (figure 12B), descendants directs des ancêtres des plastes eukaryotes, la configuration de l'appareil photosynthétique est en partie différente, d'abord par l'absence de véritable antenne dédiée aux PSI, qui sont ici associés sous formes de trimères et par la présence de phycobilisomes (PBS), des antennes protéiques non-chlorophylliennes à la place des antennes LHCII. L'organisation des thylakoïdes est différente de celles de la lignée verte, puisqu'ils ne forment pas de grana et les photosystèmes sont arrangés de manière homogène, sans compartimentation particulière.

**La photosynthèse en rouge.** Les algues rouges (figure 12C) possèdent une antenne dédiée au PSII de type phycobilisome comme les cyanobactéries, et une antenne dédiée au PSI (*a priori* monomérique) chlorophyllienne et transmembranaire, semblable à celle des végétaux verts [120]. A l'instar des cyanobactéries, les photosystèmes sont repartis uniformément au sein de la membrane thylakoïdale.

La structure de l'appareil photosynthétique des algues rouges ne ressemble donc pas tout à fait ni à celle de la lignée verte et ni à celle des cyanobactéries et semble davantage représenter un intermédiaire entre les deux.

### 1.4.3 Les phycobilisomes

Les phycobilisomes (PBS) sont les antennes collectrices principales des cyanobactéries, du PSII des algues rouges, ainsi que d'autres organismes tels que les glaucophytes [121]. Ce sont d'importantes structures protéiques situées sur la face externe des thylakoïdes (stroma), solubles en solution aqueuse. Ils sont composés de phycobiliprotéines, qui sont des protéines liées à des pigments. Ces phycobiliprotéines sont reliées entre-elles par des polypeptides de liaison (dits protéines *linker*) de manière à former des complexes : les bras des PBS. Les phycobiliprotéines peuvent être rouges (phycoérythrine, PE) bleues ou violettes (phycocyanine, PC et allophycocyanine, APC) conférant des propriétés d'absorption de la lumière plus large que celle des végétaux verts (exemple de spectres en figure 13). La chlorophylle seule absorbe essentiellement le bleu et le rouge (d'où sa couleur verte) alors que les phycobilisomes absorbent une large gamme entre le bleu-vert et le jaune, permettant aux organismes à PBS d'utiliser les longueurs d'onde de la lumière filtrée par une colonne d'eau épaisse, qui laisse passer essentiellement la lumière bleue-verte (responsable ainsi de la couleur de l'eau).

Les transferts d'énergie au sein des antennes se font par résonance et sont imparfaits, puisqu'une faible partie (environ 10 %) est ré-émise sous forme de fluorescence. Ces transferts (schématisés sur la figure 14) sont par conséquent orientés, des courtes longueurs d'onde, de haute énergie, vers les plus longues, de plus faible énergie [122]. Les phycobiliprotéines absorbent chacune des longueurs d'onde spécifiques et sont organisées verticalement par ordre d'absorption : la PE absorbe préférentiellement autour de 500-550 nm (vert) et se trouve aux extrémités des bras des PBS. Elle transmet son énergie à la PC qui se trouve au milieu des bras, qui à son tour transmet son énergie à l'APC, au coeur du PBS. Celle-ci est proche de l'antenne-coeur du PSII et peut transférer son énergie aux chlorophylles, plus particulièrement à la paire spéciale

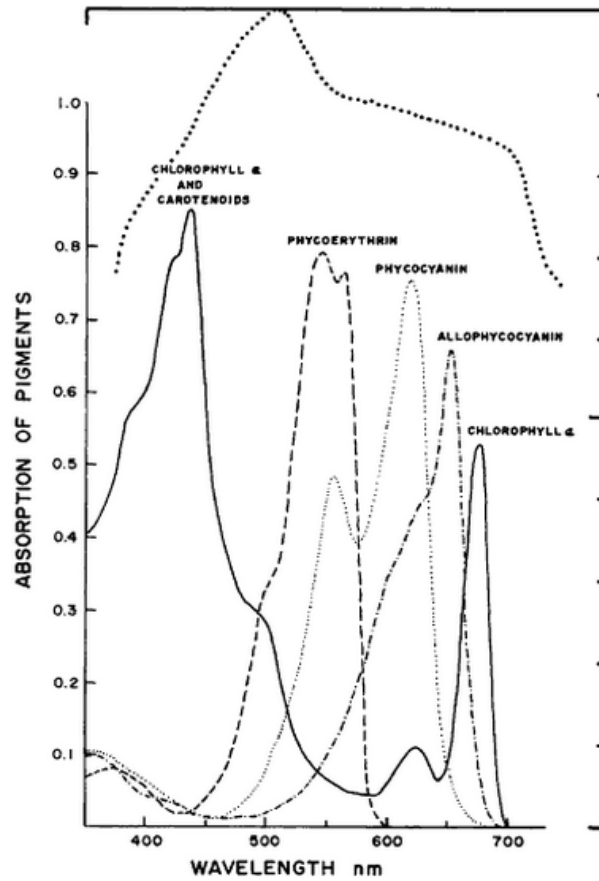


FIGURE 13: Spectres d'absorption des pigments photosynthétiques (Gantt, 1975[121]).  
Absorption spectra of photosynthetic pigments (Gantt, 1975[121]).

P680. Quand au PSI, contenant la paire de chlorophylles P700, il absorbe de la lumière de plus grande longueur d'onde, le rouge lointain (FR : *far red*). Par conséquent, le PSI peut recevoir directement de la lumière mais également être alimenté par l'énergie transférée depuis les PBS, aussi bien que celle transmise via le PSII.

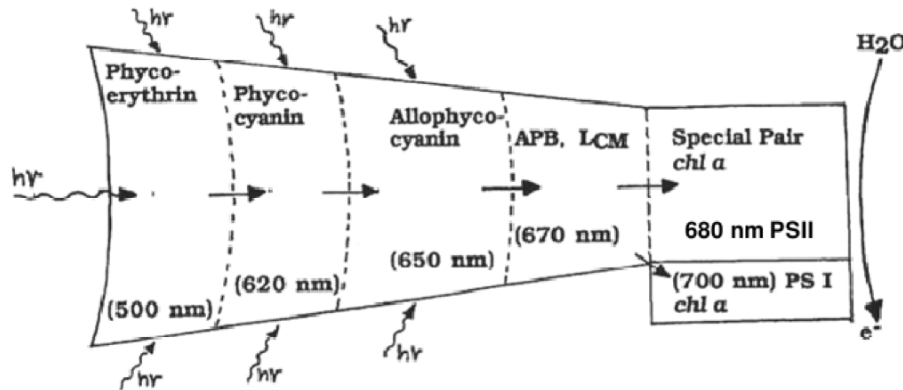


FIGURE 14: Schéma des transferts d'énergie orientés entre les phycobiliprotéines contenues dans les phycobilisomes. D'après Sidler, 1994 [123].  
Scheme of oriented energy transfers between phycobiliproteins along phycobilisome rods. Adapted from Sidler, 1994 [123].

La disposition des PBS à la surface des thylakoïdes suit différents motifs selon les organismes [124]. La figure 15, représente des PBS de *Porphyridium cruentum*, une algue rouge unicellulaire, vus en microscopie à force atomique (AFM). Leur taille est estimée à environ 50 nm de diamètre et est supérieure à celle d'un PSII (environ 15-20 nm). Ils sont ici organisés en rangs serrés. Cette très forte densité les rend responsables de la pigmentation des cellules : le plus souvent rouges chez les algues rouges et toute une gamme de couleurs entre le bleu et le rouge pour les cyanobactéries. L'arrangement des PBS chez *C. crispus* est encore inconnu.

#### 1.4.4 Régulation de la lumière

Les végétaux subissent en continu des variations d'intensité lumineuse et n'ont pas toujours des besoins constants en énergie. L'excès de lumière conduit à du stress oxydatif et à la photoinhibition du PSII. À l'inverse, le manque de lumière provoque des carences en énergie. De plus, les besoins de la cellule en NADPH ou ATP, peuvent également être variables. L'équilibre entre NADPH et ATP peut être régulé par le flux d'électrons, qui sera :

- linéaire, pour produire de l'ATP et du NADPH dans une stoechiométrie précise, tous deux étant utilisés dans le cycle de Calvin pour fixer le carbone [125].
- cyclique ou pseudo-cyclique, lorsque les électrons sont réinjectés vers le pool de plastoqui-

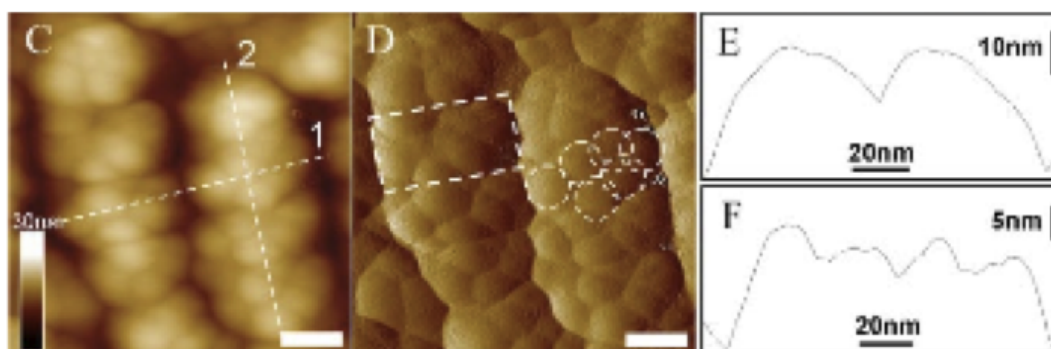


FIGURE 15: Micrographies AFM de la membrane thylakoïdale de *P. cruentum*. C. vue tridimensionnelle des PBS, D. même vue améliorée sur laquelle on distingue un PBS (rectangle) et des bras de PBS (cercles). Echelle C et D : 25 nm. E et F : coupes de PBS issues de la figure C selon, respectivement, les axes 1 et 2 [124].

AFM micrographs of thylakoidal membrane from *P. cruentum*. C. tridimensional view of PBS, D. enhanced 3D view showing one PBS (rectangle) and PBS rods (cercles). Scale C and D : 25 nm. E and F : height profiles of PBS along lines 1 and 2 in panel C, respectively [124].

nones ou le cytochrome  $b_6f$ , pour privilégier la production d'ATP [126].

Le PSI ou le PSII doivent donc être favorisés en matière d'apport d'énergie, pour créer soit de la force réductrice soit de l'énergie sous forme d'ATP.

La gestion de l'énergie par les photosystèmes s'opère par plusieurs mécanismes qui diffèrent selon les organismes.

**Le cycle des xanthophylles et la dissipation d'énergie.** Chez les plantes terrestres et les algues vertes, une stratégie à court terme est de dissiper l'énergie excédentaire par une réaction réversible localisée au niveau de pigments accessoires, des caroténoïdes, localisés au sein des antennes LHCII. Cette réaction, nommée "cycle des xanthophylles" [127] consiste à dé-époxyder la violaxanthine en zéaxanthine, via l'anthéroxanthine, afin de transformer l'énergie lumineuse en chaleur. Elle implique deux enzymes : la violaxanthine dé-époxydase et la zéaxanthine époxydase, dont l'activation est dépendante de l'acidification du lumen, résultant de la formation du gradient de protons photo-induit, parallèle au transport d'électrons. Une protéine transmembranaire, PsbS, est indispensable pour cette dissipation thermique et faciliterait le transfert d'énergie entre la chlorophylle et la violaxanthine [128].

Chez les algues brunes, le même système existe, mais ne semble pas être dépendant du gradient de protons photo-induit [129].

Chez les diatomées et dinoflagellés, il existe un cycle des xanthophylles similaire à celui des

plantes et algues vertes et brunes, les caroténoïdes en présence sont juste différents puisqu'il s'agit de diadinoxanthine qui est convertie en diatoxanthine [130]. Les enzymes chargées de cette interconversion sont spécifiques de ces caroténoïdes, ce sont la diadinoxanthine dé-époxydase et la diatoxanthine epoxidase. Il ne semble pas y avoir d'analogue à la protéine PsbS.

Les cyanobactéries ne possèdent pas d'équivalent au cycle des xanthophylles, en revanche, il a été décrit que beaucoup d'entre-elles possèdent une caroténo-protéine orange, OCP, liant une xanthophylle photo-activable, capable d'agir comme un tampon en absorbant l'énergie entre les phycobilisomes et les photosystèmes [136]. Ce mécanisme n'implique pas le gradient de protons photo-induit.

Chez les algues rouges, bien que la zéaxanthine soit présente, ainsi que de la violaxanthine en faible quantité (voir Takaishi, 2011 [137]), les enzymes impliquées dans le cycle des Xanthophylles sont absentes du génome.

**Les transitions d'état.** Il existe un autre mécanisme de régulation de la capture lumineuse, qui est le phénomène de transition d'état (schématisé en figure 16). Il permet d'équilibrer les activités des deux photosystèmes en modulant les tailles d'antennes collectrices de photons.

Chez *C. reinhardtii*, sous illumination modérée, lorsque les plastoquinones sont dans un état majoritairement réduit, signifiant que le PSII fonctionne plus rapidement que la chaîne d'électrons en aval, une kinase (Stt7) va phosphoryler les antennes LHCII qui vont alors se déplacer depuis le PSII vers le PSI, permettant de modifier la répartition de la lumière entre les deux photosystèmes [131]. Le PSI va pouvoir fonctionner plus rapidement et permettre la réoxydation des plastoquinones. Les transitions d'état ont aussi l'avantage d'envoyer l'énergie excédentaire sur le PSI plutôt que sur le PSII, l'antenne du PSI étant en effet capable de dissiper l'énergie [132] et le PSI étant particulièrement efficace en matière de photochimie [133]. Lorsque le surplus d'énergie est canalisé vers le PSI, les risques de photoinhibition du PSII sont réduits d'autant. Le mécanisme de transition d'état est contrôlé par l'état d'oxydoréduction du pool de Plastoquinones et est dépendant de la phosphorylation des protéines d'antenne LHCII. Si ce phénomène est plutôt minoritaire chez les plantes terrestres [134], en revanche, dans l'algue verte *Chlamydomonas*, il semble contribuer pour une part importante dans la régulation de la capture lumineuse [135].

Cette relocalisation des antennes permet aussi de mettre en place un flux cyclique d'électrons entre le PSI et le cytochrome *b<sub>6</sub>f*, qui va uniquement générer un gradient électrochimique de

protons et favoriser ainsi la production d'ATP sans augmenter la quantité de NADPH [131].

Chez les cyanobactéries, la transition d'état est le système majoritairement décrit, mais à défaut d'antenne chlorophyllienne, c'est le phycobilisome qui se déplace entre le PSII et le PSI.

Jusqu'à présent, l'existence de phycobilisomes dans les algues rouges suggère qu'il existe un système similaire à celui des transitions d'états [142].

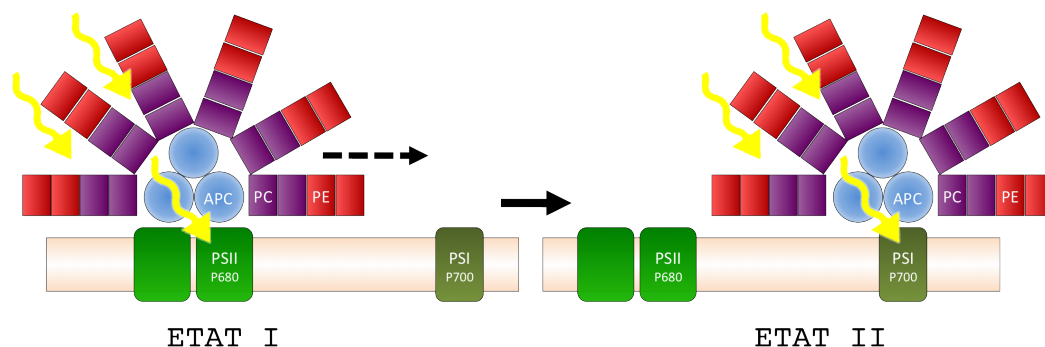


FIGURE 16: Schéma du mécanisme de transition d'état chez un organisme à phycobilisomes.  
Schematic representation of the mechanism of state-transition in a phycobilisome-containing organism.

**La gestion de la capture lumineuse chez les algues rouges.** La question de la régulation de la capture lumineuse et de la dissipation d'énergie, déjà posée dans d'autres études [138],[139],[140],[141], reste ouverte. Il n'y a pas de cycles des xanthophylles et la protéine OCP n'a pas été identifiée non plus dans le génome de *C. crispus*. D'autres mécanismes de dissipation d'énergie sous forme thermique ne sont pas connus chez les algues rouges.

Les différences existant entre l'organisation de l'appareil photosynthétique des algues rouges et celui des cyanobactéries ou de la lignée verte, aussi bien au niveau de l'architecture que de l'absence des acteurs connus relatifs à la photoprotection, suggèrent un fonctionnement original et ouvrent des perspectives de découverte de nouveaux mécanismes.

## 1.5 L'étude de la photosynthèse *in vivo*

Pour étudier les mécanismes relatifs à la photosynthèse, il existe des techniques non-invasives, donc réalisables *in vivo*, permettant de suivre l'état d'oxydoréduction de certains composants de la chaîne de transporteurs d'électrons (voir le schéma 17) et par extension de donner des informations relatives aux antennes.



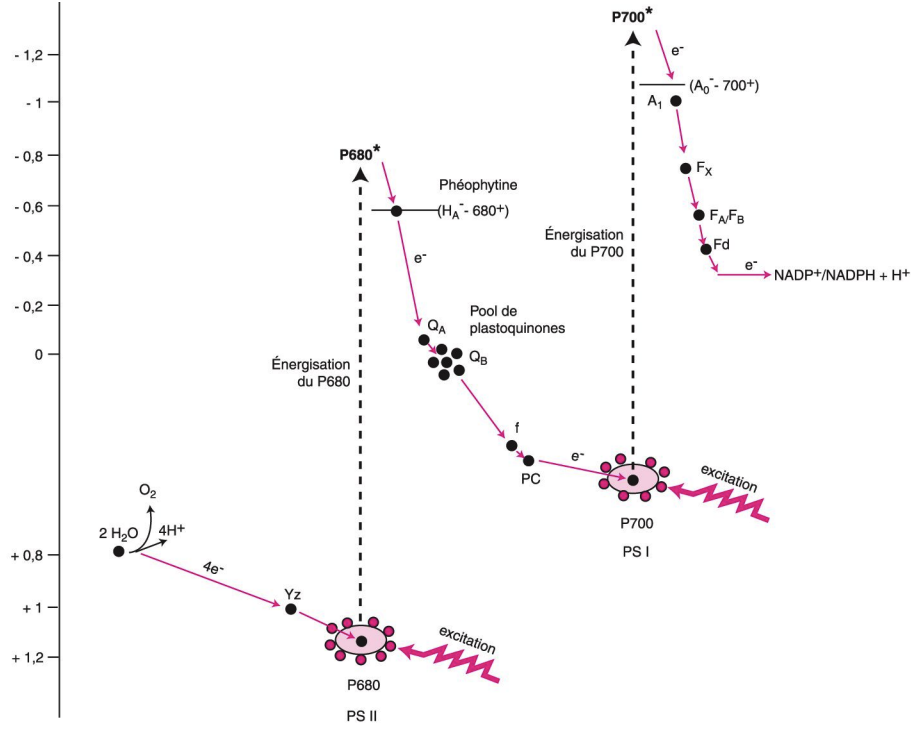


FIGURE 17: Schéma dit « en Z » présentant, sur une échelle de potentiel rédox (en volt), les transferts d'électrons entre les différents transporteurs.  $\text{Y}_Z$  : tyrosine Z (donneur primaire d'électrons), P680 : paire spéciale de chlorophylles du centre réactionnel du PSII,  $\text{f}$  : cytochrome f, PC : plastocyanine, P700 : paire spéciale de chlorophylles du centre réactionnel du PSI,  $\text{A}_0$  et  $\text{A}_1$  : chlorophylles modifiées du PSI,  $\text{F}_{X,A,B}$  : complexes de protéines fer-soufre, Fd : ferredoxine [33].

The so-called "Z scheme" showing on a redox potential scale (in volts), the different electron transfers.  $\text{Y}_Z$  : tyrosine Z (primary electron donor), P680 : special pair of chlorophylls of PSII reaction center,  $\text{f}$  : cytochrome f, PC : plastocyanin, P700 : special pair of chlorophylls of PSI reaction center,  $\text{A}_0$  et  $\text{A}_1$  : modified chlorophylls of PSI,  $\text{F}_{X,A,B}$  : iron-sulfur protein complexes, Fd : ferredoxin [33].

### 1.5.1 La fluorescence

Pour étudier la photosynthèse, une première technique repose sur une propriété physico-chimique des organismes photosynthétiques : la mesure de la fluorescence de la chlorophylle (voir Maxwell & Johnson [143] et Baker [144] pour revue). Elle est couramment utilisée chez les végétaux verts et les algues brunes et permet de mesurer l'état redox du PSII et d'estimer ainsi l'efficacité du PSII.

Le principe de la mesure de fluorescence est le suivant : lorsque qu'une molécule de chlorophylle est excitée par de la lumière, le retour à son état d'énergie initial peut se faire de trois manières :

- en transformant l'énergie lumineuse en énergie chimique (qui consiste en fait en une séparation de charges au niveau de  $P_{680}$ ),
- en dissipant l'énergie lumineuse sous forme de chaleur,
- en ré-émettant un photon de plus faible énergie : c'est la fluorescence.

Ces trois formes sont en compétition et pour caractériser les deux premières on parle alors d'extinction ou *quenching* de fluorescence. Le quenching est photochimique (PQ) pour la photosynthèse, et non-photochimique (NPQ) pour la dissipation thermique.

La mesure est réalisée ainsi (figure 18) : une faible lumière détectrice, appliquée sur un matériel biologique adapté à l'obscurité, génère un niveau basal de fluorescence,  $F_0$ , émanant des antennes et des photosystèmes. Cette lumière n'induit pas de transfert d'électrons et le système est alors dit "ouvert", prêt à recevoir de l'énergie, les accepteurs primaires d'électrons ( $Q_A$ ) étant oxydés. On déclenche par la suite un flash lumineux très intense et court de manière à saturer le PS II. Le système est alors "fermé", les accepteurs d'électrons sont réduits et l'énergie est alors restituée majoritairement sous forme de fluorescence, qui, mesurée avant la réoxydation de  $Q_A$ , atteint son niveau maximal,  $F_M$ .

La différence entre  $F_M$  et  $F_0$  est la fluorescence variable, qui reflète la quantité d'énergie que le photosystème II peut utiliser pour la photosynthèse et permet d'évaluer la proportion de centres réactionnels utilisables. Le calcul du rapport  $(F_M - F_0)/F_M = F_V/F_M$  représente le rendement quantique du PSII, souvent considéré comme le reflet de l'état physiologique de l'organisme.

### 1.5.2 Spectrométrie $P_{700}$ et $b_6f$

En complément de la fluorescence de la chlorophylle qui permet d'étudier le PSII, la spectrométrie permet d'étudier d'autres composants de la chaîne photosynthétique dont on ne peut

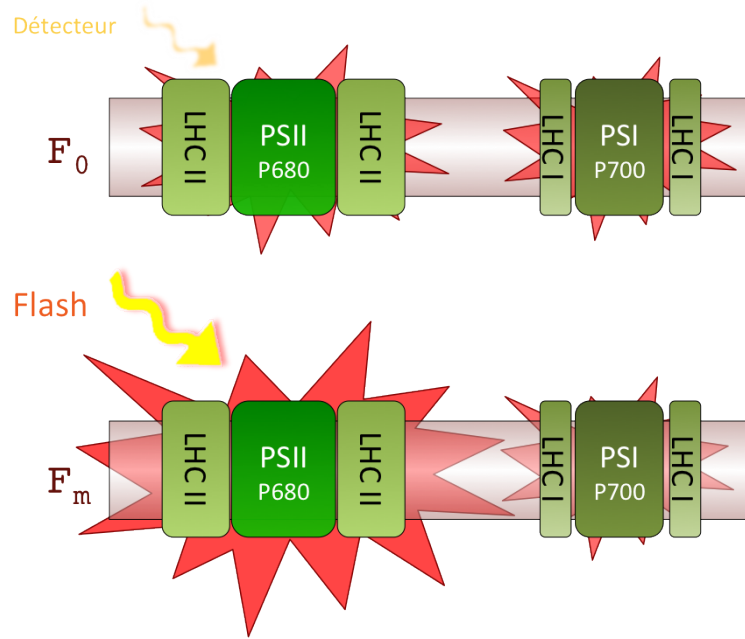


FIGURE 18: Vision simplifiée d'une mesure de fluorescence de la chlorophylle.  $F_0$  : fluorescence basale,  $F_M$  : fluorescence maximale.  
Simplified view of a chlorophyll fluorescence measurement.  $F_0$  : minimum fluorescence,  $F_M$  : maximum fluorescence.

pas suivre l'émission de fluorescence. En mesurant les variations de transmission lumineuse à 705 nm et 554 nm, il est possible de suivre respectivement, les cinétiques d'oxydation et de réduction de la paire spéciale de chlorophylles  $P_{700}$  du PSI, ainsi que l'oxydation du complexe de cytochromes  $b_6f$ . À ces longueurs d'onde, une diminution d'absorbance reflète l'oxydation

La vitesse d'oxydation de  $P_{700}$  permet d'informer sur la quantité d'électrons présents dans la chaîne en amont du PSI ainsi que d'estimer la contribution du flux cyclique d'électrons. Mesurer l'état redox des cytochromes permet en revanche d'estimer la capacité du PSI à oxyder les transporteurs situés en amont et peut refléter ainsi la taille de l'antenne du PSI.

Comme pour la fluorescence, les mesures d'état redox de PSI et des cytochromes sont relatives. Pour chaque matériel, un flash de lumière saturante permet de définir le maximum d'oxydation. Les cinétiques sont ensuite réalisées avec des intensités lumineuses plus faibles, non-saturantes.

### 1.5.3 Phénomènes de photoprotection

Chez les plantes vertes [116], le protocole décrit à la figure 19 est utilisé pour évaluer les mécanismes de photoprotection et de photoinhibition :

Une première mesure de  $F_V/F_M$  renseigne sur l'état initial de la plante à l'obscurité. Elle est

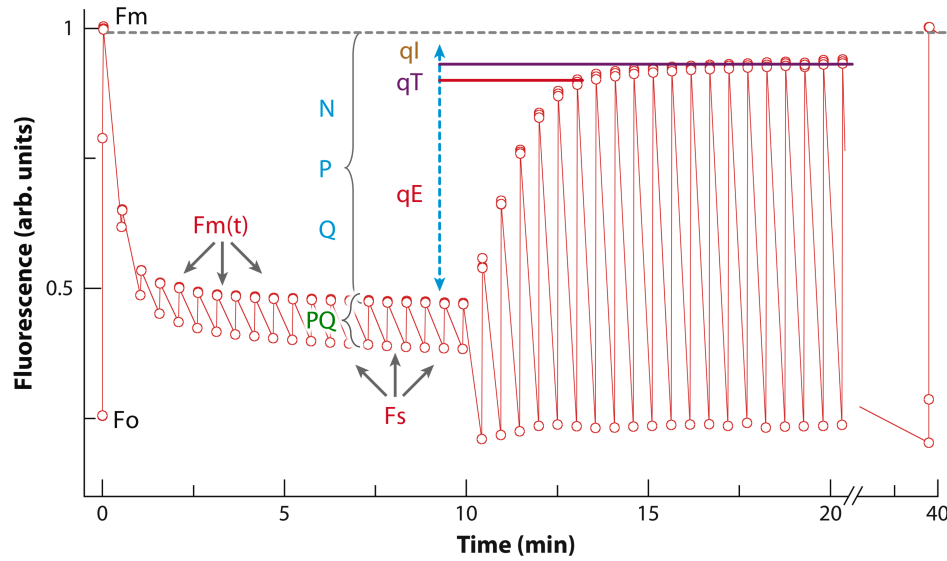


FIGURE 19: Protocole de mesure de fluorescence de la chlorophylle utilisé chez les plantes terrestres.  $F_0$  : fluorescence basale,  $F_M$  : fluorescence maximale,  $F_S$  : fluorescence “stationnaire”, NPQ : extinction de fluorescence non-photochimique, PQ : extinction de fluorescence photochimique, qE : dissipation thermique, qI : photoinhibition, qT : transition d’état (adapté d’après Eberhard *et al.*, 2008 [116]).

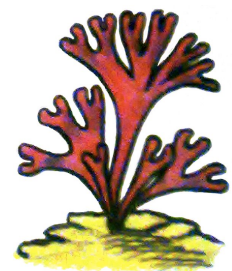
Protocol for chlorophyll fluorescence measurement in green plants.  $F_0$  : minimum fluorescence,  $F_M$  : maximum fluorescence,  $F_S$  : steady-state fluorescence, NPQ : non-photochemical quenching, PQ : photochemical quenching, qE : heat dissipation of energy, qI : photoinhibition, qT : state transition (adapted from Eberhard *et al.*, 2008 [116]).

suivie de quelques minutes d'éclairement intense - cette lumière est dite actinique, car elle permet de déclencher la photosynthèse - puis d'une remise à l'obscurité. Tout au long de l'expérience sont superposés des flashes à intervalles réguliers. On observe une décroissance de  $F_M$  tout au long de l'éclairement : c'est le quenching. Une part est dédiée à la photosynthèse (PQ), l'autre correspond à un NPQ. À l'obscurité,  $F_M$  relaxe progressivement et selon la cinétique de cette relaxation, il est possible d'identifier plusieurs composantes du NPQ :

- la première, qui relaxe en quelques dizaines de secondes, est attribuée à la dissipation thermique qE,
- la deuxième, de quelques minutes, serait due aux transitions d'état qT (chez certaines cyanobactéries il serait de l'ordre de quelques dizaines de secondes)
- la troisième, dont la relaxation est de l'ordre de quelques heures, correspond à la photoinhibition qI.

Cette dernière résulte de la destruction de protéines des PSII par une lumière excessive par rapport à la capacité de neosynthèse de ces protéines. La mesure de  $F_V/F_M$  sur une plante acclimatée à l'obscurité peut refléter finalement si elle est photoinhibée ou non.

Ce type de protocole est adapté aux plantes et algues vertes et fonctionne relativement bien sur les brunes (un exemple d'utilisation de la mesure de  $F_V/F_M$  dans Dittami *et al.*, 2009 [145]). Concernant les algues rouges, il s'est révélé inadéquat et a dû être entièrement revu, ce qui sera détaillé dans la section Résultats 3.4.



## 1.6 Avant d’attaquer...

La thématique initiale de ma thèse proposait d’étudier les effets du stress environnemental chez *C. crispus*, en utilisant une approche à grand échelle, la transcriptomique. Afin de déterminer sur quel(s) type(s) de stress(s) j’allais me focaliser, j’ai voulu effectuer des expériences de calibration visant à évaluer quelles conditions étaient suffisamment “stressantes” pour obtenir une réponse de la part de l’algue, sans pour autant être létales. J’ai alors simulé des stress en laboratoire puis j’ai mesuré des paramètres relatifs à la photosynthèse à l’aide d’un fluoromètre, afin d’évaluer l’impact du stress sur le fonctionnement de la photosynthèse chez *C. crispus*. D’abord par des mesure simple de  $F_V / F_M$ , et des *light curves*, c’est à dire la vitesse du flux d’électrons en fonction de l’intensité lumineuse. Pour des raisons alors inexpliquées, je n’arrivais pas à obtenir des résultats cohérents. Une opportunité de collaboration avec le laboratoire “Physiologie Membranaire et Moléculaire du Chloroplaste” à l’IBPC à Paris, spécialisé dans la photosynthèse, a permis de mettre en lumière un fonctionnement inattendu des phénomènes de régulation de la capture lumineuse chez *C. crispus* expliquant en partie les difficultés rencontrées lors mes expériences de fluorescence. À défaut de pouvoir calibrer mes conditions de stress, cela a éveillé un nouvel intérêt d’étude ciblé avec un aspect physiologique, complémentaire aux études globales de transcriptomique.

## 2 Matériels & Méthodes

### 2.1 Les algues

**Séquençage** Les algues ayant servi au séquençage du génome sont issues de la souche Peggy's Cove, un gamétophyte de *C. crispus* Stackhouse (Gigartinales, Rhodophyta) originaire du Canada, récolté par Juan Correa en 1985. La souche est maintenue en culture au laboratoire sous bullage à une température de 13 °C, en cycle jour/nuit 12h/12h à une intensité lumineuse de 100  $\mu\text{mol}$  de photons·m<sup>2</sup>·s<sup>-1</sup>.

**Transcriptomique** Pour les expériences de transcriptomique, des thalles de *C. crispus* ont été collectées sur le terrain, à proximité du Port de Bloscon à Roscoff (Bretagne). D'une part, trois lots d'algues fertiles ont été constitués :

- 7,9 g de gamétophytes femelles,
- 7,5 g de gamétophytes mâles,
- 5,8 g de tetrasporophytes.

L'ensemble a été ramassé au même moment - au printemps 2010, en début d'après-midi et à marée basse - et immédiatement congelé dans de l'azote liquide. Pour les expériences de RT-qPCR complémentaires, trois lots de chacun des stades ont été récoltés au même endroit, mis en culture dans des bacs d'eau de mer naturelle à une température de 13 °C pour une nuit puis congelés dans l'azote liquide au matin.

D'autre part, des lots de gamétophytes non fertiles ont également été collectés en juillet 2011 et placés *in situ* dans des bacs d'eau de mer, visant à reproduire des cuvettes de marée basse. Quatre conditions ont été mimées et présentées dans le tableau 2.

Pour les conditions 200 % NSW, de l'eau de mer naturelle a été enrichie en NaCl à raison de 33 g.l<sup>-1</sup>. Pour les conditions 35 % de soleil, un grillage a été placé au dessus des bacs. La transmission de la lumière par le grillage a été vérifiée par spectrophotométrie. Toutes les longueurs d'onde sont filtrées à 35 % de manière uniforme entre 350 et 800 nm.

Les algues ont été congelées après trois temps d'exposition :

- immédiatement après la récolte (t = 0)
- après 3 h d'exposition (midi solaire, marée basse)
- après 5 h d'exposition (après-midi, marée montante)

TABLE 2: Conditions de culture et durées d'exposition.  
Culture conditions and duration.

Nom/Name	Condition	Durée/Duration
T10	100 % NSW - 100 % soleil/sunlight	0 h
S+L-	200 % NSW - 35 % soleil/sunlight	3 h
S-L-	100 % NSW - 35 % soleil/sunlight	3 h
S+L+	200 % NSW - 100 % soleil/sunlight	3 h
S-L+/T13	100 % NSW - 100 % soleil/sunlight	3 h
T15	100 % NSW - 100 % soleil/sunlight	5 h

Cette expérience a été réalisée trois fois de manière indépendante, sur trois jours consécutifs où la marée basse est autour de 13h.

**Photosynthèse** Les expériences de spectrofluorométrie ont été réalisées à l'IBPC à Paris, avec des gamétophytes non fertiles collectés à proximité du port de Bloscon. Les algues ont été maintenues en culture dans des bacs d'eau de mer naturelle à une température de 13 °C, en cycle jour/nuit 12 h/12 h.

## 2.2 Biologie moléculaire & bioinformatique

### 2.2.1 Séquençage du génome et annotation automatique

Le génome de *C. crispus* a été séquençé selon la stratégie du *whole genome shotgun sequencing*. Les données ont été générées par un séquençage *paired-end* d'inserts clonés par la technologie Sanger sur un séquenceurs ABI3730xl. La couverture des données brutes est de 14x à partir de d'une librairie d'inserts de longueurs variables de 3 à 20 kb.

L'assemblage a été réalisé avec Arachne générant d'abord 1 266 *scaffolds* (ou super-contigs), puis après élimination des contaminations bactériennes, 921 *scaffolds* ont été retenus. L'annotation automatique a été réalisée après masquage des séquences répétées (soit 3 %) en alignant le génome sur la base de données Uniprot, en utilisant BLAT pour un alignement stringent puis des paramètres plus souples pour les régions n'ayant pas été alignées par BLAT. Les frontières exons/introns ont été affinées par Genewise.

Parallèlement, une librairie de ADNc pleine longueur a été construite et séquençée par la technologie 454 GS-FLX, produisant 333 863 lectures utiles. Ces séquences ont été alors alignées sur le génome avec BLAT puis Est2genome, permettant aux logiciels de prédiction de gènes Geneid et SNAP de générer des modèles de gènes *ab initio*. Les alignements avec des banques de



séquences d'ARNm d'algues rouges issues des bases de données EMBL ont permis de compléter l'annotation générant les 9 658 prédictions finales avec le programme GAZE, qui utilise toutes les données précédemment décrites pour prédire les modèles de gènes.

L'annotation manuelle sera détaillée dans la partie Résultats 3.1.2.

### 2.2.2 Transcriptomique - stades de vie

**Extraction des ARNs pour les grandes quantités d'algues.** Les trois lots d'algues congelées, pour l'expérience "stade de vie", sont broyées finement dans de l'azote liquide et reprises dans un tampon d'extraction à raison de 10 ml·g<sup>-1</sup> de masse fraîche. Les ARNs sont extraits par séparation de phase en présence de chloroforme, puis la phase aqueuse est précipitée au chlorure de lithium. Une seconde extraction au chloroforme est réalisée, suivi d'une précipitation à l'éthanol en présence d'acétate de sodium (selon le protocole de Apt [146]).

**Séquençage - CNS (Evry).** Une librairie de ADNc a été construite par la plateforme CNS (Evry) selon les instructions du fabricant (Illumina). Les ARN messagers sont d'abord isolés des ARNs totaux à l'aide de billes magnétiques portant des étiquettes oligo(dT). Ils sont ensuite fragmentés puis retrotranscrits en ADNc à l'aide d'amorces aléatoires. Des adaptateurs de séquençage sont ligués au fragments de ADNc qui sont par la suite séquencés par un séquenceur nouvelle génération Solexa (Illumina).

**Assemblage et comptage - ABIMS (FR2424), Roscoff.** Les séquences ont été nettoyées sur la base de la qualité avec le logiciel Btrim [147] puis débarrassées des adaptateurs avec Cutadapt [148], celles de mauvaise qualité ont été éliminées puis alignées sur la version 2 de prédictions de gènes issues du génome. Seules les séquences correspondant à un seul modèle de gène ont été prises en compte. Le nombre de séquences par gène a été ensuite quantifié à l'aide du package Python HTSeq-count ([www-huber.embl.de/users/anders/HTSeq/](http://www-huber.embl.de/users/anders/HTSeq/)).

### 2.2.3 Transcriptomique - stress

**Préparation de la banque d'ARNm - MGX (Montpellier).** La banque d'ARNm est préparé par la plateforme MGX à l'aide du kit *TrueSeq RNA sample preparation Kit* d'Illumina, selon les spécifications du fabricant : les ARNs poly-adenylés sont sélectionnés à l'aide de billes magnétiques portant des oligo(dT) puis fragmentés chimiquement. le premier brin de ADNc est

synthétisé avec le kit SuperScript II et des amorces aléatoires, puis le second brin de ADNc est synthétisé et les extrémités des fragments double-brins sont réparés. Une base A est ajoutée à ces extrémités permettant de liguer des adaptateurs. Une étape d'amplification par PCR permet de sélectionner les fragments portant les deux types d'adaptateurs.

**Séquençage - MGX.** La banque de fragments d'ADN est hybridé à une *Flow-cell* (lame de verre sur laquelle est réalisé le séquençage) puis chaque fragment est amplifié pour obtenir des *clusters* (populations clonales de fragments d'ADN). Lors de cette étape des "ponts" d'ADN sont formés et sont ensuite linéarisés et les extrémités libres des fragments sont bloquées. Les primers de séquençage sont ensuite hybridés sur ces extrémités.

Le séquençage proprement dit est réalisé sur un *HiSeq 2000* d'Illumina, grâce la technique de séquence par synthèse. Chaque nucléotide incorporé est marqué d'un fluorochrome, excité au moyen de 2 lasers et détecté par 4 caméras numériques.

**Mapping (MGX) et comptage.** L'alignement des séquences (ou lectures) sur le génome brut de *C. crispus* est réalisé par le logiciel CASAVA V1.8.1 fourni par Illumina. Seules les lectures qui s'alignent de manière unique, c'est à dire à un seul endroit du génome, sont conservées. Le comptage du nombre de reads par gène est effectué à partir du fichier d'annotation généré par le Génoscope, contenant les coordonnées de chaque gène, à l'aide du package Python HTSeq-count ([www.huber.embl.de/users/anders/HTSeq/](http://www.huber.embl.de/users/anders/HTSeq/)).

#### 2.2.4 RT-PCR quantitative

**Extraction des acides nucléiques.** Les ARNs et ADNs de chaque échantillon sont extraits respectivement à l'aide des kits RNeasy et DNeasy plant (Qiagen). Les algues sont d'abord broyées finement dans l'azote liquide, la poudre est resuspendue dans le tampon d'extraction fourni dans le kit, puis une étape de centrifugation est ajoutée de manière à éliminer un maximum de débris cellulaires et faciliter le reste de l'expérience, dont la suite du protocole se déroule selon les spécifications du fabricant.

**Quantification et retrotranscription.** L'ARN de chaque lot est quantifié au spectrophotomètre Nanodrop 2000 (Thermo®) et son intégrité est vérifiée par électrophorèse capillaire à l'aide du Bioanalyzer 2100 (Agilent) selon les instructions fournies par le fabricant. Les ARNs sont dilués

de manière à avoir la même concentration pour chaque échantillon puis retrotranscrits en ADNc à l'aide du Superscript II RT kit (InVitrogen).

**Quantification par PCR.** Les expériences de PCR quantitative sont réalisées en utilisant le SYBR Green PCR master kit (Applied Biosystems) et un thermocycleur Chromo 4 (Bio-Rad).

Les amorces utilisées pour les réactions sont contrôlées par le logiciel Primer Express (Applied Biosystems). Les séquences pour chaque gène sont décrites dans le tableau 29 et la table 2 dans le pré-manuscrit de l'article présenté dans la section 3.3.

Le programme de PCR utilisé est le suivant :

10 min de dénaturation à 95°C,

puis les 40 cycles suivants :

- 15 s de dénaturation à 95°C,
- 60 s d'hybridation-élongation à 60°C.

Une courbe de standardisation est générée à l'aide d'une série de cinq dilutions d'ADN génomique (tableau 3 section 3.3), sachant que la masse d'une copie de génome haploïde est de 0,16 pg. En considérant qu'il n'existe qu'une seule copie de chaque gène dans le génome, le nombre de copies de gène est déterminé en équivalents génome pour chaque dilution.

TABLE 3: Gamme étalon utilisée pour les expériences de PCR quantitative.  
Standard scale used for quantitative PCR experiments

Quantité d'ADN par puits (ng)	0,005	0,028	0,17	1	6
Nombre de copies (équ. génome)	29	174	1 042	6 250	37 500

Ces valeurs sont utilisées comme référence pour quantifier les ADNc de chaque gène pour chaque échantillon.

Le principe de la réaction de PCR quantitative est de mesurer la fluorescence émise à chaque cycle de PCR, puis de déterminer le nombre de copies des gènes testés à partir de la gamme d'étalonnage. Afin de vérifier la spécificité des amorces, une courbe de dissociation est réalisée en fin de PCR. L'efficacité des amorces est calculée d'après la pente des droites de standardisation. Une réaction sans matrice est réalisée pour chaque couple d'amorce pour déterminer le bruit de fond. Chaque combinaison de matrice - couple d'amorces est réalisé en triplicats techniques en plus des triplicats biologiques. Les valeurs brutes étant par la suite moyennées, corrigées en retranchant la quantité d'une séquence se trouvant à cheval sur un intron, permettant de

soustraire les éventuelles contaminations d'ADN génomique et normalisées par rapport à IF4A-2, qui a été déterminé comme meilleur gène de normalisation (ce qui sera décrit en détail dans la partie 3.3).

## 2.3 Biophysique

Pour les études de photosynthèse, les algues ont été adaptées à l'obscurité pendant 1h avant les mesures, de manière à limiter tout phénomène de photoprotection et oxyder au maximum les accepteurs d'électrons. Les variations de fluorescence ont été analysées sur des thalles coupés à la taille du porte-échantillons. Durant les expériences les échantillons ont été placés sous bullage avec de l'air saturé en humidité.

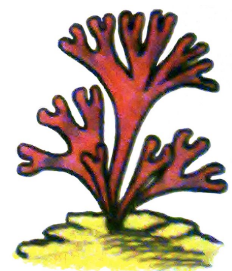
**Fluorescence** Les mesures de fluorescence et d'absorption ont été effectuées avec un spectrophotomètre JTS (Biologic). La fluorescence est échantillonnée avec des flashes de très courte durée ( $4\ \mu s$ ) ayant un effet actinique négligeable (c'est à dire, suffisamment faible pour ne pas modifier de manière significative l'état redox de l'appareil photosynthétique). Le rendement de fluorescence  $F_M$  est mesuré à la fin d'un pulse de lumière saturante rouge de 200 ms, induisant la réduction totale de  $Q_A$ . La lumière actinique est fournie par une LED dont le pic d'émission est à 520 nm et de demi-bande  $\sim 25$  nm, excitant à la fois la chlorophylle et les phycobilisomes. Les flashes de lumière détectrice verte ont une longueur d'onde de 554 nm, qui excite préférentiellement les PBSs, sauf dans la figure 7 (section 3.4) où les mesures sont effectuées à l'aide de flashes à 554 et 440 nm.

- $F_0$  est la fluorescence d'un matériel adapté à l'obscurité détectée par un flash vert de courte durée et de faible intensité.
- $F_M$  et  $F'_M$  représentent la fluorescence maximale mesurée immédiatement après un pulse saturant, respectivement sur un matériel adapté à l'obscurité et préilluminé, détecté deux fois, 100 et 200  $\mu s$  après le pulse.
- $F_S$  est la fluorescence juste avant le pulse saturant.

**Spectrométrie** L'état d'oxydoréduction du photosystème I est mesuré grâce aux changements d'absorbance à 705 nm par des flashes détecteurs de 15  $\mu s$ . L'oxydation de P700 est induite par de la lumière rouge lointain (FR) fournie par une diode dont le maximum d'émission se situe à 740 nm. Le maximum d'absorbance est mesuré en superposant à la lumière FR un pulse rouge

saturant de 20 ms permettant d'oxyder totalement P700.

**Fluorescence à 77 K** Les spectres basse température ont été réalisés sur un spectrophotomètre USB2000 Ocean Optics. Des disques de thalle *C. crispus* sont placés dans un porte-échantillon métallique qui est immergé dans de l'azote liquide. Le porte-échantillon est connecté à une fibre optique en forme de Y qui conduit la lumière d'excitation vers l'échantillon d'une part, et collecte la fluorescence émise d'autre part. La lumière d'excitation est fournie par une lampe Schott KL 1500 et les longueurs d'onde sont choisies par des filtres appropriés.



## 3 Résultats & Discussion

### 3.1 Le génome de *Chondrus crispus*

Un projet de séquençage représente un certains nombres d'étapes-clefs, et malgré l'apparente facilité actuelle, nous verrons qu'il n'est pas toujours évident à mener. L'organisme dont on veut séquencer le génome est tout d'abord un sujet de réflexion important. Les arguments qui ont mené au choix de *C. crispus* sont résumés dans l'introduction 1.1.

Une fois l'organisme choisi, le projet argumenté de séquençage du génome de *C. crispus* a été proposé au centre de séquençage français Genoscope (Evry), via l'appel d'offre annuel début 2006. Un projet pilote a d'abord été réalisée afin d'obtenir une couverture 1x dans le but d'évaluer si le séquençage est possible. Le pilote a été validé et le projet accepté par le Genoscope en septembre 2007. Une librairie d'inserts d'ADN de 3 à 20 kb a été préparée à Roscoff et envoyée au Genoscope. Le séquençage a alors débuté, suivi par l'analyse bioinformatique.

En 2008, la version finale de l'assemblage a été terminée. C'est l'étape critique car plus la taille du génome est grande et plus le génome a subi de duplications, plus elle est difficile. Près d'un an a été nécessaire pour finaliser l'assemblage de *C. crispus*, dont le génome est considéré comme "assez petit et peu dupliqué".

La première version de l'annotation automatique a été achevée mi-2009 et a permis d'initier l'analyse et l'annotation experte, étapes demandant le plus grand effort humain. Un consortium international de près de 40 laboratoires a été mobilisé pour cette tâche (liste en annexe 20 et chaque personne a choisi une catégorie de gènes à annoter selon ses affinités et son expertise. Un atelier a alors été organisé à Roscoff afin que tous les membres du consortium puissent se familiariser avec l'interface d'annotation Apollo [149] et afin de discuter de la stratégie d'annotation. S'en est suivie la rédaction de l'article qui a été soumis en été 2012. La figure 20 résume les étapes du séquençage du génome de *C. crispus*.

Cinq ans se sont donc écoulés depuis le séquençage proprement dit et l'obtention d'un génome exploitable et publiable. Malgré l'amélioration des stratégies de séquençage, des bases de données de plus en plus fournies et des outils bioinformatiques plus performants aidant notamment à l'annotation, cette durée est difficilement compressible pour des génomes eukaryotes.

L'article du génome, soumis à PNAS, est présenté à la fin du manuscrit en section ??.

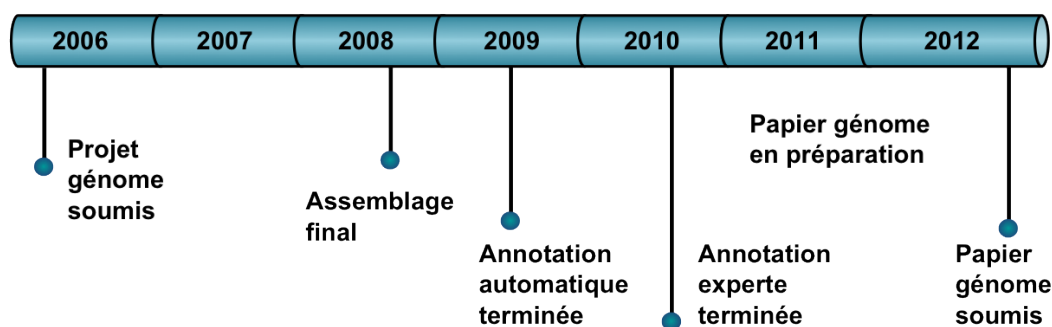


FIGURE 20: Chronologie du projet génome de *Chondrus crispus*.  
Timeline of the *Chondrus crispus* genome project.

### 3.1.1 Le génome en chiffres

La taille du génome de *C. crispus* est de 105 Mbp, et l'assemblage final a permis de générer 925 *scaffolds* (portions d'ADN assemblé), numérotés par ordre de taille, du scaffold 1 de longueur 1,2 Mbp, jusqu'au scaffold 925 long de 2 kb. La dernière version d'annotation prédit 9 658 gènes. Ce nombre est faible en comparaison avec d'autres organismes photosynthétiques (voir la table 4). Par exemple, *Arabidopsis thaliana* dont le génome est de taille similaire, possède trois fois plus de gènes [150] et *Chlamydomonas reinhardtii*, dont le génome est également de même taille, possède plus de 15 000 gènes [37]. *C. Crispus* a un nombre gènes similaire à d'autres algues vertes unicellulaires, telles que *Ostreococcus tauri*[38] ou *Chlorella variabilis*[43], alors que leurs génomes sont très réduits (en taille totale). *C. crispus* possède à peine deux fois plus de gènes que *Cyanidioschyzon merolae* [39], l'algue rouge unicellulaire dont le génome est 6.5 fois plus petit.

En comparant certaines familles de gènes que l'on retrouve normalement chez d'autres organismes, tels que les gènes de ménage (protéines ribosomales, enzymes participant à la glycolyse), il semblerait que presque tous les gènes soient présents chez *C. crispus*, mais en une seule copie, alors que chez *Arabidopsis*, il est fréquent de retrouver plusieurs copies de chaque gène.

La structure du génome de *C. crispus* présente des originalités pour un organisme multicellulaire : une faible quantité d'introns de petite taille, en effet, 88 % des gènes seraient monoexoniques. Cette caractéristique est retrouvée dans les données concernant l'autre génome d'algue rouge, *C. merolae*, qui possède à peine plus d'un exon par gène, et est également une caractéristique de génomes compacts, par exemple *O. tauri*. Mais le génome de *C. crispus* est loin d'être compact, puisque seulement 8 % des séquences sont codantes. Ces séquences sont organisées en denses îlots de gènes, noyés dans de vastes régions intergéniques d'éléments transposables. Un

TABLE 4: Comparaison du génome de *C. crispus* avec d'autres génomes d'algues. [37] [38] [150] [43] [35] [39]. Size : taille du génome, Mpb : million de paire de bases, gene nb : nombre de gènes.  
Comparison of the genome of *C. crispus* with other algal genomes. [37] [38] [150] [43] [35] [39]. Mbp : million of base pairs, gene nb : gene number.

Genome	<i>C. crispus</i>	<i>C. reinhardtii</i>	<i>O. tauri</i>	<i>A. thaliana</i>	<i>C. variabilis</i>	<i>E. siliculosus</i>	<i>C. merolae</i>
Size (Mbp)	105	121	12,6	140,1	46,2	214	16,5
Gene nb	9 658	15 143	7 892	26 341	9 791	16 281	5 331
Exon/gene	1,2	8,3	1,57	5,2	7,3	6,8	1,005

exemple caractéristique de la structure du génome de *C. crispus* est représenté par le scaffold 62 sur la figure 21.

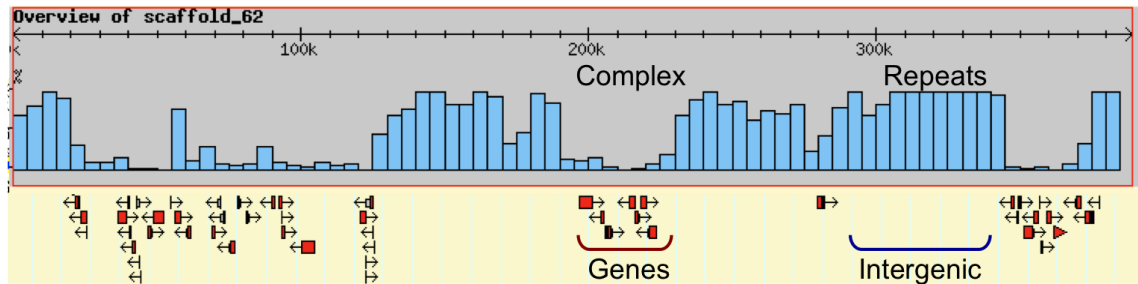


FIGURE 21: Structure du scaffold 62, typique du génome de *C. crispus*. Complex : région de haute complexité contenant les gènes. Repeats : région de faible complexité. Intergenic : région intergénomique.  
Structure of the scaffold 62, typical of the *Chondrus crispus* genome. Complex : high complexity region containing genes. Repeats : intergenic low complexity region.

Les annotations automatiques et manuelles combinées ont permis d'assigner une fonction putative à 2 786 gènes sur les 9 658 modèles de gènes prédits, ce qui ne représente que 29 % du génome. Les deux tiers restants n'ont de correspondance avec aucun des gènes présents dans les bases de données publiques (NCBI [151] et UniProt [183]).

### 3.1.2 L'annotation

L'annotation d'un génome consiste à identifier les gènes présents dans un organisme à partir de sa séquence d'ADN. Dans un premier temps, elle est faite de manière automatique, par une série d'algorithmes et de filtres permettant de prédire les cadres ouverts de lectures et visant à éliminer les séquences improbables, notamment celles qui sont jugées trop courtes. Pour appuyer ces primo-prédictions, les ARN de *C. crispus* ont été séquencés en 454 générant 300 000 EST, en plus des 4000 EST de protoplastes et de thalles séquencées pour une la première expérience



de stress chez *C. crispus* en 2006 [17]. Ces séquences ainsi que d'autres EST d'algues rouges provenant des banques de séquences publiques ont été alignées sur le génome de *C. crispus* afin d'identifier les gènes réellement transcrits. Ceci a permis de vérifier et compléter les premiers modèles de gènes hypothétiques.

En revanche, ce type d'identification peut être erronée ou imprécise : certains gènes, par exemple les plus petits, peuvent passer à travers les filtres et ne pas être reconnus comme des cadres ouverts de lecture. Il est ainsi nécessaire de contrôler le résultat de l'annotation automatique manuellement, gène par gène, il s'agit de l'annotation experte. L'objectif est de corriger ou préciser l'identité de certains gènes, mais aussi de rétablir les structures mal prédites en apportant des modifications au niveau des codons *start* ou *stop*, des sites d'épissage, voire de changer tout le cadre de lecture.

Deux types d'annotation, complémentaires, peuvent être mises en oeuvre et seront décrites dans les deux paragraphes qui suivent.

**Approche "à l'aveugle"** Cette approche, sans *a priori*, permet de favoriser les chances de découvrir des éléments originaux dans le génome. Trois autres annotateurs ont également effectué ce type d'approche. Chacun a choisi de parcourir arbitrairement un ou deux scaffolds : le scaffold 68 a été annoté par Simon Dittami, le scaffold 200 par Jonas Collén et enfin les scaffolds 119 et 120 ont été annotés par Claire Gachon. J'ai choisi le *scaffold* 62, dont la longueur est de 389 098 paires de bases et contient 53 prédictions de gènes. Chaque prédiction a été alignée contre la banque de données NCBI [151] et lorsque l'homologie avec un gène connu était suffisamment probante, une fonction a été assignée au modèle prédit. Le tableau 31, en annexe, liste l'ensemble des gènes prédits.

Au total, 20 prédictions correspondent à des gènes déjà connus dans les banques, dont 5 déjà annotées lors de la recherche de gènes relatifs au stress. Cinq prédictions n'ont pas pu être validées avec précision mais contiennent des domaines connus. Enfin, 24 prédictions n'ont pas pu être validées, tous sauf une n'ont aucune correspondance dans les banques de données et donc ont été annotées comme des protéines hypothétiques.

Malgré la grande quantité de génomes séquencés et le nombre de protéines dont les fonctions ont été identifiées, notamment chez des organismes qui paraissent évolutivement proches, le génome de *C. crispus* recèle de nombreuses inconnues.

**Approche avec *a priori* : les gènes liés au stress.** J’ai utilisé, en parallèle, une approche avec *a priori* où l’objectif était d’analyser le génome de *C. crispus* à la recherche de gènes d’intérêt. Ma thématique étant relative au stress oxydatif, mon équipe (génomique fonctionnelle et stress abiotique) m’a donc confiée l’annotation des gènes acteurs du métabolisme oxydatif et d’autres gènes connus dans la bibliographie pour être impliqués dans la réponse au stress, notamment chez les végétaux terrestres et les straménopiles.

Les figures en annexe 32 et 33 présentent la liste des gènes identifiés dans le génome. Je les ai replacés, lorsque cela est possible, dans leur contexte métabolique.

Ainsi, le cycle “eau-eau”, présenté dans l’introduction en figure 10, montre comment la cellule peut transformer l’oxygène radicalaire hautement réactif, produit notamment par l’appareil photosynthétique, en composant non toxique pour la cellule : l’eau. Les acteurs de ce cycle de détoxification sont les ROSEs (ROS Scavenging Enzymes), enzymes capturant les espèces réactives de l’oxygène, en catalysant une série de réactions d’oxydoréduction et sont donc présentes dans le génome de *C. crispus*. Par extension, j’ai également identifié d’autres types d’oxydo-réductases, avec des activités potentiellement antioxydantes, dont les fonctions ne peuvent pas toujours être définies d’après leur seule séquence protéique. Leur nom reste donc volontairement généraliste.

Le stress oxydatif n’est cependant pas la seule contrainte que doit subir la cellule. En cas de température trop élevée ou de dessiccation, l’équilibre thermodynamique de la cellule est mis en jeu, pouvant avoir pour conséquence un mauvais repliement de certaines protéines. Il existe donc une catégorie de protéines, dites chaperonnes, dont la fonction est d’assurer le bon repliement des autres protéines. Elles ont également un rôle important en cas de forte production de protéines.

La première famille contient les Heat Shock Proteins (HSP), initialement appelées ainsi car elles sont exprimées en quantité très importante lors d’un choc thermique. Elles sont nommées selon leur masse moléculaire, de 10 à 100 kDa ou parfois selon leur fonction. Il existe des systèmes de chaperonnes, agissant de concert, permettant de replier les protéines dès la fin de leur traduction. Les systèmes suivants ont été bien décrits chez les procaryotes et possèdent généralement des homologues chez les eucaryotes [152]. La figure 22 résume le système de repliement des protéines à la sortie du ribosome, tel qu’il est décrit chez *Escherischia coli*.

DnaJ, DnaK et GrpE agissent séquentiellement pour replier les protéines dénaturées suite à un stress hyperosmotique ou thermique. La protéine mal repliée est d’abord liée à DnaJ. Le

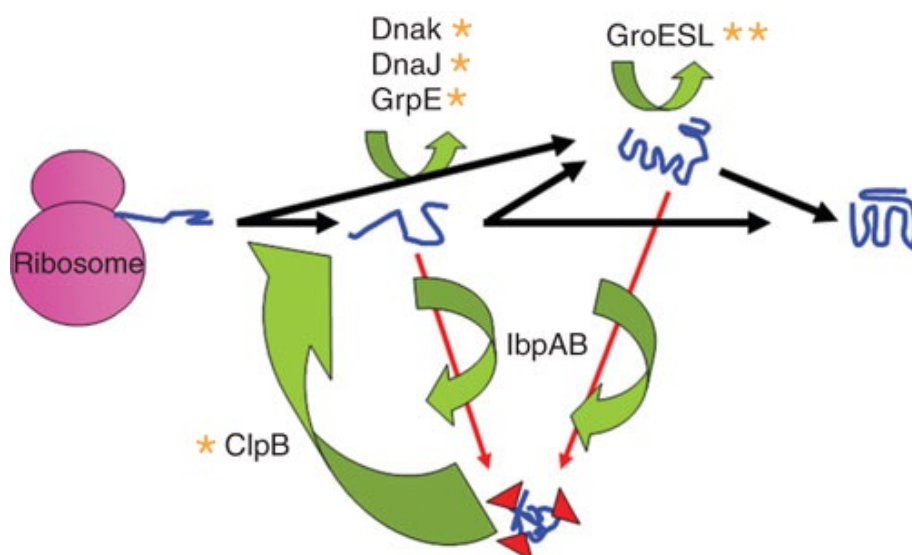


FIGURE 22: Système de repliement post-traductionnel procaryote (de Marco, 2007 [153]). Les étoiles représentent les protéines dont les gènes correspondant sont retrouvés dans le génome de *C. crispus*.  
Post-translational folding system in procaryotes (de Marco, 2007 [153]). Stars indicate the proteins which corresponding genes can be found in the genome of *C. crispus*.

complexe est stabilisé par l'hydrolyse d'une molécule d'ATP par DnaK. GrpE libère l'ADP lié à DnaK, permettant à un nouvel ATP de se lier à DnaK et de libérer la protéine en cours de repliement. Plusieurs répétitions de ce cycle sont nécessaires pour replier efficacement une protéine.

Il peut être assisté du complexe de chaperonines GroES et GroEL, ainsi que de ClpB. Les petites HSP, IbpA et IbpB, n'ont pas été identifiées dans le génome de *C. crispus*. Il existe en revanche d'autres petites HSP de taille similaire, pouvant éventuellement remplacer le couple IbpAB.

Le génome de *C. crispus* contient également tous les acteurs du complexe TCP1 (*tubulin complex protein*), qui semblerait être impliqué dans la synthèse de la tubuline, constituant majeur du cytosquelette cellulaire [154] et considéré comme un équivalent eukaryote du complexe GroES-GroEL [155].

Un autre complexe de chaperonnes a été identifié chez *C. crispus*, le complexe TBC (Tubulin-binding cofactors), impliqué aussi dans la mise en place de la tubuline. Tous les partenaires sauf un, TBCA, ont été identifiés dans le génome. Il s'agirait d'un peptide de petite taille (moins de 100 acides aminés) ce qui handicape la recherche par alignement de séquences.

La comparaison du génome de *C. crispus* et celui d'*Arabidopsis thaliana*, montre que pour des tailles de génomes similaire, respectivement 105 Mb et 140 Mb, *A. thaliana* possède un nombre estimé de gènes trois fois supérieur à celui de *C. crispus*. Si l'on s'intéresse plus précisément à chaque catégorie de gènes relatifs au stress (voir la table 5), ces gènes semblent toujours beaucoup plus nombreux chez *A. thaliana* :

- neuf ascorbate peroxidases contre deux chez *C. crispus*,
- cinq MDHAR et six TXR contre une seule copie de chaque chez *C. crispus*,
- une trentaine de glutaredoxine contre quatre seulement chez *C. crispus*.

Il en est de même pour les chaperonnes et heat shock proteins. Ainsi pour effectuer les mêmes fonctions, *C. crispus* révèle beaucoup moins de redondance fonctionnelle, comme observé chez *C. merolae*[156][39].

Il est possible que les protéines de *C. crispus* soient plus polyvalentes que celles des végétaux terrestres. Les algues possèdent en effet des tissus dont l'architecture est plus simple et homogène que celle des plantes. Elles n'ont pas de réel système vasculaire ni de cuticule et leur thalle n'est pas organisé en asymétrie abaxiale/adaxiale, seuls les organes reproducteurs possèdent des cellules différenciées. Cette simplicité apparente des tissus pourrait rendre inutile une spécialisation des composants cellulaires. Paradoxalement, *E. siliculosus*, dont l'architecture en filaments unisériés est encore plus simple, possède un génome deux fois plus grand et des gènes souvent dupliqués. De plus, il ne faut pas oublier que seule l'expérience permettra de valider l'expression de toutes ces protéines.

Si l'on s'intéresse à l'origine phylogénique des gènes de *C. crispus*, on s'aperçoit qu'elle peut être très variable selon la catégorie de gènes considérée et ne respecte pas l'arbre consensus de Adl [2] qui a été réalisé à partir de quelques gènes seulement, considérés comme des marqueurs évolutifs. Par exemple on remarque l'existence de deux types de superoxide dismutases, à cuivre et manganèse, tout comme chez *A. thaliana*. Seulement les SOD à manganèse se rapprochent davantage de celles des straménopiles et semblent avoir été transférées aux straménopiles lors de l'endosymbiose secondaire, où une algue rouge aurait été phagocytée par une autre cellule eukaryote pour donner l'ancêtre des algues brunes et diatomées.

Chez d'autres algues rouges, telles que *Porphyra*[157], dont le génome n'est pas connu, ou *Cyanidioschizon*[39], il n'y aurait pas de SOD à cuivre, alors que deux types, à cuivre et manganèse, seraient présents chez *C. crispus*. Ces SOD à cuivre ont des fortes similitudes avec

TABLE 5: Comparaison des gènes relatifs au stress entre les génomes de *C. crispus*, *A. thaliana* et *C. merolae* [150][39]. ROS scavenging enzyme : enzyme de capture des espèces réactives de l'oxygène. Mpb : million de paire de bases,  $\Sigma$  ROSEs : nombre total d'enzymes ROSE,  $\Sigma$  Cpn : nombre total de chaperones,  $\Sigma$  genes : nombre total de gènes prédits.  
Comparison of genes related to stress between the genomes of *C. crispus*, *A. thaliana* and *C. merolae*. [150][39]. Mbp : million of base pairs,  $\Sigma$  ROSEs : total number of ROSEs,  $\Sigma$  Cpn : total number of chaperones,  $\Sigma$  genes : total number of gene models.

ROS scavenging enzyme (ROSE)	<i>C. crispus</i>	<i>A. thaliana</i>	<i>C. merolae</i>
ascorbate peroxidase (APX)	2	9	1
Cu/Zn superoxide dismutase (SOD)	2	4	0
Fe/Mn superoxide dismutase (SOD)	2	5	3
catalase (CAT)	2	3	1
glutathione reductase (GR)	1	2	1
monodehydroascorbate reductase (MDHAR)	1	5	1
dehydroascorbate reductase (DHAR)	2	4	0
glutathione peroxidase (GPX)	2	8	0
peroxiredoxin (PRX)	3	6	3
glutaredoxin (GRX)	4	$\pm 30$	2
thioredoxin reductase (TRxR)	1	6	3
$\Sigma$ ROSEs	24	83	15
<b>Chaperones (Cpn)</b>			
HSP 100 (ClpB)	1	3	2
HSP 90	3	9	4
HSP 70	6	17	3
HSP 60	2	3	2
HSP 40 (DnaJ)	$>7$	93	$>6$
HSP 20 (small HSP)	5	16	2
Protein disulfide isomerase	4	23	3
Peptidylprolyl isomerase	9	45	11
Calnexin	1	3	1
$\Sigma$ Cpn	36	212	28
Genome size (Mbp)	105	140	16,5
$\Sigma$ genes	9 658	26 341	5 331

des SOD de bactéries marines (*Actinosynnema mirum* et *Micromonospora* sp.). Il pourrait les avoir acquises par transfert horizontal à partir de ces bactéries. D'autres enzymes, telles que la catalase ou la glutathione reductase présentent également de fortes similitudes avec des enzymes bactériennes. Par ailleurs, la glutathione peroxidase et l'une des peroxidases sont très proches de leurs homologues cyanobactériens, reflétant probablement une acquisition lors de l'endosymbiose primaire.

Analyser les structures de gènes chez *C. crispus* a également été l'occasion d'essayer d'identifier les peptides signaux. Ces séquences d'une vingtaine d'acides aminés se trouvent le plus souvent dans la partie N-terminale de la protéine, et permettent d'adresser cette dernière à un compartiment cellulaire spécifique. Il existe ainsi des peptides signaux pointant vers la mitochondrie, le chloroplaste ou le reticulum endoplasmique. Parmi les résultats que j'ai obtenus, il semblerait que certaines enzymes connues pour être chloroplastiques chez d'autres végétaux, n'aient pas de signal d'adressage chez *C. crispus*. C'est le cas notamment de l'ascorbate peroxidase et de la dehydroascorbate reductase. Pour déterminer si mes protéines sont exportées dans des organites particuliers, j'ai utilisé des programmes dont les algorithmes ont été conçus à partir de données consensus des banques publiques (par exemple : SignalP3.0 [158]) qui pourraient reconnaître uniquement les signaux ciblant vers la mitochondrie, ou tout simplement ne pas être adaptés du tout à notre génome, ce qui pourrait expliquer en partie les différences d'adressage observées.

Le génome de *C. crispus* possède ainsi l'ensemble des gènes impliqués dans la réponse au stress connus chez d'autres organismes, leur origine est hétérogène, certains gènes étant plus proches de gènes prokaryotes, d'autres plutôt eukaryotes. Il n'a cependant pas été possible d'identifier avec certitude l'adressage des protéines correspondant aux gènes annotés.

**Approche avec *a priori* : les antennes collectrices de lumière.** Pour poursuivre l'annotation manuelle du génome et dans le but de caractériser l'appareil photosynthétique de *C. crispus*, l'annotation du génome plastidial et des gènes relatifs à la photosynthèse a été réalisé avec Catherine Leblanc (UMR 7139) et Frédéric Partensky (UMR 7144). Un récapitulatif des gènes annotés est présenté en annexe dans les tableaux 21 à 28.

L'architecture de l'antenne du photosystème II chez les algues rouges est variable selon les espèces, elle peut être hemidiscoïdale comme chez *Rhodella violacea*, hemiellipsoïdale comme chez *Porphyridium cruentum* (des unicellulaires extrêmophiles), ellipsoïdale comme pour *Antithamnion glanduliferum* (une florideophycea comme *C. crispus*) ou résulte d'un mélange de plusieurs types, par exemple : ellipsoïdale et hemidiscoïdale chez *Porphyra umbilicalis*[159][160]. Deux exemples de phycobilisomes sont présentés dans la figure 23, la première étant complexe avec plusieurs bras organisés de manière hémidiscoïdale, l'autre étant plus simple avec uniquement un seul bras [161]. L'architecture des phycobilisomes de *C. crispus* n'a pas été décrite jusqu'à présent, mais les gènes codants pour les phycobiliprotéines sont nombreux et variés,

particulièrement les linkers et semblent indiquer que le phycobilisome de *C. crispus* possède une structure sophistiquée.

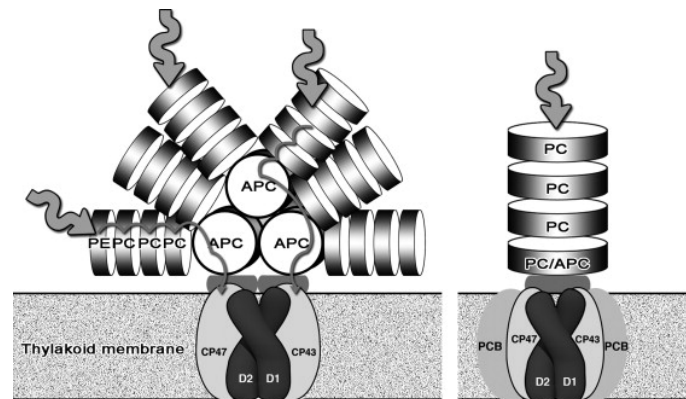


FIGURE 23: Schéma d'un phycobilisome hémidiscoïdal (à gauche) et d'un complexe de capture lumineuse en forme de tige (à droite)[161].  
Scheme of a hemidiscoidal phycobilisome (left side) and of rod shaped phycobiliprotein light harvesting complex (right side)[161].

L'antenne du PS I est constituée de sept sous-unités de *chlorophyll-binding proteins*. Elles présentent peu d'identité (30-40 %) avec les trois sous-unités - appelées LHCr pour *red light-harvesting complex* (complexe de capture lumineuse rouge) - de *Cyanidioschyzon merolae* et sont encore plus éloignées des antennes d'autres organismes photosynthétiques non rouges, suggérant une histoire évolutive divergente et une organisation autour du PS I différente.

Le génome de *C. crispus* possède de nombreux autres gènes codant pour des protéines capables de lier la chlorophylle, telles que les *high-light induced proteins* (protéines induites par les fortes lumières), HLIPs, protéines à une hélice transmembranaire, connues chez les cyanobactéries pour leur rôle dans la photoprotection [162] et/ou le stockage de la chlorophylle [163]. Jusqu'à présent les gènes codant pour les HLIPs étaient plastidiaux [164]. Parmi les sept gènes codants pour des protéines de la famille HLIP identifiés chez *C. crispus*, un seul est codé par le génome plastidial, les autres étant nucléaires.

Chez les plantes terrestres, il existe d'autres protéines transmembranaires, liées à la photoprotection, à deux hélices (les SEPs), trois hélices (les ELIPs) et les PsbS à quatre hélices[165]. Chez *C. crispus*, seule une ELIP a été identifiée dans le génome nucléaire.

L'annotation des gènes relatifs aux antennes de capture lumineuse chez *C. crispus* confirme que l'appareil photosynthétique des algues rouges présente des spécificités par rapport à celui des végétaux verts, ainsi qu'à celui des cyanobactéries. De plus au sein des algues rouges, il n'existe pas de structure d'antenne consensus, puisque les phycobilisomes et les LHCr sont différents

d'une espèce à l'autre.

**Approche avec *a priori* : les pigments.** En ce qui concerne la synthèse de pigments impliqués dans la photosynthèse, *C. crispus* possède tous les gènes nécessaires à la biosynthèse de la chlorophylle, à l'exception notable des gènes impliqués dans la voie indépendante de la lumière, suggérant que la chlorophylle ne peut être produite que durant la journée, contrairement à *Porphyra purpurea* et *P. yezoensis*, qui possèdent les trois gènes de la voie indépendante de la lumière : *chlB*, *chlL* et *chlN*.

Les gènes relatifs à la biosynthèse des caroténoïdes, existant chez les plantes terrestres et chez *Cyanidioschyzon merolae*, ont été identifiés dans le génome de *C. crispus*, sauf les deux enzymes du cycle des xanthophylles : la violaxanthine dé-époxydase et la xanthophylle époxydase, ne permettant pas la conversion violaxanthine-zeaxanthine.

Comme pour les antennes, l'annotation des gènes relatifs à la biosynthèse des pigments photosynthétiques dans *C. crispus* montre qu'il existe une variabilité entre les espèces d'algues rouges.

### 3.2 Le transcriptome de *Chondrus crispus*

Dans le cadre du séquençage du génome de *C. crispus*, une expérience de séquençage des ARN (RNAseq) a été réalisée à partir d'échantillons environnementaux des trois stades de vie macroscopiques de *C. crispus*, gamétophytes mâle et femelle et tetrasporophyte. L'objectif était alors de renforcer l'annotation manuelle, en générant davantage d'EST, dans des conditions différentes de celle de la souche qui a servi au séquençage du génome. La technologie Solexa/Illumina utilisée pour cette expérience a permis un séquençage plus profond que celui effectué précédemment en 454, puisqu'elle a fourni 60 millions d'EST, et a l'avantage d'être quantitative. Les modèles de gènes faiblement exprimés ont pu alors être validés. De plus, la profondeur du séquençage a été telle, que des petites contaminations d'ADN ou ARN plastidial ont été amplifiées et ont permis de reconstituer l'intégralité du génome du chloroplaste, dont l'ADN avait été difficile à purifier jusqu'à présent.

La thématique initiale de mes travaux de thèse prévoyait de réaliser des expériences de transcriptomique sur des algues environnementales, pour analyser leur réponse à des stress abiotiques. Le séquençage des trois stades de vie a alors été l'occasion de tester une nouvelle technique séquençage haut-débit pour effectuer de la transcriptomique quantitative chez *C. crispus*, avant



de mettre en oeuvre l'expérience "stress", sur davantage d'échantillons et donc de plus grande envergure.

Afin de caractériser le transcriptome de *C. crispus* de manière générale, pour commencer, j'ai choisi de décrire ensemble les deux jeux de données issus des expériences de RNAseq (stades de vie et stress). Puis une analyse de chacun d'entre eux séparément suivra dans les sections 3.2.2 et 3.2.3.

### 3.2.1 Caractéristiques générales

Les ARNs messagers de 21 lots d'algues ont été séquencés à l'aide de la technologie de séquençage haut-débit Solexa/Illumina. Trois lots d'algues correspondent aux trois stades de vie macroscopiques — gamétophytes mâle et femelle, tetrasporophyte — et les 18 autres lots d'algues ont été échantillonnés à trois moments de la journée et soumis à des conditions de salinité et lumière variables. Vingt millions d'ESTs brutes (lectures) de 32 nucléotides ont été générées pour chacun des trois stades de vie, et entre 31 et 65 millions de séquences de 50 nucléotides pour chacun des autres échantillons. Ces différences sont dues à l'évolution de la technologie entre les deux expériences, les ARN extraits à partir des trois stades de vie ayant été séquencés deux ans avant l'expérience "stress et cycle diurnal".

Ces lectures ont par la suite été nettoyées de leurs adaptateurs et filtrées afin d'éliminer des contaminations d'autres organismes (bactéries, métazoaires, autres algues) et éliminer les lectures trop courtes ou de qualité insuffisante (pour plus de détail, voir la section matériels et méthodes 2. Après alignement sur les prédictions de gènes de *C. crispus*, entre 5 et 14 millions de lectures par échantillon sont ainsi utiles, correspondant 21 à 26 % du nombre total de séquences générées. Le détail est présenté dans la table 6.

Une grande partie des gènes de *C. crispus* est couverte par ces lectures puisque 93 % des modèles de gènes sont représentés par au moins une lecture. La distribution moyenne du nombre de lectures par gène (table 7) est très hétérogène : la gamme dynamique est de  $10^6$ , avec une médiane de 353 lectures par gène, une moyenne de 1498, proche du troisième quartile, et un maximum de 588 525 lectures par gène. La très grande majorité des gènes est donc représentée par "seulement" quelques centaines de lectures.

Une importante partie des gènes est très faiblement exprimée, 19 % des gènes exprimés ayant moins de 50 lectures. La majorité (80 %) de ces gènes ne correspondent à aucune protéine connue dans les bases de données. Les autres gènes, annotés, sont relatifs à des éléments transposables

TABLE 6: Quantité de séquences utiles pour chaque échantillon de *C. crispus*. Sample : nom d'échantillon. Raw reads : lectures brutes. Mapped reads : lectures alignées sur la référence. T10, T13, T15 : échantillons collectés à 10h, 13h et 15h respectivement (cycle diurnal). S- : 100 % de salinité, S+ : 200 % de salinité, L- : 35 % de lumière, L+ : 100 % de lumière. T : tetrasporophyte, M : gamétophyte mâle, F : gamétophyte femelle.

Quantity of usable sequences for each sample of *C. crispus*. T10, T13, T15 : samples collected à 10AM, 1PM and 3PM respectively (diurnal cycle). S- : 100 % of salinity, S+ : 200 % of salinity, L- : 35 % of light, L+ : 100 % of light. T : tetrasporophyte, M : male gametophyte, F : female gametophyte.

Sample	Raw reads	Mapped Reads	% Mapped
T10a	50 835 456	11 357 992	22,3
T10b	38 345 204	8 298 558	21,6
T10c	65 947 135	14 470 137	21,9
T15a	50 567 193	11 682 501	23,1
T15b	44 232 630	10 764 073	24,3
T15c	48 887 212	11 475 911	23,5
S-L+/T13a	42 826 916	10 742 230	25,1
S-L+/T13b	39 715 564	9 718 990	24,5
S-L+/T13c	46 899 030	11 248 171	24,0
S+L-a	54 257 972	13 094 120	24,1
S+L-b	51 758 800	12 199 978	23,6
S+L-c	34 362 380	8 849 234	25,8
S-L-a	42 665 904	9 002 508	21,1
S-L-b	31 318 668	7 134 850	22,8
S-L-c	45 111 502	11 215 119	24,9
S+L+a	44 281 241	9 870 168	22,3
S+L+b	37 164 987	8 861 469	23,8
S+L+c	38 727 101	9 433 606	24,4
T	20 315 375	5 726 424	28,6
M	21 650 365	5 822 764	26,9
F	23 701 397	9 533 402	40,2

(helicases, pol-proteins, transposases), ainsi qu'à des protéines de fonction inconnue. Compte tenu de leur faible expression, il est intéressant de se questionner sur la notion de bruit de fond lorsque l'on réalise une expérience avec une gamme dynamique aussi importante que celle obtenue ici. L'alignement des lectures ayant été fait de manière à exclure tous les lectures ambiguës — c'est à dire, s'alignant à plusieurs endroits du génome — l'expression de certains gènes présentant de très fortes homologues de séquences a pu être sous-estimée. Néanmoins, analyser ces gènes à faible expression permettrait de préciser certaines prédictions en identifiant davantage de pseudo-gènes et améliorer ainsi la qualité du génome de *C. crispus*.

Une petite fraction du transcriptome concerne des gènes dont l'expression est beaucoup plus

TABLE 7: Le transcriptome de *C. crispus* en chiffres.  
The transcriptome of *C. crispus* in numbers.

	Reads	% of genome
Total genes	9 658	
Expressed	9 015	93.3 %
Annotated	2 786	28.9 %
Quartile 1	45	
Median	353	
Quartile 3	1191	
Max	588 525	
Mean	1 498	
$1 \leq \bar{x} < 50$	1 851	19.2 %

élevée que le reste. On peut considérer ainsi les 20 gènes dont l'expression moyenne entre tous les échantillons est supérieure à 50 000 lectures (table 8 et figure 24). La moitié d'entre-eux sont annotés. Parmi eux, l'ARN ribosomal 18S, normalement non poly-adenylé, ne devrait pas se retrouver dans les librairies séquencées, puisque les ARNs ont été filtrés préalablement avec des amorces poly-T qui capturent les queues poly-A des ARN messagers. Deux possibilités peuvent expliquer sa présence : son expression est tellement élevée qu'il parvient à se retrouver dans les banques malgré la filtration, ou la présence effective d'une queue poly-A, pour les destiner à la dégradation, comme observé dans la levure[166]. Il n'a pas été pris en compte pour la normalisation des données et a été écarté des analyses. La sur-représentation des protéines chaperones avec huit heat shock protein semble liée à leur très forte expression dans les échantillons ayant été exposés à la lumière du soleil. Néanmoins, des motifs d'expression peuvent être distingués :

- deux HSP (70 et 90) sont exprimées fortement dans tous les échantillons, même dans ceux qui n'ont pas été exposés au plein soleil. Leur expression devient encore plus forte dans les échantillons collectés après trois heures au soleil.
- six autres HSP (20, 90, 100) sont peu exprimées dans les échantillons congelés immédiatement après récolte. Leur expression devient très importante après 3h en plein soleil.

Trois gènes (CHC\_T00006259001, CHC\_T0000191000 et CHC\_T00002243001) codant pour des protéines non-identifiées suivent le même patron d'expression que les six HSP précédentes. Ceci permet d'émettre des hypothèses concernant les fonctions de ces protéines inconnues, qui pourraient avoir des rôles semblables aux HSP et participer à la protection contre l'excès de chaleur ou de lumière.

Il est possible de distinguer deux groupes d'échantillons dont la variable discriminante serait l'exposition au soleil. Le premier groupe contient des échantillons récoltés le matin et congelés immédiatement (T10a, T10b, T10c, M, F, T) alors que le deuxième groupe contient les autres algues exposées au soleil pendant plusieurs heures.

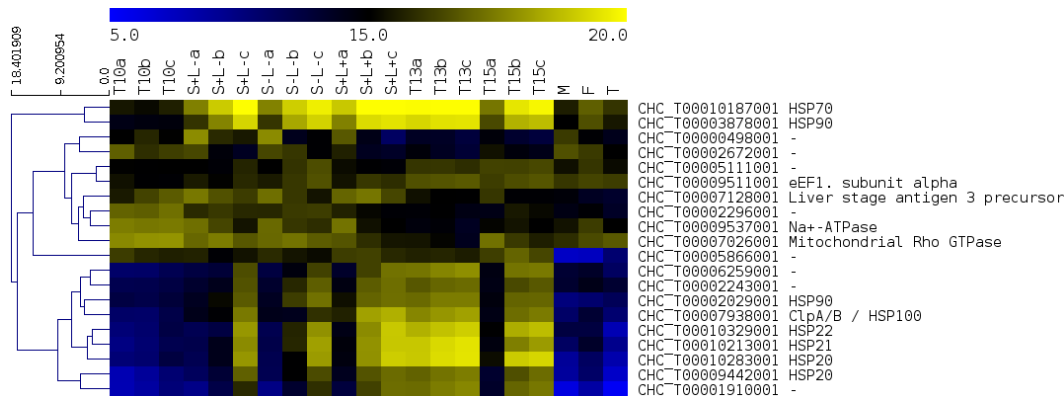


FIGURE 24: Clustering hiérarchique des 20 gènes les plus exprimés chez *C. crispus*.  
Hierarchical clustering of the 20 most expressed genes in *C. crispus*, using TMeV.

A l'issu de cette analyse générale, il est possible de considérer que le séquençage des ARN messagers d'algues environnementales, dont les fonds génétiques ne sont probablement pas identiques au génome de référence, est réalisable. Il aurait été néanmoins intéressant de comparer le taux de couverture et de lectures utiles entre ces échantillons environnementaux et des ARN issus de la souche Peggy's Cove.

### 3.2.2 Les stades de vie de *Chondrus crispus*

Lorsque nous avons reçu les premières données de séquençage, peu d'expériences de transcriptomique avaient été réalisées avec les techniques de séquençage nouvelle génération et rares étaient les publications sur les algues rouges, apportant peu de recul sur les résultats qu'il est possible d'obtenir, surtout sur des échantillons environnementaux. Un des objectifs de la transcriptomique par séquençage et de comparer l'expression des gènes entre différents échantillons. Pour effectuer une expérience robuste, et pouvoir conclure sur des différences d'expressions entre les conditions ayant un sens biologique, il est important d'effectuer l'expérience sur des réplicats biologiques. Compte tenu des coûts de séquençage à l'époque (2010) l'expérience a été réalisée avec un seul échantillon contenant chacun un mélange d'individu de chaque stade de vie (gamétophyte mâle, gamétophyte femelle, tetrasporophyte), de manière à lisser la variabilité entre

TABLE 8: Les 20 gènes les plus exprimés (en moyenne).  $\bar{x}$  reads : moyenne du nombre de séquences. % of reads : pourcentage du nombre total de lectures.  
The 20 most expressed genes (in average).  $\bar{x}$  reads : average number of sequences.  
% of reads : percentage of total number of reads.

Gene code	$\bar{x}$ reads	% of reads	Nom/Name
CHC_T00010187001	588 525	4,01	HSP70
CHC_T00003878001	301 036	2,05	HSP90
CHC_T00010283001	201 872	1,38	HSP20
CHC_T00010213001	167 565	1,14	HSP20
CHC_T00010329001	158 936	1,08	HSP20
CHC_T00007026001	113 546	0,77	Mitochondrial Rho GTPase
CHC_T00007938001	97 242	0,66	HSP100 (ClpA/B)
CHC_T00009537001	79 464	0,54	Na <sup>+</sup> -ATPase
CHC_T00007128001	74 770	0,51	Liver stage antigen 3 precursor
CHC_T00002029001	69 800	0,48	HSP90
CHC_T00009511001	67 494	0,46	eEF1, subunit $\alpha$
CHC_T00006259001	65 159	0,44	—
CHC_T00002243001	56 571	0,39	—
CHC_T00000498001	56 149	0,38	—
CHC_T00009442001	55 985	0,38	HSP20
CHC_T00001910001	54 535	0,37	—
CHC_T00002296001	54 279	0,37	—
CHC_T00005866001	52 518	0,36	—
CHC_T00005111001	51 014	0,35	—
CHC_T00002672001	48 738	0,33	—
CHC_T00010078001	218 995	1,49	18 S rRNA

individus et faire ressortir les différences les plus significatives. Cette première expérience avait pour objectif de tester la faisabilité du RNAseq sur *C. crispus* et d'acquérir des compétences en matière d'analyse de grands jeux de données .

**Normalisation.** Dans un premier temps, je me suis intéressée à la normalisation des données. Dans une publication récente de Dillies *et al.*, 2012 [167], plusieurs méthodes de normalisation ont été comparées, donnant chacune des résultats différents. Pour nos expériences, dont les données ont été analysées antérieurement, nous avons opté pour la normalisation sur le nombre total de lectures alignées. Cette méthode n'est pas la meilleure surtout s'il s'agit de comparer l'expression des gènes au sein du même échantillon. Néanmoins, pour effectuer une analyse d'expression différentielle entre plusieurs échantillons, elle semble suffisante.

Je me suis également interrogée sur la question du bruit de fond, à savoir quel est le nombre minimal de lectures pour considérer qu'un gène est réellement exprimé. Le seuil étant variable selon la profondeur de séquençage, le type d'alignement effectué. J'ai arbitrairement retiré de l'analyse les gènes dont la valeur maximale d'expression dans au moins un échantillon était inférieur à 10, ainsi 8089 gènes ont donc été considérés comme exprimés dans au moins une des conditions.

**Statistiques.** Enfin, pour appuyer les résultats, je me suis intéressée à l'analyse statistique pour ce types de données sans réplicats biologiques.

Si l'on considère qu'un échantillon à séquencer est une population de fragments d'ADNc et que chaque gène est une espèce pour laquelle on souhaite estimer la taille, alors séquencer une "population" (transcriptome) revient à échantillonner de manière aléatoire chaque espèce (gène) dans le but d'estimer l'abondance relative de chaque espèce dans la population. Si chaque fragment d'ADNc a la même chance d'être sélectionné pour séquençage et de manière indépendante, alors le nombre de lecture pour chaque gène devrait suivre une loi binomiale [168]. L'épreuve consiste pour chaque gène à tirer au sort une lecture dans la population totale et voir s'il fait partie de l'espèce (le gène considéré). Le nombre de succès  $X$  est alors le nombre de lecture par gène pour l'échantillon considéré, le nombre d'épreuves  $n$  est le nombre total de lectures dans la population. La probabilité de succès d'une épreuve se rapporte alors à tous les échantillons considérés est le rapport entre la somme du nombre de lectures par gène et la somme du nombre total de lectures. Soit pour la comparaison de l'expression d'un gène Z entre un lot d'algues A et un lot d'algues B :

$X$  = nombre de lecture du gène  $Z$

$n$  = nombre total de lecture dans chaque échantillon

$p(Z_A)$  = probabilité qu'une lecture soit assignée au gène  $Z_A$  :

$$p(Z_A) = \frac{X_A + X_B}{n_A + n_B} \quad (1)$$

Les gènes exprimés de manière différentielle ont été identifiés en utilisant la loi binômiale pour leur assigner une valeur  $p$ , et par le calcul du taux de découverte de faux-positifs (FDR : *false discovery rate*) selon Benjamini et Hochberg [169]. Les gènes dont la valeur  $p$  est supérieure au seuil de FDR (5 %) et dont le facteur d'expression différentielle est supérieur à 2 ont été considérés comme exprimés de manière significativement différente.

**En quête de marqueurs d'expression.** Dans l'introduction, nous avons vu que malgré une architecture morphologique similaire entre les stades de vie haploïdes et diploïdes macroscopiques de *C. crispus*, il existe des différences entre eux du point de vue moléculaire et biochimiques, qui peuvent être dues à une expression différentielle de certains gènes. La suite de l'analyse n'a pas la prétention d'identifier des marqueurs génétiques mais de chercher des gènes qui sont exprimés de manière spécifique dans un stade donnée.

Au départ, je n'avais qu'un seul échantillon de chacun des trois stades de vie mais l'expérience "stress", réalisée *a posteriori*, a permis de générer des données supplémentaires sur les gamétophytes. J'ai commencé par prendre en compte tous les échantillons correspondants à des gamétophytes pour l'analyse. Mais la surreprésentation des gènes liés au stress (HSP notamment) dans les échantillons ayant subi un traitement (lumineux ou salin) a conduit à les écarter de l'analyse. Par conséquent je n'ai considéré que les algues ayant été congelées immédiatement après récolte. Le pool "gamétophytes" est donc représenté par cinq échantillons : les trois gamétophytes non-fertiles T10a, T10b, T10c et les deux gamétophytes fertiles M et F.

La comparaison des échantillons a été faite en deux temps : d'abord tetrasporophyte contre le "pool" gamétophytes, puis mâle contre femelle.

- 2052 gènes ont été identifiés comme différentiellement exprimés entre tetrasporophytes et gamétophytes, 947 étant surexprimés chez le tetrasporophyte et 1105 chez les gamétophytes.
- 1070 gènes ont été identifiés comme différentiellement exprimés entre mâle et femelle, 585 surexprimés chez le gamétophyte mâle, 485 chez le gamétophyte femelle.

Deux fois plus de gènes sont différentiellement exprimés de manière significative entre gamétophytes et tetrasporophyte, qu'entre gamétophyte mâle et gamétophyte femelle. Il apparaît donc, que dans ce cadre expérimental, les différences d'expression entre les stades de vie sont plus importantes entre les phases haploïdes et la phase diploïde, qu'entre les deux sexes.

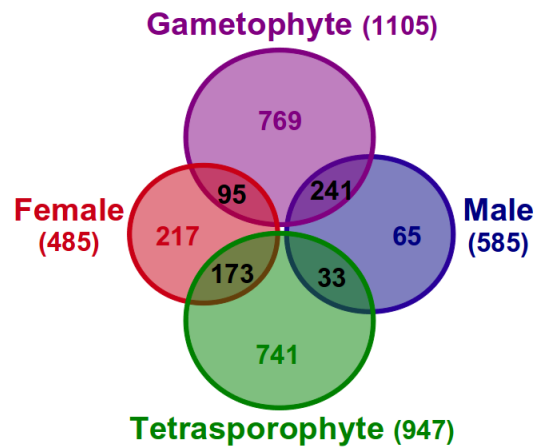


FIGURE 25: Diagramme de Venn du nombre de gènes surexprimés dans les stades de vie chez *C. crispus*. Male : gamétophyte mâle, femelle : gamétophyte femelle. Venn diagramm of the number of overexpressed genes in the life stages of *C. crispus*. Male : male gametophyte, female : female gametophyte.

**Validation expérimentale.** Avant de procéder à la suite de l'analyse, j'ai réalisé une expérience de validation de la pertinence des résultats, à l'aide d'une autre technique de quantification d'expression, la RT-qPCR, mieux connue et utilisée en routine dans le laboratoire.

La première étape pour procéder à la RT-qPCR est d'identifier des gènes dits de normalisation, qui ont un niveau d'expression stable dans les différentes conditions et servent de contrôle interne afin de quantifier l'ARN contenu dans chaque échantillon. Pour cela, les gènes couramment utilisés en tant que gènes de normalisation dans des organismes marins eukaryotes et prokaryotes ont été listés à partir de la littérature [170] [171]. Les niveaux d'expression de ces gènes ont été comparés entre les trois stades de vie de *C. crispus*. Les 12 gènes les plus stables ont été choisis, puis quantifiés par RT-qPCR. L'ensemble de l'expérience est décrite dans la section 3.3.

Le meilleur candidat pour normaliser les données de RT-qPCR est le gène codant pour le facteur d'initiation de la traduction IF4A-1.

Parmi les gènes exprimés de manière très contrastée entre les trois lots d'algues, treize gènes potentiellement spécifiques de chaque stade de vie ont été retenus à partir du jeu de données de RNAseq (la table 29 en annexe liste les amorces utilisées). Six d'entre eux, nommés GP1 à GP6,



sont fortement exprimés chez les deux gamétophytes, quatre autres (T1 à T4) sont clairement surexprimés chez les tetrasporophytes. À cela, s'ajoutent deux gènes davantage exprimés chez les mâles (M1 et M2) et un chez les femelles (F). Ces gènes additionnels ne sont cependant pas aussi différentiellement exprimés que les dix premiers. L'expression de ces gènes a par la suite été mesurée par PCR quantitative et comparée aux valeurs obtenues en RNAseq (figure 26). Les gènes GP1 à GP5 et T1 à T3 présentent des profils d'expression similaires à ceux observés en RNAseq. Les gènes GP6, M1, M2, F et T4 n'ont pas suivi la même tendance. GP6 est une peroxidase, son expression peut être conditionnée par un stress et n'est visiblement pas un bon marqueur. M1 et M2 n'ont pas d'homologues connus, F possède un domaine notch homologue à des protéines transmembranaires animales et T4 est une phosphatase. Les huit gènes ayant les mêmes profils en RT-qPCR et RNAseq semblent être de bons marqueurs pour discriminer les gamétophytes des tetrasporophytes. GP1 à 5 n'ont pas d'homologues connus, T2 et T3 non plus. Seul T1 est identifié, c'est une sulfurylase de type II, impliquée potentiellement dans la synthèse des carraghénanes. Ces huit gènes permettent de supporter la validité de la technique de RNAseq et suggèrent que ce qui différencie la phase haploïde de la phase diploïde chez *C. crispus* est inédit par rapport aux connaissances existant dans les bases de données publiques.

**Un peu de littérature...** Nos connaissances concernant l'expression spécifique de gènes relatifs au sexe et au cycle de vie dans les algues florideophytes est limitée. Afin de poursuivre la validation des gènes dont l'expression est spécifique d'un stade de vie dans les données de *C. crispus*, nous avons comparé nos résultats avec les rares données sur les algues rouges présentes dans la littérature.

La plupart des gènes étudiés chez les autres algues n'ont aucun homologue dans *C. crispus*. Par exemple, le gène femelle spécifique GMF-01 dans *Gracilaria lemaneiformis* [172], et un éventuel gène tetrasporophyte spécifique, SSH466, dans *G. lemaneiformis* [173], n'ont pu être identifiés dans le génome de *C. crispus*. Les cinq marqueurs génétiques du sexe de *Gracilaria chilensis* [174] n'ont pas été trouvés non plus.

En revanche, un homologue de Rab11, spécifique des femelles dans *G. lemaneiformis* [175] a été trouvé chez *Chondrus* (CHC\_T00009363001), une monoubiquitine (GUBQ1) surexprimée dans les tetrasporophytes par rapport aux femelles gamétophytes dans *G. lemaneiformis* [176] a aussi un homologue chez *Chondrus* (CHC\_T00010349001), et une HSP90 femelle spécifique dans *Griffithsia japonica* [177] a aussi été identifiée chez *C. crispus* (CHC\_T00003878001).

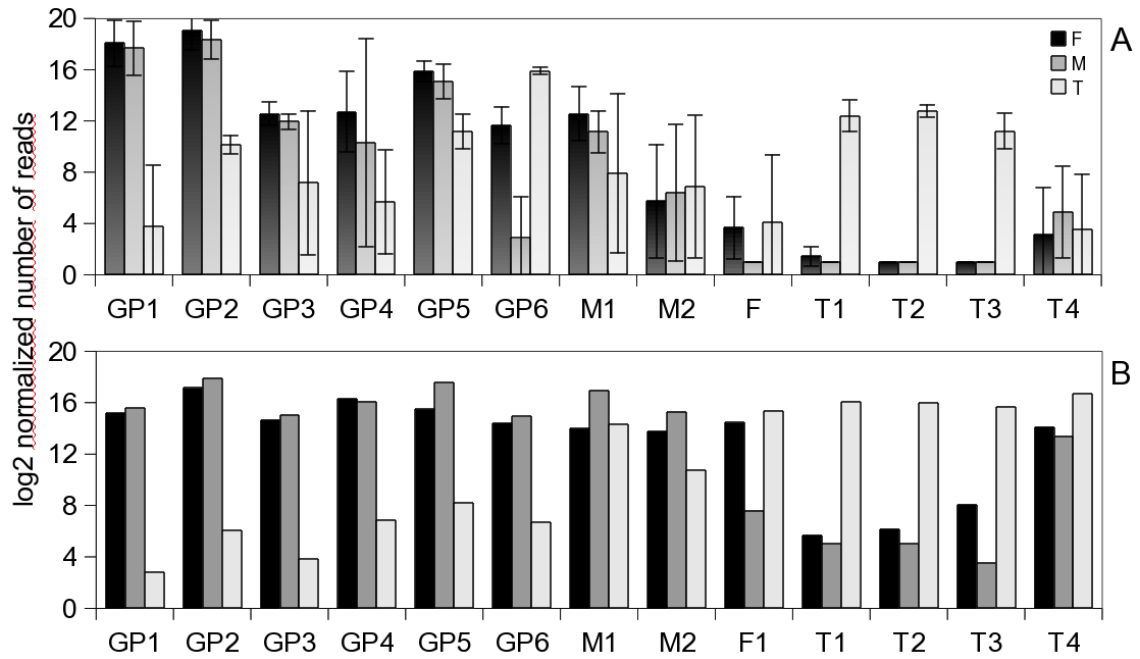


FIGURE 26: Profils d'expression des gènes candidats marqueurs d'expression de cycle de vie chez *C. crispus*. A : mesures en RT-qPCR normalisées avec IF4A, B : données issues de l'expérience de RNAseq. GP : gametophyte, M : gamétophyte mâle, F : gamétophyte femelle, T : tetrasporophyte.

Expression profiles of candidate expression marker genes of life cycle stages in *C. crispus*. A : RT-qPCR values normalized with IF4A, B : RNAseq data. GP : gametophyte, M : male gametophyte, F : female gametophyte, T : tetrasporophyte.

Néanmoins, la comparaison des niveaux d'expression de ces gènes dans *C. crispus*, présentés dans le tableau 9, ne montrent pas de différences significatives entre les stades de vie.

Il existe également des différences de réponse entre les gamétophytes et les tetrasporophytes face à l'algue verte pathogène *Acrochaete operculata*. Ainsi, suite à l'attaque du pathogène, les gamétophytes répondent par la production d'un burst oxydatif, alors que chez les tetrasporophytes, ce burst est négligeable [12]. L'enzyme responsable de ce burst, la NADPH oxydase, a été identifiée dans le génome (CHC\_T00010210001) et son expression comparée entre les trois stades de vie. Avec 254 lectures chez le tetrasporophyte contre 621 à 2255 lectures chez les gamétophytes, ce résultat est corrélé avec les expériences d'activité de cette enzyme.

Les expériences biochimiques réalisées sur d'autres échantillons de *C. crispus* sont valables sur les échantillons utilisés dans mes expériences de transcriptomique. En revanche, les connaissances sur les gènes dont l'expression est spécifique d'un stade de vie chez d'autres espèces d'algues rouges, même proches, ne peuvent pas se transférer à *C. crispus*.

TABLE 9: Expression en RNAseq des gènes homologues de quelques marqueurs de cycle de vie issues de la littérature et identifiés chez *C. crispus*. G1-3 : gamétophytes non fertiles (échantillons T10a, T10b, T10c) , M : gamétophyte mâle, F : gamétophyte femelle, T : tetrasporophyte.  
RNAseq expression of genes homologous to life cycle markers found in the literature and identified in the genome of *C. crispus*. G1-3 : non fertile gametophytes (samples T10a, T10b, T10c), M : male gametophyte, F : female gametophyte, T : tetrasporophyte.

Code	Name	G1	G2	G3	M	F	T
CHC_T00010349001	Ribosomal-ubiquitin fusion	9592	7815	6455	976	1009	1090
CHC_T00003878001	HSP90	18238	22179	24045	32683	97451	46019
CHC_T00009363001	Rab11 Small GTPase	3618	4168	3115	3982	5345	6619
CHC_T00010210001	NADPH oxydase	1432	1886	2255	1580	621	254

**Une voie de biosynthèse.** Une autre différence biologique majeure entre la génération diploïde et haploïde de *C. crispus* réside dans la voie de biosynthèse des polysaccharides de paroi cellulaire. Un modèle [187] propose un début de biosynthèse commun impliquant une sulfotransférase, puis deux voies distinctes impliquant à la fois des sulfotransférases et des sulfurylases. Un type de sulfotransférases et deux types de sulfurylases ont été identifiées dans le génome, chacune pouvant être codée par un ou plusieurs gènes. Leur expression est présentée dans le tableau 10.

TABLE 10: Expression en RNAseq des gènes de sulfotransférases et sulfurylases identifiés chez *C. crispus*. M : gamétophyte mâle, F : gamétophyte femelle, T : tetrasporophyte. Dernière colonne : G/T gamétophyte/sporophyte spécifique.  
RNAseq expression of sulfotransferases and sulfurylases genes identified in the genome of *C. crispus*. M : male gametophyte, F : female gametophyte, T : tetrasporophyte. Last column : G/T gametophyte/sporophyte specific.

Type	Code	M	F	T	
<b>Sulfotransferase</b>	CHC_T00008846001	4526	8159	1145	G
	CHC_T00008342001	3765	7289	406	G
	CHC_T00008402001	5	20	2863	T
	CHC_T00008834001	205	92	2403	T
	CHC_T00008796001	5	0	293	T
	CHC_T00009000001	991	1471	2059	-
	CHC_T00009100001	288	273	498	-
	CHC_T00008775001	306	293	279	-
	CHC_T00009431001	296	191	439	-
	CHC_T00006844001	1621	1777	1247	-
<b>Gal-2,6-Sulfurylase I</b>	CHC_T00008516001	58133	29970	14884	G
<b>Gal-2,6-Sulfurylase II</b>	CHC_T00009416001	0	5	6806	T
	CHC_T00008940001	15	440	12944	T
	CHC_T00009418001	8	7	9802	T
	CHC_T00008308001	3	2	1950	T
	CHC_T00008371001	0	15	1008	T
	CHC_T00009181001	245	97	430	-

Les gènes codant pour des sulfotransférases ont des expressions contrastées entre les gamétophytes et tetrasporophytes : les gènes CHC\_T00008846001 et CHC\_T00008342001 étant plutôt gamétophyte spécifiques et pourraient catalyser les réactions conduisant aux  $\mu$ - et  $\nu$ -carraghénanes, alors que les gènes CHC\_T00008402001 et CHC\_T00008834001 sont clairement surexprimés chez le tetrasporophyte et seraient plutôt impliqués dans la synthèse de  $\lambda$ -carraghénane. Concernant la sulfotransférase CHC\_T00008775001, l'expression est moins contrastée entre les échantillons et pourrait participer à la réaction commune aux deux voies de biosynthèse.

Concernant les sulfurylases, une seule d'entre-elles semble gamétophyte spécifique (la sulfurylase I : CHC\_T00008516001), une autre (CHC\_T00009181001) est exprimée de manière moins différentielle dans les trois échantillons, les cinq restantes étant hautement surexprimées chez le tetrasporophyte.

Ces données permettent de replacer hypothétiquement certaines de ces enzymes dans la voie de biosynthèse présentée en figure 27. Pour la synthèse des  $\lambda$ - et  $\theta$ -carraghénanes, les possibilités sont multiples, arbitrairement les gènes les plus exprimés ont été placés sur le schéma. En revanche, il semble y avoir peu de choix pour les  $\mu$ - et  $\nu$ -carraghénanes, puisque deux sulfotransférases sont nécessaires à leur synthèse, et seules deux d'entre-elles sont plus exprimés dans les gamétophytes. Quand à la synthèse de  $\kappa$ - et  $\iota$ -carraghénanes, une seule enzyme a été identifiée, des études biochimiques devraient aider à discriminer lequel des deux carraghénanes est synthétisé par la sulfurylase I.

Une majorité des enzymes identifiées sont exprimés fortement chez le tetrasporophyte, notamment les sulfurylases, potentiellement impliquées dans la synthèse de  $\theta$ -carraghénane qui est une forme hypothétique de carraghénane. Peu d'enzymes sont gamétophyte spécifiques. cela suggère que certaines enzymes impliquées dans la voie de biosynthèse des  $\kappa$ - et  $\iota$ -carraghénanes sont encore mal connues et font potentiellement partie des gènes sans homologues connus.

**En quête d'inconnu.** Pour identifier les gènes impliqués dans les mécanismes les plus importants dans chaque stade de vie, nous avons d'abord effectué une annotation automatique des gènes, de manière à pouvoir leur associer des termes GO (Gene Ontology [184] [185]) avec le programme Blast2GO [186]. C'est une manière d'assigner des catégories fonctionnelles, permettant d'organiser et classer les gènes et d'estimer les catégories les plus représentées comportant des gènes significativement différentiellement exprimés. Une étape d'enrichissement a été réa-

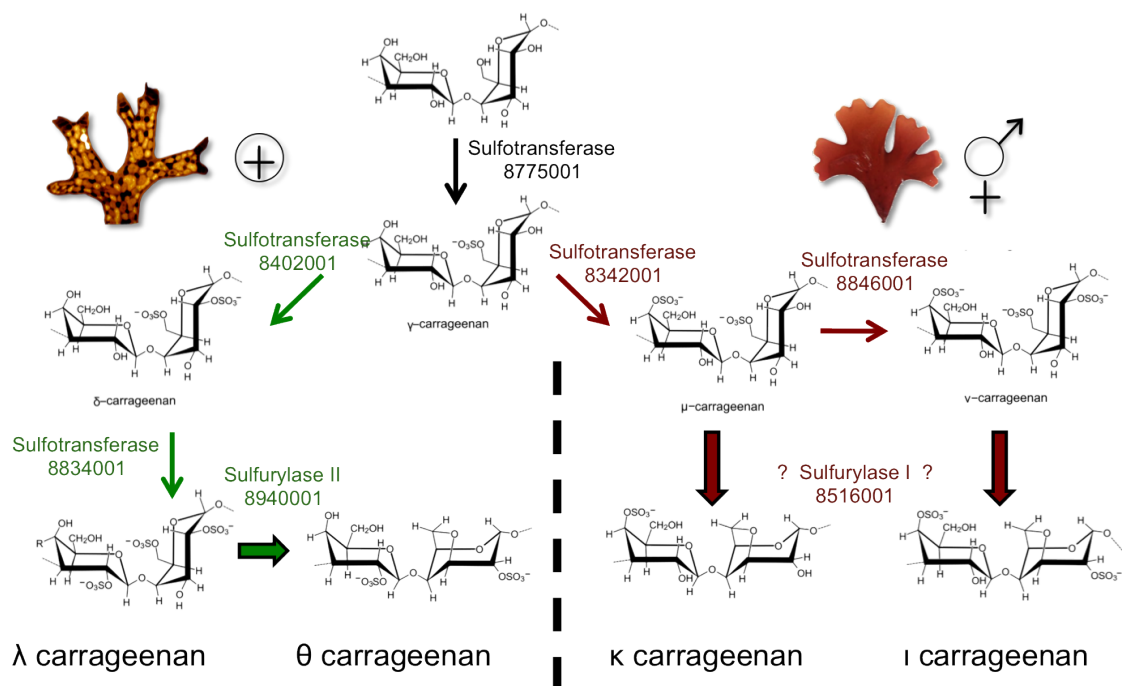


FIGURE 27: Voies de biosynthèse hypothétiques des carraghénanes et la position possible de quelques enzymes. En noir : voie commune, en vert : voie spécifique des tetrasporophytes, en rouge : voie spécifique des gamétophytes, flèche fine : sulfotransferase, flèche épaisse : sulfurylase.

Hypothetical biosynthesis pathways of carrageenans and possible position of enzymes. In black : common pathway, in green ; tetrasporophyte specific pathway, in red : gametophyte specific pathway, thin arrow : sulfotransferase, thick arrow : sulfurylase.

lisée, permettant d'identifier quelles catégories GO sont les plus représentées dans le jeu de données, les résultats de cet enrichissement sont présentés en table 11. Pour les gamétophytes, trois grandes catégories ressortent particulièrement : celles liées à la photosynthèse (photosynthèse GO :0015979, photosystème GO :0009521, thylakoïde GO :0044436), celles liées au métabolisme oxydatif et au stress (processus d'oxydation-réduction GO :0055114, activité oxidoréductase GO :0016491, activité antioxydante GO :0016209, réponse au stress GO :0006950) et enfin celles liées à la régulation d'enzymes (notamment des inhibiteurs et régulateurs de peptidases et métalloenzymes : en revanche, pour les gènes différentiellement exprimés dans le tetrasporophyte, aucune catégorie particulière n'est surreprésentée.

La comparaison des deux gamétophytes, mâle et femelle, montre un enrichissement en catégories GO déjà représentées lors de la comparaison gamétophyte-tetrasporophyte. L'origine de GO relatives à la photosynthèse vient donc du gamétophyte femelle, alors que la surreprésentation des régulateurs d'enzymes (GO :0061134, GO :0061135, GO :0030414, GO :0004866, GO :0004857) et du métabolisme oxydatif (GO :0016209, GO :0004601, GO :0016684) est issue du gamétophyte mâle.

Je me suis également intéressée aux gènes annotés les plus exprimés chez le tetrasporophyte et les gamétophytes, répertoriés respectivement dans les tables 12 et 13. Dans le tetrasporophyte, on retrouve majoritairement des enzymes impliquées dans le métabolisme des polysaccharides de paroi (trois sulfotransférases et sept sulfurylases), un homologue de protéine membranaire à domaine Notch, et une protéine à domaine homeobox. Dans les gamétophytes, il semble que les peroxydases, les heat shock proteins et des kinases soient majoritairement représentées.

L'analyse des gènes les plus différentiellement exprimés entre mâle et femelle (respectivement dans les tables 14 et 15) montre une surexpression de kinases chez le mâle (bien que leur niveau absolu d'expression soit faible) alors que chez la femelle, des enzymes de métabolisme des sucres (glycosyl-transférases) et relatives à la photosynthèse (composants du PSI) sont davantage exprimés.

TABLE 11: Catégories GO (gene ontology) surreprésentées dans les données RNAseq. G/T : entre gamétophytes et tetrasporophytes, M/F : entre gamétophyte mâle et gamétophyte femelle. Proc : processus biologique, Fn : fonction, Comp : compartiment cellulaire.  
GO categories (gene ontology) over represented in the RNAseq data. G/T : between gametophyte and tetrasporophyte, M/F : between male gametophyte and female gametophyte Proc : biological process, Fn : function, Comp : cellular compartment.

GO term	Category	Type	G/T	M/F
GO :0055114	oxidation-reduction process	Proc	G	
GO :0050896	response to stimulus	Proc	G	
GO :0006950	response to stress	Proc	G	
GO :0006952	defense response	Proc	G	
GO :0007165	signal transduction	Proc	G	
GO :0023052	signaling	Proc	G	
GO :0015979	photosynthesis	Proc	G	
GO :0006979	response to oxidative stress	Proc	G	
GO :0016491	oxidoreductase activity	Fn	G	
GO :0016209	antioxidant activity	Fn	G	M
GO :0043531	ADP binding	Fn	G	
GO :0004601	peroxidase activity	Fn	G	M
GO :0016684	oxidoreductase activity, acting on peroxide as acceptor	Fn	G	M
GO :0046906	tetrapyrrole binding	Fn	G	
GO :0005506	iron ion binding	Fn	G	
GO :0020037	heme binding	Fn	G	
GO :0061134	peptidase regulator activity	Fn	G	M
GO :0061135	endopeptidase regulator activity	Fn	G	M
GO :0030414	peptidase inhibitor activity	Fn	G	M
GO :0004866	endopeptidase inhibitor activity	Fn	G	M
GO :0004857	enzyme inhibitor activity	Fn	G	M
GO :0048551	metalloenzyme inhibitor activity	Fn	G	M
GO :0008191	metalloendopeptidase inhibitor activity	Fn	G	M
GO :0010576	metalloenzyme regulator activity	Fn	G	M
GO :0003849	3-deoxy-7-phosphoheptulonate synthase activity	Fn	G	
GO :0044436	thylakoid part	Comp	G	F
GO :0009579	thylakoid	Comp	G	F
GO :0009521	photosystem	Comp	G	F
GO :0034357	photosynthetic membrane	Comp	G	F
GO :0009522	photosystem I	Comp	G	F
GO :0009521	photosystem	Comp		F
GO :0009523	photosystem II	Comp		F



TABLE 12: Les 23 gènes annotés les plus différentiellement exprimés (plus de 10 fois) chez le tetrasporophyte de *C. crispus*. G : gamétophytes (moyenne de T10a, T10b, T10c, M et F) T : tetrasporophyte, DE-fold : facteur d'expression différentielle. À ces gènes s'ajoutent 58 gènes aux fonctions inconnues, non-répertoriés ici. The 23 annotated genes with the most differential expression level (10-fold and more) in the tetrasporophyte of *C. crispus*. G : gametophytes (mean of T10a, T10b, T10c, M and F), T : tetrasporophyte, DE-fold : differential expression factor. The 58 genes with unknown functions are not listed.

Code	T	G	DE-fold	Name
CHC_T00010108001	5872	115	52	Animal heme peroxidase
CHC_T00008796001	293	4	76	Carbohydrate sulfotransferase
CHC_T00008402001	2863	68	43	Carbohydrate sulfotransferase
CHC_T00008834001	2403	94	26	Carbohydrate sulfotransferase
CHC_T00007034001	166	2	75	Cysteinyl-tRNA synthetase
CHC_T00006674001	249	1	414	DEAD box ATP-RNA helicase
CHC_T00008657001	59	2	39	Exosome component 4 Rrp41-like
CHC_T00008945001	68	6	12	Exosome component 4 Rrp41-like
CHC_T00009416001	6806	15	472	Gal-2,6-Sulfurylases II
CHC_T00008308001	1950	1	1571	Gal-2,6-Sulfurylases II
CHC_T00009418001	9802	13	791	Gal-2,6-Sulfurylases II
CHC_T00008371001	1008	3	343	Gal-2,6-Sulfurylases II
CHC_T00008940001	12944	117	113	Gal-2,6-Sulfurylases II
CHC_T00009021001	23	0	23	Gal-2,6-Sulfurylases II
CHC_T00008447001	12	0	12	Gal-2,6-Sulfurylases II
CHC_T00007649001	895	52	17	Homeobox protein PKNOX1-like protein
CHC_T00006936001	331	5	71	Maltopentose forming amylase
CHC_T00006804001	1019	4	246	Mastigoneme-like protein
CHC_T00007400001	26256	761	35	Notch-like protein 1a
CHC_T00000119001	4392	429	10	Nucleolar RNA-binding protein
CHC_T00007266001	925	72	13	Vanadium-dependent bromoperoxidase
CHC_T00006547001	5475	71	78	YjeF-like protein
CHC_T00007782001	1643	17	101	Zinc finger protein

TABLE 13: Les 39 gènes annotés les plus différentiellement exprimés (plus de 10 fois) chez les gamétophytes de *C. crispus*. G : gamétophytes (moyenne de T10a, T10b, T10c, M et F) T : tetrasporophyte, DE-fold : facteur d'expression différentielle. À ces gènes s'ajoutent 125 gènes aux fonctions inconnues, non-répertoriés ici. The 39 annotated genes with the most differential expression level (10-fold and more) in the gametophytes of *C. crispus*. G : gametophytes (mean of T10a, T10b, T10c, M and F), T : tetrasporophyte, DE-fold : differential expression factor. The 125 genes with unknown functions are not listed.

Code	T	G	DE-fold	Name
CHC_T00008400001	2	17	11	Acyl-CoA oxidase
CHC_T00008546001	82	1369	16	Acyl-CoA-binding protein
CHC_T00010179001	85	32361	375	Animal heme peroxidase
CHC_T00009226001	30	3860	125	Animal heme peroxidase
CHC_T00006910001	0	20	20	Asp carbamoyltransferase catalytic subunit
CHC_T00006929001	29	785	27	tRNA (Asn/Gln) amidotransferase
CHC_T00007625001	11	269	25	Atypical/BCR protein kinase
CHC_T00000834001	55	595	11	cAMP-3',5'-cyclic phosphodiesterase 4D-like
CHC_T00008342001	406	12586	31	Carbohydrate sulfotransferase
CHC_T00008846001	1145	14454	12	Carbohydrate sulfotransferase
CHC_T00007700001	68	736	11	Chromosome condensation regulator RCC1
CHC_T00008966001	39	522	13	CDK regulatory subunit
CHC_T00010296001	1122	12736	11	DEAD box-like
CHC_T00008867001	2	736	478	DNA topoisomerase I
CHC_T00007483001	64	662	10	FACT complex subunit SPT16-like
CHC_T00009353001	12	419	34	Haloalkane dehalogenase
CHC_T00008448001	8	237	31	Haloalkane dehalogenase
CHC_T00010329001	248	3897	16	Heat Shock Protein 20, HSP20
CHC_T00008783001	202	2371	12	Histone H3
CHC_T00000712001	0	121	121	Histone H3
CHC_T00009200001	210	2665	13	Histone H4
CHC_T00007162001	0	51	51	NADH-dependent FMN reductase DszD
CHC_T00009164001	193	3217	16	NB-ARC protein
CHC_T00009020001	213	3241	15	NB-ARC protein
CHC_T00010065001	3	56	18	PAP2/haloperoxidase-like protein
CHC_T00007682001	80	2244	27	Peptidoglycan-binding lysin domain
CHC_T00007401001	3	46	15	Protein kinase
CHC_T00007862001	55	646	12	Pseudo-response regulator
CHC_T00006557001	3	45	14	tRNA-ribosyltransferase-isomerase
CHC_T00008538001	0	103	103	Serine/threonine protein kinase
CHC_T00008589001	2	61	40	Serine/threonine protein kinase
CHC_T00000229001	0	17	17	SNF2 family helicase
CHC_T00006930001	41	462	11	Tissue inhibitor of metalloproteinase 2b
CHC_T00007050001	3	46	15	TRAF - NF-kappa-B activator
CHC_T00000768001	0	17	17	Transposase
CHC_T00010345001	94	1051	11	Vacuolar-type H <sup>+</sup> -ATPase, Subunit V1-F
CHC_T00008417001	49	614	12	WD40-repeat containing protein
CHC_T00008276001	39	433	11	WD40-repeat containing protein
CHC_T00008426001	103	1084	10	WD40-repeat containing protein

TABLE 14: Les 12 gènes annotés les plus différentiellement exprimés (plus de 10 fois) chez le gamétophyte mâle de *C. crispus*. M : gamétophyte mâle, F : gamétophyte femelle, DE-fold : facteur d'expression différentielle.. À ces gènes s'ajoutent 31 gènes aux fonctions inconnues, non-répertoriés ici.

The 23 annotated genes with the most differential expression level (10-fold and more) in the male gametophyte of *C. crispus*. M : male gametophyte, F : female gametophyte, DE-fold : differential expression factor. The 31 genes with unknown functions are not listed.

Code	M	F	DE-fold	Name
CHC_T00000184001	21	2	11	Aldo-keto reductase
CHC_T00009008001	22	2	11	Animal heme peroxidase
CHC_T00007625001	438	18	25	Atypical/BCR protein kinase
CHC_T00009252001	28	2	14	Casein kinase I, CKA1
CHC_T00009010001	23	2	12	Cyclin-dependent kinase B2-1
CHC_T00008867001	1 010	24	43	DNA topoisomerase I
CHC_T00008946001	81	8	10	HAD-superfamily hydrolase
CHC_T00006964001	21	1	21	Multi-Sensor hybrid histidine kinase
CHC_T00010065001	105	4	27	PAP2/haloperoxidase-like protein
CHC_T00008538001	87	5	18	Serine/threonine protein kinase
CHC_T00010081001	12	1	12	U2 small nuclear ribonucleoprotein A
CHC_T00008276001	489	25	20	WD40-repeat containing protein

TABLE 15: Les 25 gènes annotés les plus différentiellement exprimés (plus de 10 fois) chez le gamétophyte femelle de *C. crispus*. M : gamétophyte mâle, F : gamétophyte femelle, DE-fold : facteur d'expression différentielle. À ces gènes s'ajoutent 65 gènes aux fonctions inconnues, non-répertoriés ici.

The 25 annotated genes with the most differential expression level (10-fold and more) in the female gametophyte of *C. crispus*. M : male gametophyte, F : female gametophyte, DE-fold : differential expression factor. The 65 genes with unknown functions are not listed.

Code	M	F	DE-fold	Name
CHC_T00010261001	22	241	11	40S ribosomal protein S12
CHC_T00010057001	6	195	32	Acetyl-CoA carboxylase
CHC_T00009356001	5	86	17	Acyl carrier protein
CHC_T00010108001	1	16	16	Animal heme peroxidase homologue
CHC_T00007807001	4	53	13	ATPase (AAA+ superfamily)-like
CHC_T00007395001	379	4 112	11	Carbohydrate binding module (family 6)
CHC_T00009301001	88	2 122	24	Cellulose synthase, family GT2
CHC_T00008940001	6	177	29	Gal-2,6-Sulfurylases II
CHC_T00009549001	58	921	16	Glycosyltransferase, family GT47
CHC_T00007649001	2	72	35	Homeobox protein PKNX1-like
CHC_T00009452001	49	585	12	Lipase domain protein
CHC_T00006532001	7	72	10	Megakaryocyte stimulating factor
CHC_T00009504001	34	785	23	Mitogen-activated protein kinase 5 MPK5
CHC_T00000119001	18	809	44	Nucleolar RNA-binding protein
CHC_T00009380001	81	834	10	Photosystem I P700 chlorophyll a apoprotein A1
CHC_T00008611001	9	140	15	Photosystem I P700 chlorophyll a apoprotein A2
CHC_T00000122001	9	1 077	118	Proteoglycan 4 isoform 1 precursor
CHC_T00010122001	24	421	17	Ribosomal protein L4, partial
CHC_T00008682001	38	559	14	Serine/threonine protein kinase
CHC_T00000114001	3	35	11	SusC/RagA family TonB-linked membrane protein
CHC_T00009096001	9	300	33	Ubiquitin-conjugating enzyme 1
CHC_T00007266001	2	36	18	Vanadium-dependent bromoperoxidase
CHC_T00009017001	14	245	17	Xylosyltransferase, family GT14
CHC_T00006547001	5	71	14	YjeF-like protein

**Discussion.** A notre connaissance assez peu de littérature existe sur les 'expérience à grande échelle sur d'autres algues rouges. Néanmoins, le genre *Pyropia* (en tant que *Porphyra*) a été étudié assez attentivement. Cette algue à un cycle de vie digénétique hétéromorphe, avec un sporophyte filamenteux microscopique et un gamétophyte plus grand à l'aspect foliacé. Différentes approches ont été utilisées pour étudier les différence entre gamétophyte et sporophyte, dont une approche basée sur des EST [178] : deux collections d'EST (20 779 EST en tout) ont été comparées après *clustering*. Dans cette analyse, 1 013 gènes étaient représentés dans les deux collections, dont 201 gènes qui étaient différentiellement exprimés de manière significative. De nombreuses protéines de stress était surexprimés dans le sporophyte alors que dans le gamétophyte, les éléments important pour la synthèse de protéines étaient majoritairement surexprimés. Contrairement à notre analyse, il n'y avait pas de différence observable dans les gènes potentiellement impliqués dans la synthèse de paroi. Une autre approche a aussi été utilisée pour *Porhyra*, la SSH (hybridation suppressive et soustractive) [179], pour étudier les différences entre les sporophytes portant des carpoporophytes et des gamétophytes.

D'autres études d'expression différentielle dans *Gracilariopsis andersonii* montraient dans un premier cas 117 gènes surexprimés dans les carposporophytes, incluant des glutathione S-transferase [182], et dans un second cas, 11 gènes présentaient une expression différentielle dans *Gracilariopsis* (*Gracilaria/Gracilariopsis*) *lemaniformis* [181].

Assigner des catégories fonctionnelles aux gènes est courant dans les analyses globales de type transcriptomique, cependant, au moins pour les données génomique de *C. crispus*, le nombre de gènes pour lesquels une catégorie fonctionnelle peut être assignée est assez faible, moins de 30 % étant annotés et parmi ceux-là, tous n'ont pas pu être classés. L'enrichissement en terme GO n'a pas toujours donné de résultat et il a été nécessaire de traiter les données à la main. Les listes de gènes les plus différentiellement exprimés ne sont pas tout à fait corrélées avec l'enrichissement en catégories GO. En effet, le logiciel Blast2GO prend en compte tous les gènes significatifs, c'est à dire à partir d'une surexpression minimale de 2 fois, alors que dans les tableaux que les listes présentées ne contiennent que les gènes dont la différence d'expression est extrême, à plus de 10 fois, ce qui réduit considérablement le nombre de gènes et de catégories GO.

Pour le tetrasporophyte, l'analyse des gènes les plus différentiellement exprimés a révélé de nombreuses enzymes liées à la réorganisation de la paroi cellulaire et des gènes importants dans les processus de contrôle du développement. Les tetrasporophytes utilisés pour l'expérience

étaient fertiles et portaient de nombreuses structures reproductives sur l'ensemble du thalle, contrairement aux gamétophytes dont les structures reproductives sont très localisées (apex pour les mâles et quelques petites zones pour les femelles). Les sulfurylases sont impliquées dans la dégradation des carraghénanes, la production de gamètes et de spores peut possiblement demander une réorganisation des parois cellulaires. Compte-tenu de la grande proportion des structures reproductives chez le tetrasporophyte, il est cohérent d'observer une forte expression d'enzymes de modification des carraghénanes dans ce stade de vie. En revanche, d'autres gènes de dégradation de la paroi tels que des cellulases ou carrageenases ne sont pas surexprimés. Dans les gamétophytes, les gènes connus relatifs au métabolisme des carraghénanes sont très peu exprimés. Il est possible que d'autres enzymes de biosynthèse des polysaccharides existent dans ce stade de vie, mais n'ont pas encore été identifiées et peuvent faire partie des nombreux gènes sans fonction assignée qui sont surexprimés chez les gamétophytes.

La comparaison des deux gamétophytes mâle et femelle, à la recherche de marqueurs d'expression a été asymétrique. Chez le mâle, les gènes annotés qui ressortait comme différentiellement exprimés font parti des gènes à faible expression. Il n'est donc pas évident de trouver un marqueur clair pour le mâle parmi les gènes connus. En revanche quelques gènes relatifs au métabolisme des polysaccharides semblent être plutôt femelle spécifiques. A l'instar du tetrasporophyte, le gamétophyte femelle porte des structures reproductives (alors que le mâle génère des gamètes qui sont dispersés dans le milieu) ce qui peut expliquer également la nécessité de remanier les parois cellulaires. Pour la femelle, l'enrichissement en catégories GO est soutenu partiellement par l'analyse manuelle des gènes, puisque dans les deux cas ressortent les gènes liés à la photosynthèse.

De manière générale, dans ce jeu de données, il n'y pas de gènes très fortement exprimés parmi ceux qui sont différentiellement exprimés entre mâle et femelle. Bien que les deux lots d'algues soient fertiles, la proportion de structures reproductives est faible comparé au l'ensemble du thalle qui est lui non reproductif. Trouver un marqueur fort au niveau transcriptionnel peut être difficile sur des thalles entiers. Cela suggère que le fonctionnement des deux gamétophytes, mâle et femelle, est très similaire et que les différences à l'échelle de l'individu sont subtiles. Pour identifier des différences plus nettes, il faudrait isoler les structures reproductives mâles et femelles afin d'enrichir le transcriptome en ARN spécifiques de chaque sexe. Sans oublier que le génome de référence est celui d'un gamétophyte mâle, le locus spécifique du gamétophyte femelle n'est donc présent dans les données génomiques et n'a pas encore été identifié. Ainsi,

même s'il les ARN correspondant aux gènes du locus femelle sont présents dans les séquences de RNAseq, il est difficile de les identifier au sein de toutes les lectures non-alignées sur le génome de référence mâle.

Enfin, les gènes relatifs au métabolisme général (protéines ribosomales, cycle de Krebs, respiration ou glycolyse) ne sont pas particulièrement représentés dans un stade de vie ou un autre. Le métabolisme général a donc une importance comparable chacun des stades de vie.

Cette expérience ne permet pas d'aller beaucoup plus loin dans la recherche de signification biologique. Néanmoins, elle permet de proposer des hypothèses de travail concernant les voies de biosynthèse de carraghénanes, elle a également permis de trouver des gènes candidats pour les expériences visant à identifier les gènes de normalisation pour la RT-qPCR. Enfin, elle a permis d'identifier trois gènes spécifiquement exprimés dans les tetrasporophytes, dont l'expression a été validée par RT-qPCR.

### 3.2.3 Lorsque *Chondrus crispus* stress...

Nous avons fait le choix de réaliser ces expériences dans des conditions les plus proches possibles des conditions naturelles, aussi bien au niveau des organismes que des approches expérimentales. Cela signifie que nous étudions *C. crispus* avec toute sa faune et flore associée (épiphytes, endophytes, microherbivores, etc.) et en utilisant des algues qui ne sont pas génétiquement homogènes. Cette approche écophysiologique a le potentiel de fournir des informations importantes sur le fonctionnement réel de *C. crispus* dans la nature.

Après avoir parcouru la bibliographie à la recherche de d'expériences similaires à la nôtre, les études les plus proches portaient sur de la méta-transcriptomique, c'est à dire des échantillons environnementaux comportant une multitude d'organismes, avec pour objectif d'analyser la fonctionnement de la communauté. Il semblerait que la transcriptomique quantitative de type RNAseq réalisée sur une seule espèce avec des échantillons environnementaux soit rare. L'étude qui suit est donc à notre connaissance inédite, au moins pour les végétaux, ce qui est avantageux d'une part, mais cela présente également un défi pour les analyses et les comparaisons avec d'autres études.

**Lumière, salinité et cycle diurnal.** Les expériences ont été réalisées dans six conditions, selon deux axes :

1. quatre conditions où l'on combine "stress" salin et lumineux (échantillons S+L+, S+L-, S-L+, S-L-). Chaque stress dure 3h, de 10h à 13h.
2. une cinétique diurnale, avec des algues prélevées à trois moments de la journée : 10h, 13h, 15h, (échantillons T10, T13, T15 respectivement). Les conditions sont proches des conditions naturelles, eau de mer à 100 % de salinité et plein soleil.

S-L+ est le même échantillon que T13. Un aperçu du protocole expérimental est montré en figure 28.

Pour appuyer statistiquement les résultats de séquençage, les gènes ayant moins de 10 lectures ont été retirés de l'analyse, les gènes différentiellement exprimés de manière significative ont été identifiés par ANOVA sous TigrMEV 4.8 [188] et le taux de faux-positifs a été calculé selon Benjamini & Hochberg [169]. Les tests ont été effectués entre les trois moments de la journée d'une part, et entre les quatre conditions lumière/salinité. Il est apparu qu'il n'y avait pas de différences significatives d'expression entre les échantillons ayant été dans des conditions de salinité normales et ceux dont l'eau de mer a été supplémentée en NaCl. En revanche, des





FIGURE 28: Protocole de stress *in situ*. Les algues sont placées dans des béchers d'eau de mer supplémentée (S+) ou non (S-) en NaCl, les béchers étant placés dans des caisses sont remplies d'eau de mer à température ambiante. Les caisses recouvertes d'un filtre contiennent les échantillons L-, les caisses ouvertes contiennent les échantillons L+.

Protocol for *in situ* stress. Algae are in beakers with NSW (S-) or NSW + NaCl (S+). The beakers are in boxes full of NSW at field temperature. Boxes covered with filters contain samples L-, open boxes contain samples L+.

différences d'expression importantes ont été observées entre les deux conditions lumineuses (100 % et 35 % de lumière solaire) ainsi que selon le moment de la journée (10, 13h, 15h). Les valeurs d'expression ont été transformées en  $\log_2$  puis des groupes (*clusters*) ont été identifiés sous TigrMEV en utilisant la fonction clustering hiérarchique par distance euclidienne.

Pour la cinétique diurnale, trois types de patrons d'expression ressortent de l'analyse et sont représentés dans le tableau 16. On peut constater que davantage de gènes sont surexprimés dans la matinée comparé à la mi-journée ou l'après-midi :

- 3089 gènes sont surexprimés à 10h, puis réprimés à partir de 13h,
- 629 gènes sont surexprimés à partir de 13h, et le sont toujours à 15h,
- 160 gènes sont surexprimés à partir de 15h.

Il faut noter aussi qu'il y a aucun gène significativement sous- ou sur-exprimé uniquement à la mi-journée. Il y a donc une importante reprogrammation entre le matin, au début d'une marée basse, et plus tard, lorsque les algues ont été exposées à des intensités lumineuses élevées pour une longue durée.

Pour l'axe "lumière", on distingue :

- 1064 gènes surexprimés sous la lumière filtrée à 35 % et réprimés sous pleine lumière,
- 225 gènes surexprimés par l'exposition à la pleine lumière,

Afin de mieux comprendre les données, j'ai effectué une synthèse de l'expérience avec les différents moments de la journée et les manipulations environnementales (figure 29). Des gènes

TABLE 16: Matrice des patrons d'expression des gènes au cours de la journée et selon les conditions de stress et nombre de gènes pour chaque patron. Cluster : nom du patron d'expression, + : sur-expression, - : sous-expression, Light : lumière, Salt : salinité.

Matrix of gene expression patterns during the day and for each stress condition and number of genes for each pattern. + : over-expression, - : under-expression.

cluster	10h	13h	15h	stress	total genes	identified
10h	+	-	-		3089	803
13h15h	-	+	+		807	173
15h	-	-	+		187	40
13h	-	+	-		0	0
10h15h	+	-	+		0	0
Light				+	225	49
Light				-	1064	164
Salt				+	0	0
Salt				-	0	0

sont ainsi communs entre la cinétique diurnal et les deux conditions lumineuses : 67 % des gènes réprimés par la lumière font partie du groupe de gènes surexprimés à 10h, et 75 % des gènes induits à la lumière sont également induits à partir de 13h. Ces résultats montrent que la lumière a une forte influence sur l'expression des gènes dans *C. crispus*.

**Les gènes matinaux.** Parmi les 3089 gènes induits le matin, 803 ont été annotés. Ils représentent essentiellement :

- des acteurs de la réplication et réparation de l'ADN (helicases, histones, ORC, remodelage de la chromatine),
- la synthèse d'ARN (facteurs d'épissage, RNA polymerases),
- la traduction (acides aminés, chaperones, ARN de transferts, facteurs d'initiation et d'élongation, ribosome)
- le cytosquelette (actine, tubuline)
- les polysaccharides de paroi (glycosyl hydrolase et transferases, cellulose)
- la photosynthèse (composants du PS, PBS, NDH, ATPase)

Ces gènes sont principalement des gènes de ménage et peu de gènes de stress étaient présents, indiquant que l'algue était dans une phase active de photosynthèse, division cellulaire et croissance.

Cent soixante-quatre de ces gènes sont également surexprimés en faible lumière (3 LHC, 9 enzymes liés à la synthèse des polysaccharides de paroi, 8 acteurs de la traduction, 7 pour la

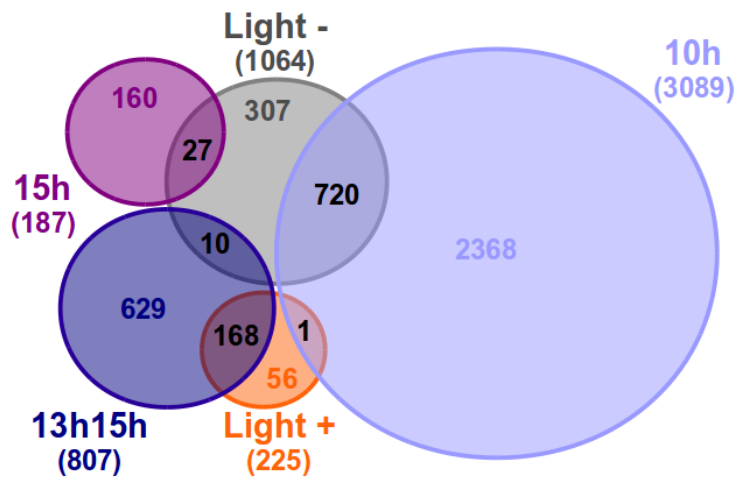


FIGURE 29: Diagramme de Venn du nombre de gènes surexprimés entre deux conditions lumineuses (Light + : de lumière solaire, Light - : gènes induits sous 35 % de lumière solaire) et selon une cinétique diurnal (10h : gènes induits le matin puis sous-exprimés à partir de 13h, 13h15h : gènes induits à partir de 13h, 15h : gènes induits à partir de 15h).  
Venn diagramm of the number of genes overexpressed between two light conditions (Light + : genes induced under 100 % of sunlight , Light - : genes induced under 35 % of sunlight) and during the day (10h : genes induced in the morning then repressed after 1PM, 13h15h : genes induced after 1PM, 15h : genes induced after 3PM).

transcription, 4 pour la réplication, 2 CDK et 3 gènes de stress).

**Les gènes tardifs.** Les gènes induits en début d'après-midi sont moins nombreux que le matin, puisqu'il y en a 807, dont 173 ayant été annotés. Les gènes de replication, transcription, traduction sont toujours nombreux à être surexprimés, mais la majorité des gènes alors induits sont des gènes de stress (détaillés dans la section suivante). Les composants de l'appareil photosynthétiques : notamment les protéines D1, D2, CP43, commencent à être resynthétisés à partir de 15h.

Il est donc possible que le métabolisme général et la photosynthèse soient réduit dans l'après-midi. Il a en effet été montré dans les algues rouges que la photosynthèse peut être réduite dans l'après-midi, par exemple chez *Gracilaria tikvahia* [190], chez *Gracilaria folifera* [191], ainsi que dans *C. crispus* [192] [112]. Il a également été montré que la croissance de *C. crispus* est plus importante le matin que pendant le reste de la journée [193].

Cela indique que le matin représente une période croissance et de photosynthèse active où de nombreux gènes sont fortement exprimés, suivi par une expression plus homogène, avec moins de gènes différentiellement exprimés, dans l'après-midi.

**Antioxydants et chaperones.** L'hypothèse de ce travail était que le stress et tout particulièrement les gènes acteurs dans la gestion des espèces réactives de l'oxygène ont une importance déterminante dans la réponse au stress chez *C. crispus*. Je me suis donc focalisé sur les gènes de stress et plus particulièrement les gènes relatifs aux enzymes antioxydantes et aux chaperones,

Parmi les gènes induits à 10h, puis réprimés l'après-midi, les glutaredoxines et glutathion peroxidases sont toutes représentées, ainsi que l'ensemble du complexe TCP1. Ces gènes sont remplacés à 13h par une ascorbate peroxidase, une monodehydroascorbate reductase, toutes les petites HSP, la moitié des HSP70, deux HSP 100 (ClpA/B). Ces gènes restent fortement exprimés à 15h, auxquels s'ajoutent une catalase et la calnexine.

Les expressions plus élevées de APX, CAT, TrxR et MDHAR dans l'après-midi, ainsi que l'APX dans les conditions de forte lumière, indiquent un stress oxydatif accru, résultant de la génération de ROS et de la diminution de l'activité photosynthétique et possiblement influencé par la forte intensité lumineuse. Dans *Gracilaria tikvahiae* [190], les activités de l'APX et la CAT diminuaient dans l'après-midi, indiquant que l'augmentation de l'expression des gènes n'est pas nécessairement corrélée à une augmentation d'activité. Néanmoins, l'augmentation de l'expression reflète l'effort que l'algue investit pour s'acclimater.

GRX et GPX (ainsi que la SOD à cuivre) sont des enzymes antioxydantes, cependant, leur patron d'expression est différent de celui des enzymes antioxydantes typiques, ce qui peut indiquer que leur rôle est différent dans *C. crispus*. Le système glutaredoxine fournit de la force réductrice et peut ainsi avoir un rôle en tant qu'antioxydant, mais il est surtout important pour la réduction des ponts disulfures dans les protéines [194]. Cette seconde fonction a peut-être davantage d'importance chez *C. crispus* en période de forte activité transcriptionnelle et par conséquent de forte synthèse de protéines.

La GPX utilise le glutathion comme donneur d'électrons et peut catalyser la réduction aussi bien du  $H_2O_2$  et des lipides hydroperoxydés, en eau et alcool [195] [196], et participe à la régulation du stress oxydatif après un stress photooxydatif et l'infection par des pathogènes [197]. Considérant son patron d'expression, il est possible que le rôle de la GPX dans *C. crispus* soit davantage lié à un stress oxydatif plutôt qu'un stress photo-induit.

L'augmentation de l'expression des gènes TCP1 dans la matinée pourrait être corrélée avec leur activité en tant que chaperones pour la tubuline [154], ainsi qu'avec la forte expression de la tubuline également observée.

Les résultats montrent également l'importance des HSP dans *C. crispus*, de grandes familles

de gènes dont les profils d'expression sont très dynamiques. Il est remarquable que l'expression matinale des gènes de ménage soit remplacée par une expression de gènes dominée par les HSP (12 % des gènes annotés l'après-midi contre moins de 1 % des gènes annotés du matin). Il a également été montré que la réponse impliquant des HSP chez *C. crispus* est très dynamique dans différents types de stress, incluant la forte lumière et la salinité [16].

Ces résultats mettent l'accent sur l'importance de la lumière (et potentiellement du cycle diurnal) sur l'expression des gènes dans *C. crispus*. L'étude précédente de Collén *et al.* (2007) indiquait que le stress osmotique était déterminant dans l'expression des gènes *C. crispus* [16]. Les expériences présentées ici, plus proches des conditions naturelles, mettent davantage l'accent sur les effets de la forte intensité lumineuse. Les différences entre les deux études pourraient aussi être liées à la réponse plus dynamique à la lumière des échantillons environnementaux que les échantillons cultivés au laboratoire (plus de détails dans la section 3.4.

TABLE 17: Expression des gènes de ROSEs selon les conditions lumineuses et le cycle diurnal.

Expression of ROSEs genes in the two light conditions and during the day.

gene	Light+	Light-	10h	13-15h	15h
APX	x			x	x
APX					
CAT					
CAT					x
TRxR				x	x
DHAR			x		
GRX			x		
GRX			x		
GRX			x		
GRX			x		
GPX		x	x		
GPX			x		
GR			x		
GST					
Mn SOD					
Mn SOD					
Cu SOD					
Cu SOD			x		
MDHAR				x	x
PMSR					
PMSR					
PRX					
PRX					

TABLE 18: Expression des gènes de HSP selon les conditions lumineuses et le cycle diurnal.  
Expression of HSP genes in the two light conditions and during the day.

gene	Light+	Light-	10h	13-15h	15h
CPN10				x	x
CPN10				x	x
HSP20				x	x
HSP20				x	x
HSP20				x	x
HSP20				x	x
HSP20				x	x
HSP20				x	x
HSP20		x	x	x	x
HSP40/DnaJ				x	x
HSP40/DnaJ				x	x
HSP40/DnaJ					
HSP40/DnaJ		x	x		
HSP40/DnaJ					
HSP40/DnaJ			x		
HSP40/DnaJ			x		
CPN60				x	x
CPN60					
HSP70				x	x
HSP70				x	x
HSP70				x	x
HSP70	x			x	x
HSP70					
HSP70					
HSP70				x	x
HSP70					
HSP70-BP					
HSP90					
HSP90					
HSP90					
ClpA/B	x			x	x
ClpA/B				x	x
HSP100					
GrpE				x	x
GrpE		x	x		
HSF			x		

TABLE 19: Expression des gènes de Chaperones selon les conditions lumineuses et le cycle diurnal.

Expression of Chaperone genes in the two light conditions and during the day.

gene	Light+	Light-	10h	13-15h	15h
PPIA					
PPIA					
PPIA					
PPIA					
PPIA			x		
PPIA					
PPIA					
PPIA			x		
PDI					x
PDI				x	x
PDI					
TBCB					
TBCC		x			
TBCD					
TBCE			x		
TCP1 $\alpha$			x		
TCP1 $\beta$			x		
TCP1 $\delta$			x		
TCP1 $\eta$			x		
TCP1 $\gamma$			x		
TCP1 $\theta$			x		
TCP1 $\zeta$			x		
Calnexin					x

### 3.3 À la recherche des gènes de normalisation de RT-qPCR

#### Préambule

L'article qui suit décrit la méthodologie permettant de déterminer les meilleurs gènes de normalisation pour les expériences de RT-PCR quantitative chez *C. crispus*. La RT-qPCR est une technique sensible pour quantifier l'expression des gènes, mais elle dépend de la stabilité de l'expression des gènes utilisés comme référence pour la normalisation des données [199] [200]. Jusqu'à présent, le gène codant pour l'actine avait été utilisé comme référence pour normaliser les données de RT-PCR quantitative chez *C. crispus* [13], étant un gène dont l'expression attendue est potentiellement stable. Cependant, en compilant les données chez d'autres espèces végétales, on constate qu'il n'existe pas de gène(s) de référence universel(s) commun(s) à tous les organismes [205] [206]. Il est donc nécessaire de procéder à une détermination des gènes les plus appropriés pour chaque modèle. La transcriptomique est une part importante du travail mené sur *C. crispus* et la PCR quantitative a pour objectif de permettre d'exploiter plus finement les données haut-débit. Il est apparu nécessaire d'effectuer ce travail méthodologique visant à identifier les gènes de normalisation chez *C. crispus*.

#### Résumé

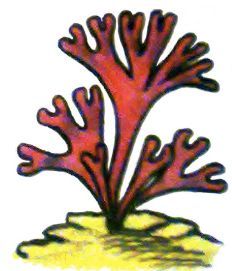
Afin identifier les meilleurs gènes de référence pour normaliser les données de RT-qPCR chez l'algue rouge *Chondrus crispus*, l'expression de 14 gènes a été quantifiée dans 56 échantillons d'algues soumis à 18 conditions différentes. L'originalité de l'étude porte sur le type de matériel biologique utilisé, doit deux lots d'algues, le premier étant issu de la souche séquencée cultivée et soumis à des conditions de stress en laboratoire et le deuxième comportant des algues récoltées sur le terrain et stressées *in situ*, issues des mêmes échantillons que les algues ayant servi pour l'expérience de transcriptomique "stress salin et lumineux" (voir la section 3.2.3).

La stabilité de l'expression des gènes entre les échantillons a été évaluée selon trois méthodes : la première consistait à comparer les valeurs de Ct (*cycle threshold*), c'est à dire du nombre de cycles de PCR à un seuil donné. Les deux méthodes suivantes étaient basées sur des analyses statistiques en utilisant les algorithmes geNorm et NormFinder.

Les deux groupes d'échantillons (culture et terrain) présentaient des comportements différents. Les gènes étant tous stables entre les algues cultivées en laboratoire, alors que pour les échantillons environnementaux, des différences d'expression ont pu être observées. Ainsi selon



le type d'échantillon dont l'expression de gènes doit être quantifiée, il faut s'attendre à des différences de comportement. Selon les méthodes d'évaluation utilisées, les candidats les plus stables étaient différents, lorsque les deux lots d'échantillons étaient analysés séparément. En revanche, si l'on considère l'ensemble des algues, deux gènes semblent être de bonnes références de normalisation, en concordance selon les deux méthodes statistiques : IF4A-1 et eRF3.



# RT-qPCR normalization genes in the red alga *Chondrus crispus*

Nathalie Kowalczyk<sup>1,2</sup>, Sylvie Rousvoal<sup>1,2</sup>, Cécile Hervé<sup>1,2</sup>, Catherine Boyen<sup>1,2</sup>, Jonas Collén<sup>1,2\*</sup>

**1 CNRS - UMR7139 Végétaux marins et biomolécules, Station Biologique de Roscoff, place G. Teissier, Roscoff, Brittany, France**

**2 UPMC - Université Pierre et Marie Curie - Paris 6, Station Biologique de Roscoff, place G. Teissier, Roscoff, Brittany, France**

\* E-mail: collen@sb-roscoff.fr

## Abstract

*Chondrus crispus* is a common red macroalga living on the rocky shores of the North Atlantic Ocean. It has a long research history, being a major source of carrageenan, a thickener widely used in the food industry, but also for physiological and ecological studies. To establish it as a model for red algae, its genome has been sequenced, allowing the development of molecular tools such as quantification of gene expression, including RNAseq and RT-qPCR. To determine appropriate genes for RT-qPCR normalization, the expression of 14 genes were monitored in 18 conditions using two pools of algal samples: samples from the sequenced strain, cultured and stressed in laboratory conditions and *C. crispus* collected on the shore and stressed *in situ*.

The expression stability of the genes between the samples was evaluated by comparing the Ct range and using geNorm and NormFinder. The candidate genes included translation related proteins (initiation factors IF4A-1 and IF4A-2, elongation factor EF1 $\alpha$  and eRF3, an eukaryotic polypeptide chain release factor), cytoskeleton proteins (two  $\beta$ -tubulins,  $\alpha$ -tubulin and actin), enzymes involved in the pentose phosphate pathway (glucose 6-phosphate dehydrogenase), protein recycling process (ubiquitin and ubiquitin-conjugating enzyme) and glycolysis (isocitrate dehydrogenase).

The two pools of samples showed different expression patterns. Most of the genes were stable in the algae cultivated in the laboratory, whereas environmental samples showed a more important variation in gene expression. When analyzing the two pools separately, the ranking of the most stable genes were different from one method to another. When considering all samples, the two statistical methods were concordant, revealing translation initiation factor 4A-2 and eukaryotic polypeptide chain release factor

3 as pertinent normalization genes. This study highlights thus the importance of testing reference genes according to the experiments as well as the genetic background of the organism.

## Introduction

*Chondrus crispus* is a red macroalga, widely represented in the intertidal and subtidal zones of the rocky shores of the North Atlantic Ocean. Because of its ecological abundance and its economical importance, it has been one of the principal sources for carrageenan, a gellifying and thickening molecule used in the food industry, it also has a relatively important research history. Its habitat is a dynamic environment undergoing rapid changes of large amplitude in physical and chemical parameters due to the tidal cycles, combined with diurnal and weather variations. Studies about stress in *C. crispus* have been carried out, first with physiological approaches [1] [2] [3], then with molecular tools such as large scale quantitative transcriptomics [4] [5].

In previous reports [4] [6] [7], RT-qPCR has been used for targeted expression studies in *C. crispus*. This technique is sensitive to quantify gene expression, highly reliable [8] [9], but depends of the stability of the genes used as reference for data normalization. Different methods of identifying normalization genes have been developed so far, such as geNorm [10], NormFinder [11] and BestKeeper [12]. These algorithms have been used in many reports, in a wide range of organisms and tissues, first in animals (dolphin [13], fish [14], worm [15]), then in plants (rice [16], grapevine [17], banana [18]). More recently it has also been used on the brown algal model *Ectocarpus siliculosus* [19].

Until now, actin has been used as a normalization gene for *C. crispus*, based on the use of this gene as a common reference gene in other species. However, the stability of this gene has not been tested in *C. crispus*. To our knowledge, few other studies in red algae have used RT-qPCR for gene expression, however, when performed, usually without data normalization [20] [21]. In one case, an empirical gene was used as a reference [22], and in one recent study, Wu *et al.* showed that  $\beta$ -tubulin was the most appropriate normalization gene for expression studies between *Porphyra haitanensis* gametophyte and tetrasporophyte [23]. Even if normalization genes have been defined in many species, it appears that there is no universal gene common to every organism. Quantitative transcriptomics is an important part of the work on *C. crispus* and RT-qPCR aims to exploit accurately the high-throughput data. The genome of *C. crispus* has been sequenced and annotated [24]. This simplified the determination and identification

of the optimal normalization genes in *C. crispus*.

## Materials and methods

### Culture conditions and treatments

The red alga *Chondrus crispus* Stackhouse (Gigartinales, Rhodophyta) from two different origins was used for the experiments (see table 1), the samples A, F, G, H, I and M were algae collected near the port of Bloscon in Roscoff (48.724208,-3.969744 Brittany, France). The algae labelled C1 to C12 were samples from the strain of *C. crispus* "Peggy's Cove", a gametophyte collected in Canada and kept in 10 L plastic flasks in filtered and autoclaved NSW in a culture room at 13°C and bubbled with compressed air. Light was provided by fluorescent tubes with a photon flux density of 100  $\mu\text{mol}$  of photons·m<sup>2</sup>·s<sup>-1</sup> for 12 hours per day.

For the chemical treatments the algae were transferred into Petri dishes containing NSW and the additives for 3 hours. The treatments were 10 mM H<sub>2</sub>O<sub>2</sub>, 200 mg·L<sup>-1</sup> harpin, 200  $\mu\text{M}$  CdCl<sub>2</sub>, 200  $\mu\text{M}$  CuSO<sub>4</sub>, 200  $\mu\text{M}$  ZnCl<sub>2</sub>, 200  $\mu\text{M}$  AlCl<sub>3</sub>, 500  $\mu\text{M}$  dichlorvos, 500  $\mu\text{M}$  glyphosate, 500  $\mu\text{M}$  paraquat, 1 % DMSO and 100  $\mu\text{M}$  methyl jasmonate with DMSO (1 %).

The treatments for the field collected algae were full sunlight in NSW for 0, 3 and 5 h to follow the diurnal cycle. For hypersaline stress combined with light stress, algae were exposed to full sunlight in enriched NSW with 33 g·L<sup>-1</sup> of NaCl added, sunlight filtered to 35 % in NSW and sunlight filtered to 35 % in NSW enriched with 33 g·L<sup>-1</sup> of NaCl.

Three biological replicates, each containing a pool of algal thalli, were obtained for each treatment and these were used for RNA extraction.

### Nucleic acid extractions

RNA and DNA were extracted using respectively Qiagen RNeasy plant kit and DNeasy plant kit, according to the manufacturer's protocols with the following two modifications: 100 mg of frozen tissue were ground in liquid nitrogen and were resuspended in the extraction buffer. After mixing vigorously for several minutes, a centrifugation step was added to eliminate cellular debris. For the RNA extraction, a treatment with RNase-free DNase I (Turbo DNase, Ambion) was performed in order to eliminate residual genomic DNA.

## RNA quantification and cDNA synthesis

Nucleic acid concentrations were measured by absorbance at OD<sub>260</sub> using a Thermo® NanoDrop 2000 spectrophotometer. The purity of RNA samples was assessed by measuring the ratio OD<sub>260</sub>/OD<sub>280</sub> and OD<sub>230</sub>/OD<sub>260</sub> (see additional tables 1 and 2). RNA integrity was verified by capillary electrophoresis using the Agilent Bioanalyzer 2100 (additional data, Fig. 3), according to the manufacturer's instructions or by agarose gel electrophoresis (additional data Fig. 4). The RNA of each sample was diluted and 800 µg was transcribed to cDNA using oligo(dT)<sub>12-18</sub> and the Superscript<sup>TM</sup> II RT kit (Invitrogen) according to the manufacturer's protocol, and subsequently diluted with nuclease free water to 0.5 ng.µL<sup>-1</sup>.

## Real-time PCR

For each gene, a pair of oligonucleotide sequences was designed close to the 3' end of the coding sequence using Primer Express<sup>TM</sup> from Applied Biosystems (Table 2). The specificity of the oligonucleotides was tested *in silico*, using BLAST on the whole genome of *C. crispus*. The Q-PCR reactions were performed in a 96-well Chromo 4 thermocycler (Bio-Rad) with SYBRgreen PCR master kit. The protocol was : 14 min at 95°C, followed by 40 runs of 15 s at 95°C and 60 s at 60°C. Each sample was technically triplicated. *C. crispus* genomic DNA was used as a quantification reference. A 1:6 dilution series ranging from 29 to 37,500 copies of the *C. crispus* genome was prepared and tested for each gene in order to determine the amplification efficiency. The specificity of the amplification was verified with a dissociation curve obtained by heating the samples from 65°C to 95°C. In addition to the DNase I treatments, the absence of genomic DNA was checked by amplifying an intron sequence on the cDNAs. The number of copies of contaminant gDNA (less than 100 copies) was subtracted from all other values, prior to any further analysis.

## Results and discussion

### Treatments and choice of housekeeping genes

Two sets of conditions were tested in this study. In one set, algae collected on the shore were submitted to treatments *in situ* immediately after sampling in order to reproduce conditions close to natural stresses. As the diurnal rhythm has a strong influence on metabolism and physiology of red algae, a series of

samples collected at three different times of the day has been analyzed. In tide pools, light and salinity are known to vary considerably, these natural stressors will modify the gene expression. The environmental samples had a genomic background different from the sequenced strain from which the PCR primers were designed, which makes the originality of this study. The other set of experiments was carried out in laboratory controlled conditions, using samples of the sequenced strain of *C. crispus*. The agents used such as  $H_2O_2$ , a reactive oxygen species produced by many organisms, including algae, under conditions of abiotic and biotic stress, as well as methyl jasmonate, are known to induce a strong stress response at the transcriptomic level in *C. crispus* [4]. Metals like copper, cadmium, zinc and aluminium were also tested, being important pollution factors in marine environments. Harpin, dichlorvos, paraquat (methyl viologen) and glyphosate are widely used pesticides able to induce a strong expression of stress response genes [6]. Paraquat and glyphosate are herbicides, the first generates reactive oxygen species by re-rooting electrons from the photosystem I to molecular oxygen and the other inhibits the synthesis of aromatic amino-acids.

RNA was extracted from biological triplicates of algae treated as above, resulting in a total of 54 samples for 18 conditions. The abundance of the transcripts of 14 potential housekeeping genes was assayed on these cDNAs. First, candidate genes considered as normalization genes for RT-qPCR in other marine eukaryotic and procaryotic species were chosen [19] [25]. Then, among this list, 12 genes have been selected from a transcriptomics experiment done on environmental samples of *C. crispus* and correspond to the most stably expressed genes in this experiment. The candidate genes included translation related proteins (initiation factors IF4A-1 and IF4A-2, the elongation factor EF1 $\alpha$  and eRF3, an eukaryotic polypeptide chain release factor), cytoskeleton proteins (two  $\beta$ -tubulins,  $\alpha$ -tubulin and actin), an enzyme involved in the pentose phosphate pathway (glucose 6-phosphate deshydrogenase), proteins involved in the protein degradation and recycling process (ubiquitin and ubiquitin-conjugating enzyme) and also and enzyme involved in glycolysis (isocitrate dehydrogenase). Two additional genes were also selected for their high differential expression as negative controls: a NADH-ubiquinone oxidoreductase and a L-amino acid oxidase. The genes are listed in Table 2.

## Quantification and data analysis

The first analysis aimed to assess whether the transcripts levels of the candidate genes were comparable between the different conditions. The cycle threshold (Ct [9]) value variations have been calculated for

each gene and are shown in Fig. 1. All genes had different levels of expression, corresponding to Ct-range from 21.66 to 28.94, and some of them were influenced by the treatments. Actin and IF4A-2 showed higher expression values than the other genes, while G6PDH had the lowest expression. Most of the genes had more than 7 cycles of variation between the samples and one gene, *eRF3*, was equally expressed in all samples and showed very little variation in the Ct values with only 1.81 cycles of variation.

When considering the two sets of samples separately, the variations in the Ct range were smaller, 7 genes from the "cultured" pool showed less than 3 cycles of variations (see supplementary data Fig. 1A), *eRF3* was still stable but at the second position (1.81 cycles), the smallest Ct value variation was showed by Tub  $\beta$ -2 (1.16 cycles). A similar shrinking was observed for the field samples (supplementary data Fig. 3) where *eRF3* and Tub  $\beta$ -2 had the same rank and showed even smaller variations in Ct values (1.5 and 0.56 cycles respectively).

The geNorm pairwise analysis was also performed in order to test the robustness of the data. This analysis was first described by Vandesompele *et al.* [10] and is still widely used to evaluate normalization genes. The stability of the genes was tested between the different conditions and the results are shown in Fig. 2A. Genes were considered suitable when M was inferior to 0.5. Two genes fulfilled this criterium: *eRF3* and *IF4A-1*, the most stable being *IF4A-1*. The least stable genes were *IcdH*, *Nuox* and *Aaox*. The M values were also calculated for the two pools of samples separately. For the culture samples (Fig. 2B) the most stable genes were *eRF3*, *UBQ* and *IF4A-1*. For the field samples (Fig. 2C) the genes were different : *Actin*, *Tub  $\beta$ -1* and *IF4A-1*.

Another approach, using NormFinder, was also used to test the candidate genes (Fig. 3A) The ranking of the genes with intermediate M values was different, compared to the geNorm analysis. However, there was a good correlation for the two most stable genes (*eRF3* and *IF4A-1*) and the three most fluctuating. To calculate a normalization factor, NormFinder showed that the best combination of two genes was *eRF3* and actin, with a stability value of 0.195. Interestingly, when the analysis was done with culture samples the ranking was different (Fig. 3B), *eRF3* and *EF1 $\alpha$*  being the most stable, but with the environmental samples (Fig. 3C) the same genes as the full analysis were considered as most stable.

Plotting the geNorm and NormFinder analyses one against the other (Fig. 4) supports the global ranking of the 14 genes, for the whole set of samples.

## Conclusion

In this study, two methods were used to identify the best normalization genes: geNorm and NormFinder and the results were concordant. The experiment was made with two pools of samples, six samples harvested on the field and thus with a varied genetic background, and twelve samples corresponding to the sequenced strain of *C. crispus* thus having the same genetic background. The field samples came from an environment with dynamic conditions, due to tidal cycles and weather changes. The sequenced strain is cultured in very stable conditions. These two environments lead to different physiological responses to the stressing conditions to which the algae have been submitted. Field samples have a more plastic metabolism than the cultured algae, that were acclimated to stable conditions. When considering all the samples together, the large variations observed particularly when it comes to Ct value, are clearly due to the differences between the two pools of samples, which are more homogenous when considered separately than all together.

From the result of the geNorm analysis, the samples from Peggy's Cove exhibited a stable expression for all the candidate genes. The trend was similar with the NormFinder analysis, but since the differences between values was small, the ranking of the genes was slightly different. Thus if the experiment was carried out with only laboratory samples, as it is frequently done, all the tested genes could have been used as normalization genes, even those which were supposed to have differential expression. However, when considering the field samples, the stability values are higher and the ranking of the genes is different.

For expression studies concerning only environmental samples, even if the results were not fully concordant, a trend emerged, Tub  $\beta$ -1 and IF4A-1 were well ranked in the different methods used.

When tallying all methods, IF4A-1 and *eRF3* are the best candidate gene for normalization, as they have a good rank whatever the method used for calculations, especially when analyzing samples from various origins and having thus different behaviors.

## Acknowledgments

The authors would like to thank S. Dittami for his help and critical reading of the manuscript. NK is granted by the Conseil Régional de Bretagne (Rannvro Breizh) and GIS Europe Mer. This work was supported by funding from IDEALG Grants ANR-10-BTBR-04-02 and 04-04 "investissements d'avenir", Biotechnologies-Bioressources.



## Author's contributions

NK and CH extracted the RNAs and DNA, tested the quality of RNAs and performed cDNA synthesis. NK chose the housekeeping genes, found homologues in *Chondrus crispus* genome and designed oligonucleotides for Q-PCR. SR carried out the Q-PCR experiments. Analysis of the data was performed by NK. The manuscript was written by NK, CB and JC and approved by all authors.

## References

1. Collén J, Davison I (1999) Stress tolerance and reactive oxygen metabolism in the intertidal red seaweeds *Mastocarpus stellatus* and *Chondrus crispus*. *Plant, Cell & Environment* 22: 1143–1151.
2. Cabello-Pasini A, von Wobeser EA, Figueroa F (2000) Photoinhibition of photosynthesis in *Macrocystis pyrifera* (phaeophyceae), *Chondrus crispus* (rhodophyceae) and *Ulva lactuca* (chlorophyceae) in outdoor culture systems. *J Photochem Photobiol B* 57: 169–78.
3. Lohrmann N, Logan B, Johnson A (2004) Seasonal acclimatization of antioxidants and photosynthesis in *Chondrus crispus* and *Mastocarpus stellatus*, two co-occurring red algae with differing stress tolerances. *Biol Bull* 207: 225–32.
4. Collén J, Hervé C, Guisle-Marsollier I, Léger J, Boyen C (2006) Expression profiling of *Chondrus crispus* (rhodophyta) after exposure to methyl jasmonate. *J Exp Bot* 57: 3869–81.
5. Collén J, Guisle-Marsollier I, Léger J, Boyen C (2007) Response of the transcriptome of the intertidal red seaweed *Chondrus crispus* to controlled and natural stresses. *New Phytol* 176: 45–55.
6. Hervé C, Tonon T, Collén J, Corre E, Boyen C (2006) NADPH oxidases in eukaryotes: red algae provide new hints! *Current Genetics* 49: 190–204.
7. Herve C, de Franco PO, Groisillier A, Tonon T, Boyen C (2008) New members of the glutathione transferase family discovered in red and brown algae. *Biochem J* 412: 535–544.
8. Gachon C, Mingam A, Charrier B (2004) Real-time pcr: What relevance to plant studies? *J Exp Bot* 55: 1445–1454.

9. Bustin S, Benes V, Nolan T, Pfaffl M (2005) Quantitative real-time rt-pcr ? a perspective. J Mol Endocrinol 34: 597–601.
10. Vandesompele J, de Preter K, Pattyn F, Poppe B, van Roy N, et al. (2002) Accurate normalization of real-time quantitative rt-pcr data by geometric averaging of multiple internal control genes. Genome Biol 3: R34.1–34.11.
11. Andersen C, Jensen J, Ørntoft T (2004) Normalization of real-time quantitative reverse transcription-pcr data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64: 5245–5250.
12. Pfaffl M, Tichopád A, Prgomet C, Neuvians T (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: Bestkeeper ? excel-based tool using pair-wise correlations. Biotechnol Lett 26: 509–515.
13. Spinsanti G, Panti C, Lazzeri E, Marsili L, Casini S, et al. (2006) Selection of reference genes for quantitative rt-pcr studies in striped dolphin (*Stenella coeruleoalba*) skin biopsies. BMC Mol Biol 7: 32.
14. Ingerslev H, Pettersen E, Jakobsen R, Petersen C, Wergeland H (2006) Expression profiling and validation of reference gene candidates in immune relevant tissues and cells from atlantic salmon (*Salmo salar* l.). Mol Immunol 43: 1194–1201.
15. Hoogewijs D, Houthoofd K, Matthijssens F, Vandesompele J, Vanfleteren J (2008) Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in *C. elegans*. BMC Mol Biol 9: 9.
16. Jain M, Nijhawan A, Tyagi A, Khurana J (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time pcr. Biochem Biophys Res Commun 345: 646–651.
17. Reid K, Olsson N, Schlosser J, Peng F, Lund S (2006) An optimized grapevine rna isolation procedure and statistical determination of reference genes for real-time rt-pcr during berry development. BMC Plant Biol 6: 27.

18. Chen L, Zhong H, Kuang J, Li J, Lu W, et al. (2011) Validation of reference genes for rt-qpcr studies of gene expression in banana fruit under different experimental conditions. *Planta* 234: 377–390.
19. le Bail A, Dittami S, de Franco PO, Rousvoal S, Cock M, et al. (2008) Normalisation genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. *BMC Mol Biol* 9: 75.
20. Lu N, Zang X, Zhang X, Chen H, Feng X, et al. (2012) Gene cloning, expression and activity analysis of manganese superoxide dismutase from two strains of *Gracilaria lemaneiformis* (gracilariaceae, rhodophyta) under heat stress. *Molecules* 17: 4522–32.
21. Ho C, Teoh S, Teo S, Rahim R, Phang S (2009) Profiling the transcriptome of *Gracilaria changii* (rhodophyta) in response to light deprivation. *Mar Biotechnol* (NY) 11: 513–9.
22. Teo S, Ho C, Teoh S, Rahim R, Phang S (2009) Transcriptomic analysis of *Gracilaria changii* (rhodophyta) in response to hyper- and hypoosmotic stresses. *J Phycol* 45: 1093–1099.
23. Wu X, Niu J, Huang A, Xu M, Wang G (2012) Selection of internal control gene for expression studies in porphyra haitanensis (rhodophyta) at different life-history stages. *J Phycol* 48: 1040–1044.
24. Jonas Collén, Betina Porcel, Wilfrid Carré, Steven G. Ball, Cristian Chaparro, Thierry Tonon, Tristan Barbeyron, Gurvan Michel, Benjamin Noel, Klaus Valentin, Marek Elias, François Artiguenave, Alok Arun, Jean-Marc Aury, José F. Barbosa-Neto, John H. Bothwell, François-Yves Bouget, Lorraine Brillet, Francisco Cabello-Hurtado, Salvador Capella-Gutiérrez, Bénédicte Charrier, Lionel Cladière, J. Mark Cock, Susana M. Coelho, Christophe Colleoni, Mirjam Czjzek, Corinne Da Silva, Ludovic Delage, France Denoeud, Philippe Deschamps, Simon M. Dittami, Toni Gabaldón, Claire M. M. Gachon, Agnès Groisillier, Cécile Hervé, Kamel Jabbari, Michael Katinka, Bernard Kloareg, Nathalie Kowalczyk, Karine Labadie, Catherine Leblanc, Pascal J. Lopez, Deirdre H. McLachlan, Laurence Meslet-Cladiere, Ahmed Moustafa, Zofia Nehr, Pi Nyvall Collén, Olivier Panaud, Frédéric Partensky, Julie Poulain, Stefan A. Rensing, Sylvie Rousvoal, Gaelle Samson, Aikaterini Symeonidi, Jean Weissenbach, Antonios Zambounis, Patrick Wincker and Catherine Boyen (2013) Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the archaeplastida. *PNAS*, in press.

25. Thomas F, Barbeyron T, Michel G (2011) Evaluation of reference genes for real-time quantitative pcr in the marine flavobacterium *Zobellia galactanivorans*. J Microbiol Methods 84: 61–6.

## Tables

**Table 1 - Culture conditions and duration**

Culture conditions and duration.

Sample	Treatment	Duration
A	100 % NSW - 100 % sunlight	0 h
F	200 % NSW - 35 % sunlight	3 h
G	100 % NSW - 35 % sunlight	3 h
H	200 % NSW - 100 % sunlight	3 h
I	100 % NSW - 100 % sunlight	3 h
M	100 % NSW - 100 % sunlight	5 h
C1	NSW	3 h
C2	10 mM H <sub>2</sub> O <sub>2</sub>	3 h
C3	200 mg.L <sup>-1</sup> Harpin	3 h
C4	200 $\mu$ M CdCl <sub>2</sub>	3 h
C5	200 $\mu$ M CuSO <sub>4</sub>	3 h
C6	200 $\mu$ M ZnCl <sub>2</sub>	3 h
C7	200 $\mu$ M AlCl <sub>3</sub>	3 h
C8	500 $\mu$ M dichlorvos	3 h
C9	500 $\mu$ M glyphosate	3 h
C10	500 $\mu$ M paraquat	3 h
C11	NSW + 1 % DMSO	3 h
C12	100 $\mu$ M methyl jasmonate DMSO (1 %)	3 h

**Table 2 - Sequences of the primers used for quantitative PCR in *C. crispus***

CDS: coding DNA sequence, PRL : PCR product length.

Gene code	Accession number	5' (forward) - 3' (reverse)	E (%)	R <sup>2</sup>	Tm product	PRL (bp)
IF4A-1	GSCHC2T00006625001	CCCGACAAACCGTGAGAAC CAAGAAGTTAATGGCGACACC	99	0.997	82.57	82
UBQ	GSCHC2T00016204001	TCAGACTACAACATCCAGAAGG CTTACGGCACACCATACGG	100	0.9945	85.87	126
eRF3	GSCHC2T00003248001	GAGGAGTTACCGTGTCTAAG CGACATCTTCAACTGTACTAAGC	94	0.9975	81.46	117
Tub $\beta$ -1	GSCHC2T00006992001	TTGCCGCCACCTTTATCG GAACTCCATCTCATCCATACCC	91	0.995	83.54	137
G6PdH	GSCHC2T00012633001	AAACGGACCTGAAAGTAGTAATG ATGCCCAAACACGAGAACC	105	0.9645	83.32	114
EF1 $\alpha$	GSCHC2T00017465001	CTCCGCAGCAACCATTTCG AGCATGACCACTGTGTTACC	103.5	0.9945	82.35	91
IF4A-2	GSCHC2T00002822001	CTCGCTCTGCCACATTC GCCTTGTCGTCCTGGTAATC	99.5	0.994	83.23	100
UbCE	GSCHC2T00005024001	GCCACCAGTTGTCAAAGTTTG GAGGATGTCCAAGCAGATACC	100	0.9975	79.88	80
Tub $\beta$ -2	GSCHC2T00009280001	CGACCTTGTTTCCGAATATCAG TCGTTCTCATACCCTTCTTCTTC	98	0.9855	80.17	84
IcdH	GSCHCT00014646001	AATCCGATTGCCTCTATATTTGC GGTCCCATCCAACCTTTCCTC	101	0.995	80.3	66
Actin	GSCHCT00010316001	ATTTGACTGCCTTTAACTCCATC GGGTCTCAATCTCCTTCTGC	99	0.998	80.61	135
Tub- $\alpha$	GSCHCT00004177001	TGCTCAATCGCCAACTCG ATACCCTCGCCACATACC	98	0.999	82.33	113
NUox	GSCHCT00009097001	AGTGCCATACACGATTCTGC AACAACAGCGAGTCCATCC	100	0.998	82.8	72
AAox	GSCHCT00004592001	TTTGAGAATCGCACTGGTAAAC ATACTCCTTATCCATCCACTCTG	89	0.968	85.13	133
Intron	GSCHC2T00016204001	GCTTGGGTTTCGCACATTAC GCATCAAGAAGAAGATTAAGCCT	100	0.999	85.27	146

## Figure Legends

### Figure 1 - Ct values of the 14 housekeeping genes considering all tested samples

The range of expression of the 14 candidate genes are represented by the average Ct values (diamonds) and Ct ranges (bars).

### Figure 2 - Mvalue analysis of the expression stability of the 14 housekeeping genes

The 14 genes are ranked according to their M value, calculated by the geNorm software [10]. Low values of M indicate that a gene is expressed very stably. A. all samples, B. culture samples, C. field samples.

### Figure 3 - NormFinder analysis of the expression stability of the 14 housekeeping genes

The stability of expression of the 14 candidate genes was calculated using the NormFinder method [11]. A. all samples, B. culture samples, C. field samples.

### Figure 4 - NormFinder analysis vs. geNorm analysis of the 14 housekeeping genes

The stability of expression of the 14 candidate genes was compared between the NormFinder method (abscissa axis) and the geNorm method (ordinate axis). ( $R^2 = 0.94$ ).

## **Additional Files**

### **Additional file 1 — RNA quantifications for environmental samples**

Concentration of RNA in environmental samples.

### **Additional file 2 — RNA quantifications for culture samples**

Concentration of RNA in culture samples.

### **Additional file 3 — RNA Bioanalyzer gel for environmental samples**

Quality of RNA in environmental samples.

### **Additional file 4 — RNA gel for culture samples**

Quality of RNA in culture samples. Ladder range : 0.2 - 10 kb.

### **Additional file 5 — Ct-range of the 14 housekeeping genes**

A. culture samples, B. field samples.



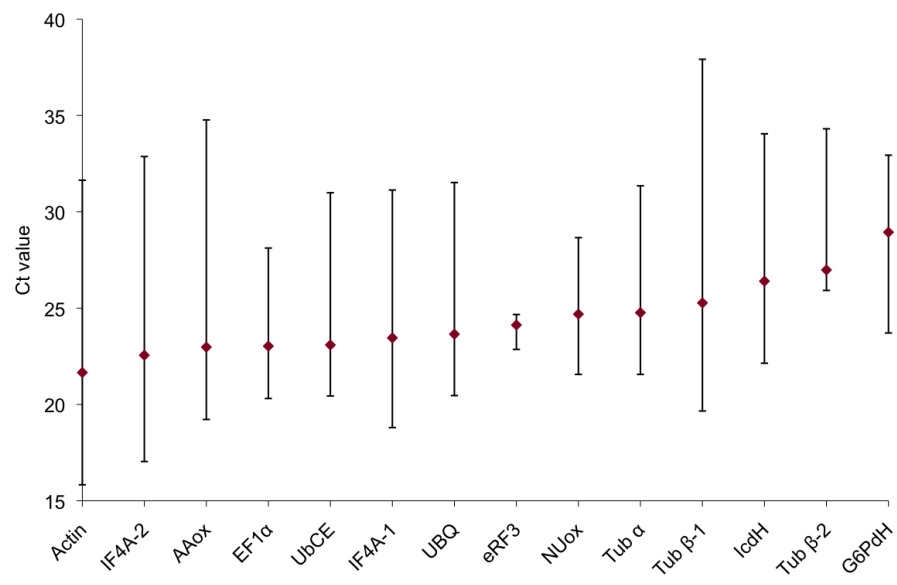


fig1

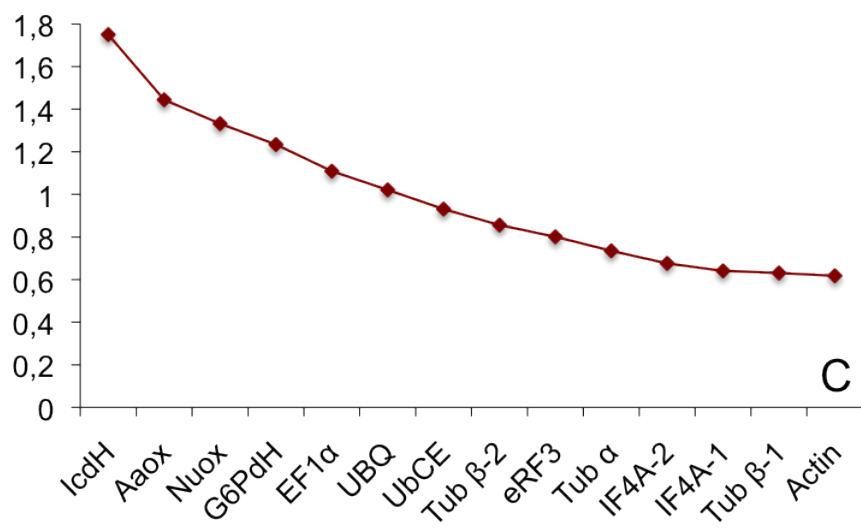
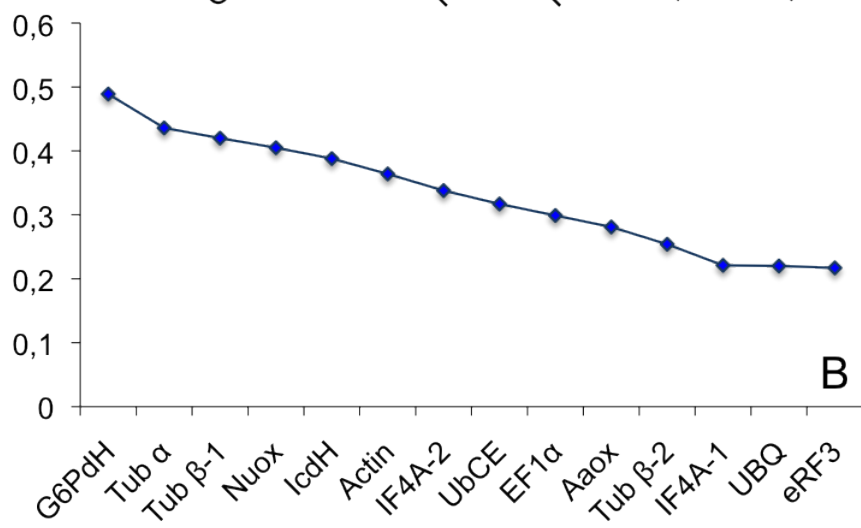
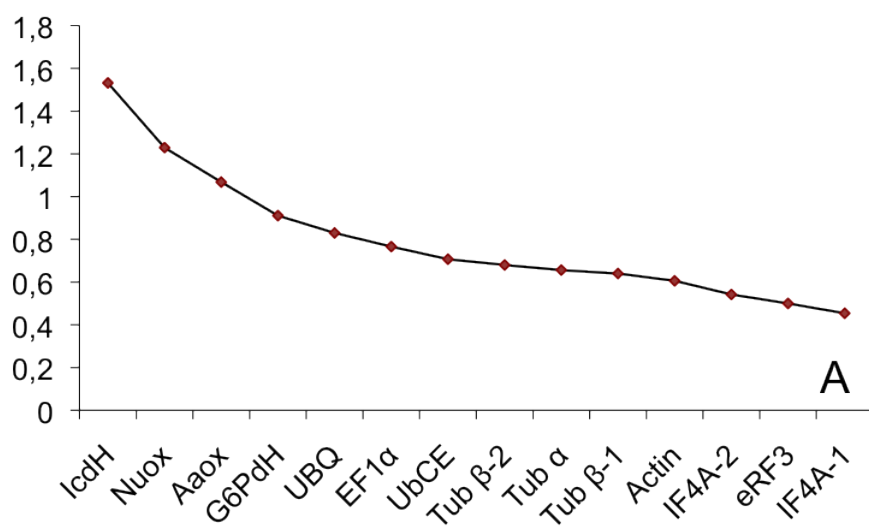


fig2

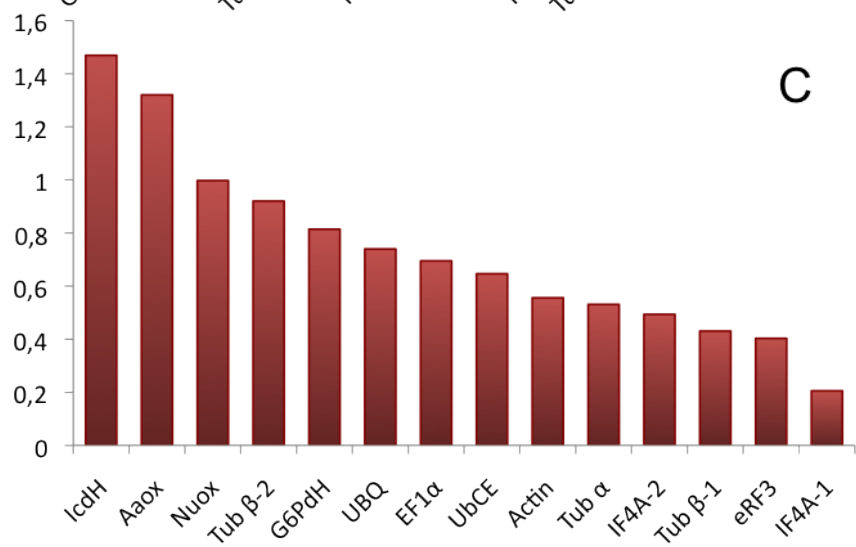
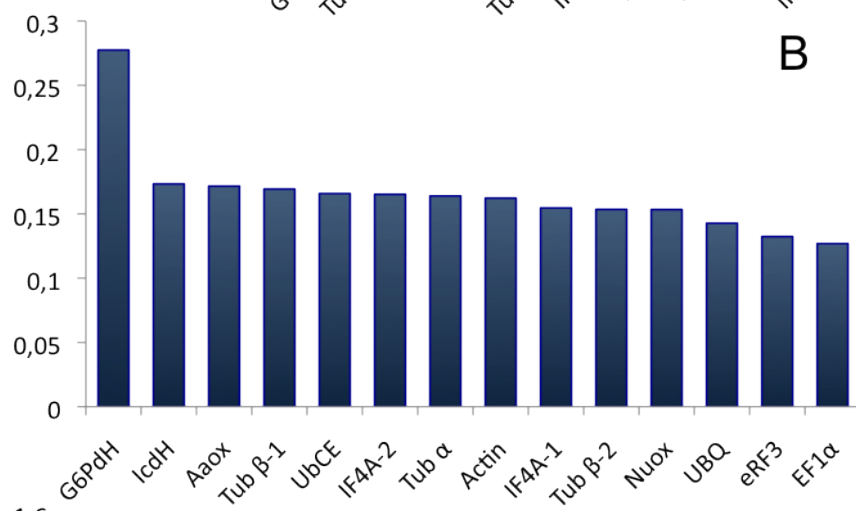
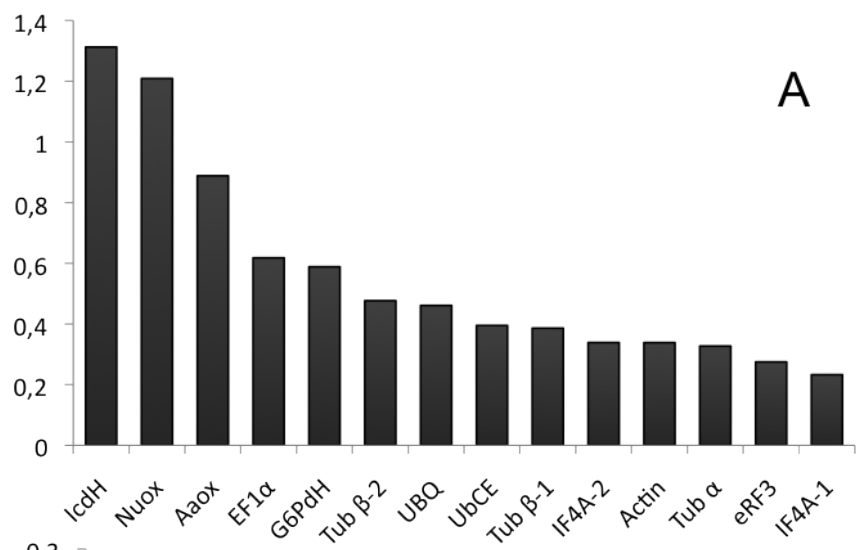


fig3

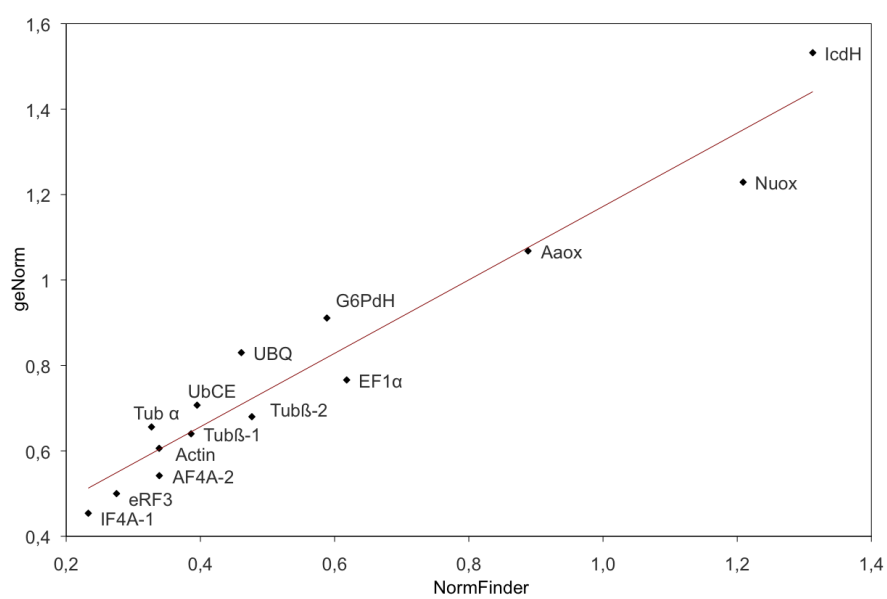


fig4

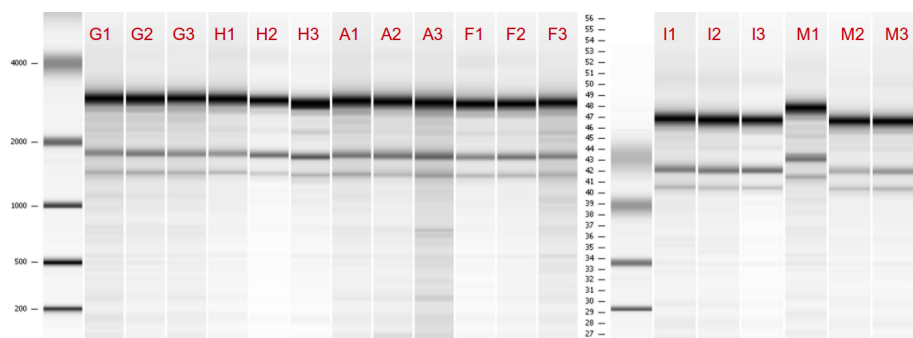
## 1 Supplementary data

Sample	[RNA] $\mu\text{g}.\mu\text{l}$	260/280	260/230
1A	0.4	1.9	1.72
2A	0.2	2.0	1.89
3A	0.2	2.0	1.78
1F	0.3	2.0	2.01
2F	0.2	2.0	1.95
3F	0.3	1.9	1.97
1G	0.3	2.0	2.00
2G	0.2	2.0	2.00
3G	0.3	2.0	1.99
1H	0.2	2.0	1.27
2H	0.3	2.0	1.97
3H	0.3	2.0	1.84
1I	0.2	2.0	1.85
2I	0.2	2.0	1.93
3I	0.2	2.0	1.86
1M	0.3	2.0	2.00
2M	0.1	2.0	1.75
3M	0.2	2.0	1.90

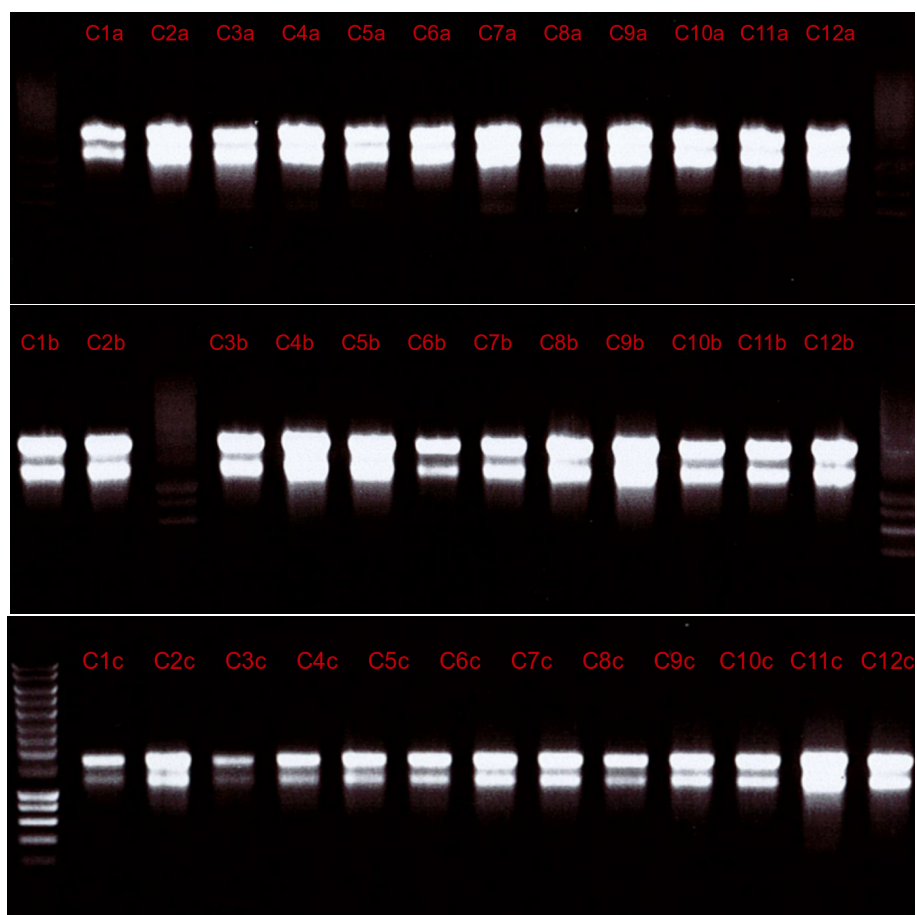
S1

Sample	[RNA] $\mu\text{g}, \mu\text{l}$	260/280
C1a	1.5	2.0
C2a	2.2	2.0
C3a	2.0	2.0
C4a	2.2	2.0
C5a	1.6	1.9
C6a	2.2	2.0
C7a	2.4	2.0
C8a	2.5	2.0
C9a	2.5	2.0
C10a	2.0	2.0
C11a	2.1	2.0
C12a	2.3	2.0
C1b	2.4	2.2
C2b	2.4	2.2
C3b	2.3	2.2
C4b	2.9	2.1
C5b	2.3	2.2
C6b	1.1	2.2
C7b	1.7	2.1
C8b	2.2	2.2
C9b	2.7	2.2
C10b	2.0	2.2
C11b	1.9	2.2
C12b	2.2	2.1
C1c	0.6	2.1
C2c	1.1	2.1
C3c	0.4	2.1
C4c	0.7	2.1
C5c	0.9	2.2
C6c	0.8	2.1
C7c	1.0	2.1
C8c	1.1	2.1
C9c	0.9	2.2
C10c	1.0	2.1
C11c	0.8	2.1
C12c	1.6	2.1

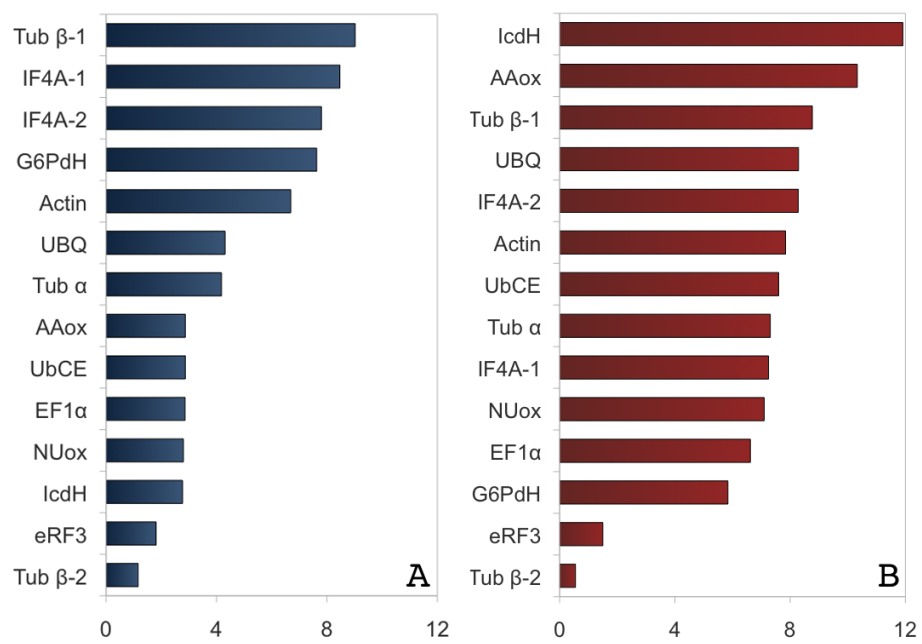
S2



S3



S4



S5



### 3.4 La photosynthèse en rouge

#### Préambule

L'objectif de l'étude suivante était de réanalyser le comportement de l'appareil photosynthétique des algues rouges face à la lumière. La photosynthèse peut être étudiée *in vivo* au moyen d'une méthode non-invasive : l'analyse de la fluorescence de la chlorophylle. Cette technique est couramment utilisée sur les végétaux verts (plantes terrestres, algues vertes) ainsi que sur les cyanobactéries, en revanche, rares sont les papiers traitants de la fluorescence chez les algues rouges. Comme décrit dans l'introduction de ce manuscrit, l'appareil photosynthétique des algues rouges est différent de celui des chlorophytes aussi bien que de celui des cyanobactéries. Les mécanismes liés à la photoprotection chez les algues rouges n'ont pas été pleinement élucidés et les acteurs connus à ce jour chez les autres organismes photosynthétiques, notamment ceux impliqués dans le NPQ sont absents.

Une des études les plus poussées, réalisée par Delphin et collaborateurs [138][139][140] sur des algues rouges unicellulaires extrêmophiles montrait une importante décroissance du rendement maximale de fluorescence lorsque les algues étaient soumises à une illumination préalable. Ils ont interprété ce quenching comme étant déclenché par l'établissement du gradient de pH photo-induit et indépendant d'une phosphorylation de protéines.

Des expériences plus récentes [214], utilisant une technique de fluorescence à très basse température résolue dans le temps, ont suggéré l'existence d'un transfert énergétique — le *spill-over* — entre les photosystèmes II et I.

Dans ce papier, nous réanalysons les phénomènes de quenching énergétique chez le modèle d'algue rouge multicellulaire *C. crispus*, en analysant les variations de fluorescence à température ambiante et à 77 K, combiné à de la spectrométrie permettant de mesurer l'état d'oxydo-réduction du cytochrome  $b_6/f$  et du PSI. En accord avec les études réalisées par Delphin et collaborateurs, nous observons une forte décroissance du rendement de fluorescence de la chlorophylle. La caractérisation de ce quenching en faisant varier le régime lumineux et en étudiant les changements de taille d'antenne du PSII et PSI nous a mené aux observations suivantes :

1. le quenching est déclenché par un simple pulse lumineux de 200 ms et met 1 minute à se mettre en place,
2. il n'est pas plus prononcé lorsque le régime lumineux est plus important,
3. il n'est pas lié à une diminution de taille d'antenne PSII, en revanche il se traduit par

l'augmentation de la taille d'antenne PSI.

4. il est entièrement réversible après une illumination rouge lointain, excitant préférentiellement le PSI,
5. il peut être induit à l'obscurité en réduisant le pool de PQ par des conditions anoxiques,
6. l'utilisation d'un agent découplant diminue son intensité mais ne permet pas de l'abolir.

Ces résultats nous ont conduits à écarter l'hypothèse de la formation d'un quencher au niveau de l'antenne-coeur du PSII, ainsi que celle de transitions d'états (migration des phycobilisomes entre PSII et PSI). En revanche nous proposons, d'une part, que le flux lumineux au niveau du PSII est régulé au moyen d'un *spill-over* énergétique vers le PSI, résultant d'un contact physique entre les deux photosystèmes. D'autre part, la mise en place de ce *spill-over* est déclenchée lors de la réduction du pool de PQ et a besoin d'un  $\Delta pH$  minimal, mais est indépendante du  $\Delta pH$  photo-induit.

## Modèle

Pour illustrer le phénomène d'extinction de fluorescence attribué au *spill-over*, nous proposons un modèle hypothétique schématisé en Fig. 30, basé sur les cinétiques observées d'induction et de relaxation.

Induction :

- au repos, dans l'état adapté à l'obscurité, le pool de PQ est oxydé, les photosystèmes sont déconnectés et la fluorescence du PSII est élevée.
- lors d'un premier pulse lumineux, le pool de PQ se réduit, les photosystèmes I et II entrent en contact, permettant de transférer une grande partie de l'énergie arrivant sur PSII vers PSI et entraînant ainsi une baisse du rendement de fluorescence du PSII.
- en prolongeant l'illumination, les photosystèmes subissent des changements conformationnels, amplifiant le transfert d'énergie de PSII vers PSI, et diminuant encore davantage la fluorescence émise par PSII.

Relaxation :

- l'illumination par de la lumière FR permet d'oxyder le pool de PQ et résulte dans la déconnection des photosystèmes. Cela se traduit par le retour partiel de l'émission de fluorescence par le PSII. En cas de réinduction par un pulse, l'accolement des photosystèmes dans leur nouvelle conformation entraîne immédiatement un transfert efficace de l'énergie de PSII vers PSI.

- en prolongeant l'illumination FR, les photosystèmes reviennent à leur conformation initiale, ce qui restaure le niveau de fluorescence maximale de PSII.

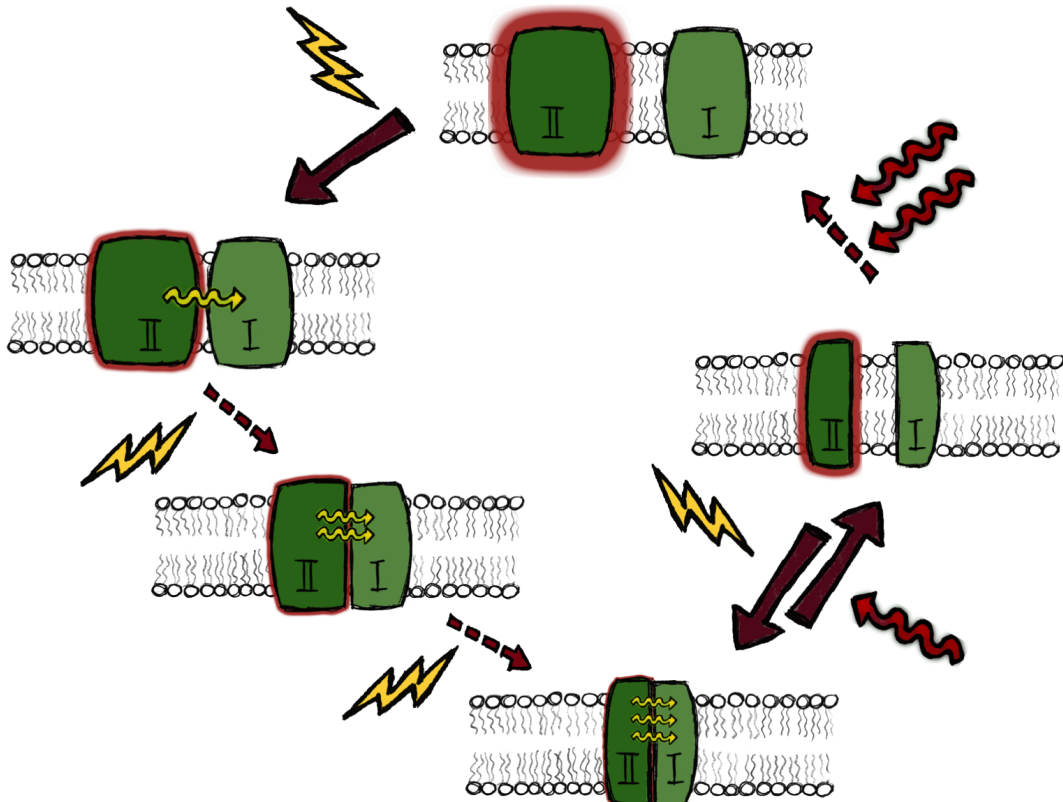
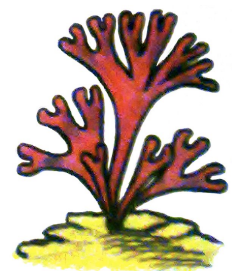


FIGURE 30: Modèle hypothétique de fonctionnement du *spill-over*.  
Hypothetical model of the spill-over mechanism.





Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: [www.elsevier.com/locate/bbabio](http://www.elsevier.com/locate/bbabio)

# Photosynthesis in *Chondrus crispus*: The contribution of energy spill-over in the regulation of excitonic flux

Nathalie Kowalczyk<sup>a,\*</sup>, Fabrice Rappaport<sup>b</sup>, Catherine Boyen<sup>a</sup>, Francis-André Wollman<sup>b</sup>, Jonas Collén<sup>a</sup>, Pierre Joliot<sup>b</sup>

<sup>a</sup> UMR 7139, Station Biologique de Roscoff, CNRS-UPMC, 29680 Roscoff, Brittany, France

<sup>b</sup> UMR 7141, Institut de Biologie Physico-Chimique, CNRS-UPMC, 75005 Paris, France

## ARTICLE INFO

### Article history:

Received 9 October 2012

Received in revised form 13 March 2013

Accepted 9 April 2013

Available online xxxx

### Keywords:

Photosynthesis

Red algae

*Chondrus crispus*

Chlorophyll fluorescence

Spill-over

Plastoquinone pool

## ABSTRACT

*Chondrus crispus* is a species of red algae that grows on rocks from the middle intertidal into the subtidal zones of the North Atlantic coasts. As such, it has to cope with strongly variable abiotic conditions. Here we studied the response of the photosynthetic apparatus of this red alga to illumination. We found that, as previously described in the case of the unicellular alga *Rhodella violacea* (E. Delphin et al., Plant Physiol. 27 118 (1998) 103–113), a single multi-turnover saturating pulse of light is sufficient to induce a strong quenching of fluorescence. To elucidate the mechanisms underlying this fluorescence quenching, we combined room temperature and 77 K fluorescence measurements with absorption spectroscopy to monitor the redox state of the different electron carriers in the chain. In addition, we studied the dependence of these various observables upon the excitation wavelength. This led us to identify energy spill-over from Photosystem II to Photosystem I rather than a qE-type non-photochemical quenching as the major source of fluorescence quenching that develops upon a series of 200 ms pulses of saturating light results, in line with the conclusion of Ley and Butler (Biochim. Biophys. Acta 592 (1980) 349–363) from their studies of the unicellular red alga *Porphyridium cruentum*. In addition, we show that the onset of this spill-over is triggered by the reduction of the plastoquinone pool.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

*Chondrus crispus* is a red macroalga living in the intertidal zone and widely represented on the rocky shores of the North Atlantic Ocean. Its economical and ecological importance has made it a historical model. It has been harvested for ages for its cell wall polysaccharides, the carrageenans, used in the food industry as thickening and gelling molecules.

The phylogenetic position of *Chondrus* among red algae makes it an interesting organism from the evolutionary point of view. According to the hypothesis of a unique primary endosymbiosis, all Archaeplastida are the result of a single event, where a eukaryotic cell has phagocytized an ancestor of current cyanobacteria, gaining the capacity to perform photosynthesis. Red algae are the sister group of the green lineage, which contains green algae and land-plants. This position gives them noticeable properties, particularly concerning their photosynthetic apparatus. On

the contrary to green organisms, red algae have non-appressed thylakoid membranes [1], with evenly distributed photosystems (PS) I and II. Like cyanobacteria, their PSII antenna is an extrinsic protein complex known as the phycobilisome (PBS) [2], composed of several blue and red pigments. A membrane bound Chl containing complex (LHCr for red-LHC), distantly related to green LHCl [3], acts as the light harvesting system of their monomeric PSI.

Photosynthesis allows plants, algae and some bacteria to transform CO<sub>2</sub> into organic matter, (CH<sub>2</sub>O)<sub>n</sub>, using solar energy. Light fluence and quality are not constant during the day and these organisms have developed different strategies to ensure an optimal energy capture. Their optimum photosynthesis is, in actual fact, a trade-off between a large absorption cross section to maximize the excitonic flux and a finely tuned excitation quenching to minimize the harmful effects of excess energy. In order to respond to an excess of incoming photon flux, green algae and plants have evolved different strategies whereby the excitonic flux to photosystems is tuned, and three different types of NPQ (for non-photochemical quenching) are commonly distinguished:

- qE is the dominant form of quenching that is associated with the acidification of the lumen [4]. It requires deep oxidation of carotenoid via the Xanthophyll cycle [5,6] and the presence of PsbS [7] (or its homolog in unicellular green algae), a transmembrane

Abbreviations: Chl, Chlorophyll; DCMU, 3-(3,4-dichlorophenyl)-1,1-dimethylurea; FR, far red; HA, hydroxylamine; LHC, light harvesting complex; NPQ, non-photochemical quenching; NSW, natural sea water; OCP, orange carotenoid protein; PBS, phycobilisome; PQ, plastoquinone; PS, photosystem; qE, energy-dependent quenching

\* Corresponding author at: UMR 7139, CNRS-UPMC, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, Brittany, France. Tel.: +33 298292323.

E-mail address: [nathalie@sb-roscoff.fr](mailto:nathalie@sb-roscoff.fr) (N. Kowalczyk).

0005-2728/\$ – see front matter © 2013 Elsevier B.V. All rights reserved.  
<http://dx.doi.org/10.1016/j.bbabio.2013.04.004>

Please cite this article as: N. Kowalczyk, et al., Photosynthesis in *Chondrus crispus*: The contribution of energy spill-over in the regulation of excitonic flux, Biochim. Biophys. Acta (2013), <http://dx.doi.org/10.1016/j.bbabio.2013.04.004>

protein. qE leads to the thermal dissipation of light energy. It is a fast relaxing process, a few minutes in the dark, concomitant with the relaxation of the light-built pH gradient.

- qT is the component related to state transition, a lateral displacement of light harvesting antennae along the membrane that contributes to balancing the excitonic flux between PSII and PSI [8]. It is known to be regulated by the redox state of the plastoquinone (PQ) pool and to involve the phosphorylation of the mobile antennae. The relaxation of qT quenching needs more time than qE, at least 20 min for the entire process.
- qI is the long term relaxing (several hours) component of NPQ [9,10].

The molecular mechanisms underlying NPQ in cyanobacteria and green species differ significantly. Cyanobacteria possess neither the enzymatic machinery required for the (de)phosphorylation identified in plants and green algae [11] nor the one involved in the xanthophyll cycle. Instead, the orange carotenoid protein (OCP) [12], a light-activated protein, bears a carotenoid and acts as an excitation and heat dissipator. In addition, state transitions occur on a much shorter time scale in cyanobacteria [11] although they also would be controlled by the redox state of the PQ pool. Yet instead of mobile membrane-bound LHClI antennae, the process would imply a migration of PBS on the external surface of thylakoids between PSII and PSI [11] and a change in spill-over between the two photosystems [13]. The occurrence of the latter regulation is supported, for instance, by the study of Mc Connell et al. who showed that, upon state transitions, the excitation energy absorbed by chlorophylls on the one hand, and by phycobilin pigments, on the other hand, is redistributed in an independent manner [14].

Some red algae live in the intertidal zone, which is a very dynamic environment. They have to cope with changing fluence and quality of light many times a day, depending on the tidal height, wave action, varying shading from other organisms and the cloud cover. However, the mechanisms related to the regulation of light capture and photoprotection in red algae are not yet fully understood (see [15,16] for reviews), especially the phenomenon known as NPQ, which has been widely described in green organisms [9,10].

According to their genomes [17,18], red algae do not possess any of the proteins characterized to date as being involved in NPQ in other photosynthetic species, like those involved in the xanthophyll cycle, PsbS, or OCP. However, studies have shown that NPQ does occur. In *Rhodella violacea* and *Porphyridium cruentum*, unicellular extremophile red algae, the maximum fluorescence yield undergoes a pronounced decrease upon illumination. In *R. violacea*, this quenching was interpreted as being triggered by the light-induced  $\Delta\text{pH}$  [19,20] and independent of protein phosphorylation [21].

Ley & Butler showed, using absorption and fluorescence spectra of frozen cells, that, in *P. cruentum*, most of the energy absorbed by the PBS is transferred to PSII but that energy transfer occurs from PSII to PSI. They estimated that, in the open state (i.e. when the primary quinone electron acceptor in PSII is oxidized), about half of the energy is transferred to PSI. This fraction increases to near unity when PSII photochemical traps are closed [22]. In addition, they observed that preilluminating the cells with green light resulted in an increased photo-oxidation rate of  $P_{700}$  when compared to that observed in dark-adapted cells and that this specific photochemical enhancement was only observed when exciting PBS [23], suggesting that, although PBS is specifically connected to PSII, energy transfer may occur from PSII to PSI which requires a physical association between PSI and PSII [24]. This physical contact, controlled by a redox parameter would account for the modulation of the PSII fluorescence yield [25]. More recently, Yokono and coworkers, using time resolved fluorescence at  $-196^\circ\text{C}$  on *P. cruentum* and macroalgae, also suggested the occurrence of an energy spill-over between PSII and PSI [26], thereby accounting for the excitation quenching observed under more physiological conditions.

In this paper we reinvestigate the mechanisms of excitation quenching in red algae using the macroalga *C. crispus* as a model. As previously described in the case of the unicellular alga *R. violacea* [19,20] a single multi-turnover saturating pulse of light is sufficient to induce a strong quenching of fluorescence. To elucidate the mechanisms underlying this fluorescence quenching, we combined the room temperature and 77 K fluorescence measurements and absorption spectroscopy to monitor the redox state of the different electron carriers in the chain. We thereby characterized the dependence of this quenching on the illumination regime and analyzed the changes in the PSI and PSII antenna size induced by the quenching.

This leads us to propose that, along the lines of the model that emerged from the studies of *P. cruentum* (see above), on the one hand, the excitonic flux at the level of PSII is regulated by the extent of energy spill-over toward PSI and that, on the other hand, this spill-over is triggered by the reduction of the PQ pool.

## 2. Materials & methods

### 2.1. Biological material

Non-fertile gametophytes of *C. crispus* Stackhouse (Gigartinales, Rhodophyta) were collected near the port of Bloscon in Roscoff, Brittany (France) and were kept in a thermostated natural sea water (NSW) bath at  $13^\circ\text{C}$  in 12 h/12 h light/dark regime under a light fluence of  $20\ \mu\text{E m}^{-2}\text{ s}^{-1}$ .

Algae were dark adapted for 1 h before measurements. The time course of fluorescence changes was analyzed with a thallus cut to the size of the sample holder. During the experiments, the samples were kept in a small volume of NSW.

### 2.2. Functional analysis

Fluorescence and absorption changes were measured using a JTS spectrophotometer (Biologic) in which fluorescence is sampled using short (4  $\mu\text{s}$  duration) flashes with negligible actinic effect. Fluorescence yield  $F_M$  is measured at the end of a 200 ms pulse of saturating red light that induced a full reduction of  $Q_A$ . If not stated otherwise, actinic green light was provided by a LED peaking at 520 nm (half band  $\sim 25$  nm) that excites both chlorophyll and PBSs. Experiments were performed using green detecting flashes at 554 nm that excite preferentially the PBSs, except in Fig. 7 in which the kinetics of fluorescence changes were measured using detecting flashes at 554 and 420 nm.

$F_0$ : fluorescence of a dark adapted (at least 1 h) material detected by a weak and short flash of green light.  $F_S$ : fluorescence measured right before a saturating pulse.  $F_M$  and  $F'_M$ : maximum fluorescence, measured immediately after a saturating pulse, respectively on a dark adapted and a preilluminated material.

$P_{700}$  redox state was measured as changes in the absorbance at 705 nm (detecting flash 15  $\mu\text{s}$  duration).  $P_{700}$  oxidation was induced by a far red (FR) actinic illumination provided by a LED peaking at 740 nm. The absorption corresponding to the fully oxidized  $P_{700}$  was measured by superimposing to the FR light a pulse of saturating red light (20 ms) that induces a full oxidation of  $P_{700}$ . The absorption changes associated with the oxidation of c-type cytochrome were measured at 555 nm. As described in more details in the text, the actinic light was provided by green light LED and Wratten filters # 16 and # 44A (see Fig. S1).

Fluorescence emission spectra were measured with a USB2000 Ocean Optics spectrophotometer. The disks of *C. crispus* thalli were in a metallic sample holder that was immersed in liquid nitrogen. The sample holder was connected to a Y-shaped light guide that conducts the excitation light to the sample and collected the emitted fluorescence. The exciting light was provided by a Schott KL 1500 lamp and the wavelengths selected using appropriate filters.

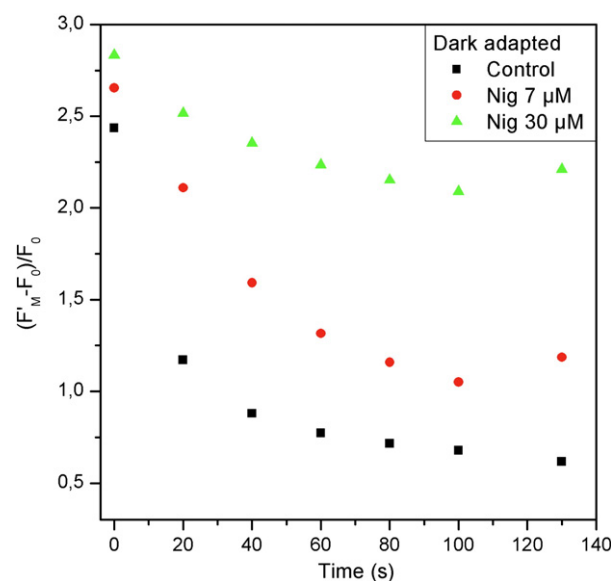


### 3. Results & discussion

#### 3.1. A fluorescence collapse induced by a train of saturating light pulses

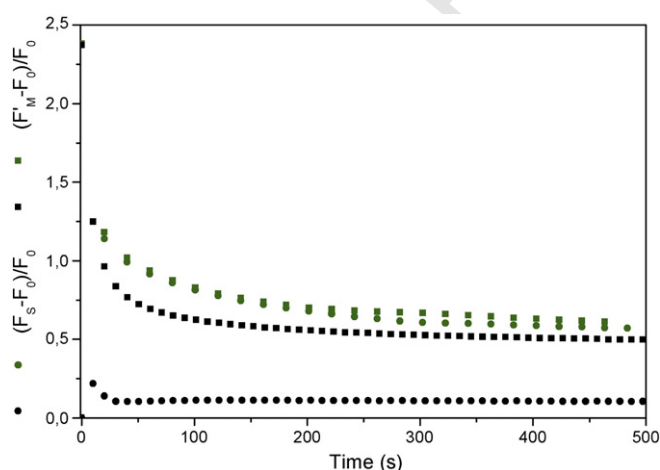
We first assessed the variation of the maximum fluorescence yield under conditions usually promoting photoprotective processes, i.e. high light conditions ( $k_i$  PSII  $\sim 1000 \text{ s}^{-1}$ ). As shown in Fig. 1, the fluorescence yield  $(F'_M - F_0) / F_0$ , measured after 200 ms pulses of saturating light, steadily decreased from 2.4 for the dark adapted state to 0.7 after a few minutes of illumination. Thus, as in *R. violacea* [19] fluorescence quenching processes develop in *C. crispus*. Notably, after  $\sim 200 \text{ s}$  of illumination the difference between  $F'_M$  and  $F_S$  progressively increased, likely indicating the activation of the  $\text{CO}_2$  assimilation cycle as described by [27].

A remarkable observation made by Delphin et al. [20] with *R. violacea* is that a mere train of light pulses, in the absence of any continuous illumination, is sufficient to induce a strong quenching. As shown in Fig. 1, the same applies to *C. crispus*. Actually, we observed that the pulse train induced an even more pronounced decrease of the  $(F'_M - F_0) / F_0$  level than the superposition of pulses to a continuous light. Moreover, a single pulse of 200 ms duration (i.e. shorter than the 800 ms pulse used in [20]) was sufficient to promote a large fraction of this quenching as evidenced by the decrease of  $(F'_M - F_0) / F_0$  observed upon the second pulse in the series ( $\sim 1.25$  to be compared to 2.4 for the dark adapted state). Finally, the overall time course of NPQ was similar in the presence or absence of the continuous exciting light. To our knowledge this is specific to red algae suggesting that the process governing the onset of NPQ in this part of the photosynthetic realm is unique. As a first step to characterize the origin of the quenching we compared its sensitivity to nigericin, a  $\text{H}^+/\text{K}^+$  exchanger commonly used to dissipate transmembrane  $\Delta\text{pH}$ . Fig. 2 shows the consequences of the addition of nigericin at two different concentrations, 7 and 30  $\mu\text{M}$ , both being, to our experience, larger than the concentration required to reach saturation of the uncoupling effect in whole leaves from land-plants (2  $\mu\text{M}$ ). First we note that the  $(F'_M - F_0) / F_0$  measured upon the first pulse increased upon the addition of nigericin, showing that i) the dark adapted state is constitutively quenched to a certain extent, and ii) the extent of this quenching decreases upon the addition of nigericin. Since the hydrolysis of ATP by the chloroplast ATP synthase builds up a  $\Delta\text{pH}$  in the dark [28,29], the effect of nigericin addition on the  $F'_M$  of the dark-adapted state is consistent with the model proposed by Delphin et

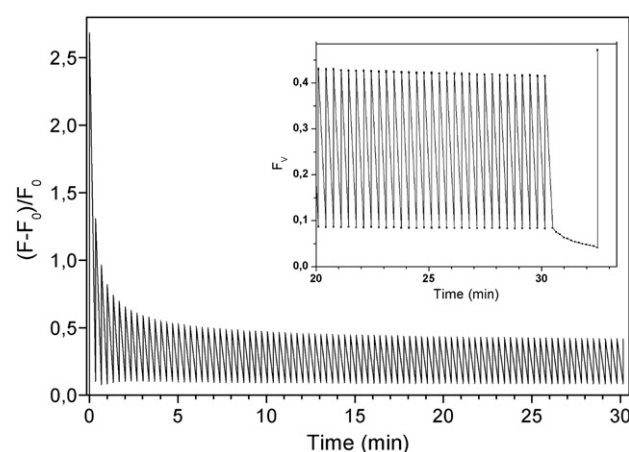


**Fig. 2.** Fluorescence changes  $(F'_M - F_0) / F_0$  induced by a series of pulses in the presence or absence of nigericin (Nig). The concentrations of Nig were 7  $\mu\text{M}$  (circles) and 30  $\mu\text{M}$  (triangles). A thallus, dark adapted for more than 1 h, has been illuminated by a train of 6 pulses 20 s apart. At the end of the series, a last pulse is fired after 30 s of darkness. The 3 experiments have been performed using contiguous fragments of the same thallus.

al. [19] according to which the quenching depends on the luminal pH. Accordingly, the quenching observed after the first pulse, in the absence of nigericin, would stem from the light-induced increase of the luminal pH, as proposed by Delphin et al. Yet, several pieces of data are not fully consistent with this framework. First, although the rationale for the fact that uncouplers relieve a  $\Delta\text{pH}$  induced quenching is that they promote a faster relaxation of the light-induced  $\Delta\text{pH}$ , the remaining quenching observed after a series of pulses still relaxes slowly even at the highest nigericin concentration used as illustrated in Fig. 2 by the very moderate recovery of  $(F'_M - F_0) / F_0$  after 30 s of relaxation in the dark. Second, a single pulse of 200 ms duration is sufficient to induce a pronounced quenching but yet seems unlikely to allow enough photochemical turnovers to promote a marked acidification of the lumen. Third, as shown in Fig. 3, which displays the development and relaxation of the quenching on longer time scales, the quenching had hardly relaxed after several



**Fig. 1.** Fluorescence changes measured on a thallus of *Chondrus crispus*. Green symbols: a thallus, dark adapted for more than 1 h, has been illuminated with green light ( $k_i$  PSII  $\sim 1000 \text{ s}^{-1}$ ) to which pulses of saturating light (200 ms duration) were superimposed allowing one to probe the  $F'_M$  values. Black symbols: the thallus, dark adapted for more than 1 h has been illuminated by a train of saturating pulses (0.1 Hz). The two experiments were performed using contiguous fragments of the same thallus.



**Fig. 3.** Fluorescence changes measured on a dark adapted thallus illuminated for 30 min by a train of pulses. Pulse duration: 200 ms. Dark intervals between pulses: 20 s. The curve displays the fluorescence measured immediately before and after the pulse ( $(F_S - F_0) / F_0$  and  $(F'_M - F_0) / F_0$ ), respectively. The inset shows the last pulses of the series on an expanded time scale. At the end of the pulse series, the fluorescence decay is measured during the following 2 min of darkness. A last pulse is fired at the end of this period of darkness to probe the recovery of  $(F'_M - F_0) / F_0$ .

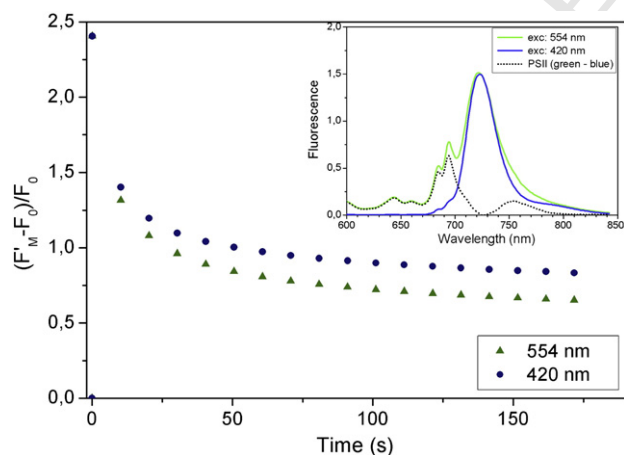
minutes of darkness following a 30 min illumination. We estimated the half-time of the relaxation of the quenching as ~15 min, and this makes it a much slower process than the relaxation of the light-induced  $\Delta pH$  in land-plants [30]. Even though the life-time of the luminal acidification is not known in red algae it is unlikely to be as long-lived as the quenching reported here, in as much as the building up of a light-induced proton motive force is expected to fully activate the ATP synthase [31,32] which, in turn, should result in a short-lived proton gradient [33]. We thus conclude that, in addition to the luminal pH, another parameter likely contributes to determining the occurrence and the extent of the quenching in *C. crispus*.

As mentioned in the introduction, there are three types of hypothesis that one should consider to account for fluorescence quenching [34]: a change in the absorption cross section of PSII resulting from the migration of a mobile antenna from PSII to PSI, the formation of a quencher within the PSII photochemical units, and, finally, an increased connectivity between PSII and PSI photochemical units whereby excitation spills over to the photochemical trap of lowest energy i.e. PSI.

We first tested the hypothesis of a lateral migration of the PBSs which could, as proposed in cyanobacteria [35], be preferentially associated with PSII or PSI, depending on the redox state of the plastoquinone. One way to assess the occurrence of putative changes in the relative excitonic connectivity between PBSs and PSII or PSI is to specifically excite PBSs and to probe the relative extent of PSII and PSI fluorescences. As shown in the inset of Fig. 4, the low temperature fluorescence emission spectrum displays no contribution of PBS when excited in blue, as expected since PBS hardly absorb at this wavelength. Notably, the contribution of fluorescence emitted by PSII photochemical units when excited at 420 nm drastically decreased thus showing that this wavelength preferentially excites PSI, most presumably via LHCr, a membrane bound Chl containing complex that acts as the light harvesting system of monomeric PSI [24]. Conversely, when exciting at 554 nm, the low temperature emission spectrum clearly displayed the emission from PBSs, in the 650 nm region [24], and those ascribed to PSII at 685 and 695 nm [24,25]. We repeated the experiments shown in Fig. 1 where the quenching was induced by a series of saturating pulses (0.1 Hz)

and the fluorescence was probed by weak flashes exciting preferentially either PBSs (554 nm) or LHCr (420 nm). As shown in Fig. 4, both antenna systems underwent a quenching of similar extent (see below for a discussion of the slightly more pronounced quenching when using green detecting light), thus showing that it does not specifically affect either one. This argues against a dramatic change in the connectivity between PBSs and PSII since, in this case, the quenching would be specifically observed when exciting in their absorption band, i.e. at 554 nm. We are thus, at this stage, left with two remaining types of hypothesis: the occurrence of spill-over or the light-induced formation of a quencher.

For energy spill-over to occur from PSII to PSI, the two photochemical units must be excitonically coupled. In this case, an exciton within the PSII antenna has a given probability to be transferred to and trapped by the PSI photochemical unit. This probability is expected to increase when PSII centers are unable to trap the incoming excitation flux owing to their reduced  $Q_A$ , so that an increase in the PSI antenna size occurring upon the reduction of  $Q_A$  is commonly considered to be the signature of energy spill-over. We thus measured the dependence of the PSI turnover rate on the redox state of  $Q_A$ , following the steps of Ley and Butler in their study of the excitonic connectivity between PSII and PSI photochemical units in the red alga *P. cruentum* [23]. To assess this rate, we measured, at room temperature, the extent of the oxidation of the electron donors to PSI induced by short and weak light pulses. We submitted the sample to a 200 to 400  $\mu s$  weak pulse of actinic light to induce a limited number of PSI turnover. As a probe of this turnover number, we used the relative extent of cytochrome *f* and/or cytochrome *c*<sub>6</sub> oxidation 300  $\mu s$  afterwards i.e. after completion of the oxidation phase of cyt *c* and *f*. As shown in the Supplementary data (Fig. S2) no significant re-reduction of cytochrome *f* occurred at this delay time. After normalization to the full oxidation extent, measured after a saturating actinic pulse, the fraction of cytochrome oxidized by the short and weak pulse did not exceed 10%. We checked that it was linearly related to the duration of the pulse, or in other terms a reliable measure of the number of PSI photochemical turnovers. These conditions being satisfied, the extent of cytochrome oxidation provides a measure that is proportional to the number of photons absorbed by the PSI photochemical units and thus, the light intensity being kept constant, to its absorption cross section. In order to assess whether PBS can transfer excitonic energy to PSI we used two different types of exciting light, hereafter denoted as A and B, as described in the Material & methods section. A-type pulses (in the 530–560 nm region, see the spectra in Fig. S1) preferentially excited PBSs whereas B-type pulses (in the 510–540 nm region) excites Chl, carotenoids and PBSs. We then compared the extent of cytochrome oxidation in the absence and presence of DCMU (30  $\mu M$ ) and hydroxylamine (3 mM) to induce the stable reduction of  $Q_A^-$  (see Table 1A). In both cases, nigericin was present in order to relax the quenching as much as possible (see above). With both the A and B type pulses, the fraction of oxidized cytochrome increased in the presence of  $Q_A^-$ , as expected in the spill-over hypothesis. Moreover, this increase was more pronounced when exciting more specifically PBSs (two fold for the A type



**Fig. 4.** Fluorescence changes induced by a series of pulses separated by 10 s of dark, measured using blue or green detecting flashes. Green symbols: the fluorescence yield was measured using detecting flash at 554 nm. Blue symbols: the fluorescence yield was measured using detecting flash at 420 nm.  $(F_M - F_0)/F_0$  ratio is lower under blue detecting flashes than under green detecting flashes excitation (1.8 and 2.49 respectively) owing to the larger contribution of PSI fluorescence in conditions where only Chl's are excited. Variable fluorescence  $(F_M - F_0)/F_0$  measured using blue detecting flashes has been normalized to that obtained with green detecting flashes. Inset: fluorescence emission spectra measured at 77 K: a dark adapted thallus has been frozen in liquid nitrogen and the fluorescence emission spectra measured using green (green line) or blue (blue line) excitation. Fluorescence emission at 725 nm has been normalized for the two spectra. Dotted line: difference between the spectra obtained using green and blue excitation. This spectrum reflects the PBS + PSII emission spectra.

**Table 1**

Fraction of oxidized cyt *c*<sub>556</sub> and cyt *f* induced by a short and weak pulse. A-type pulses (in the 530–560 nm region, see the spectra in Fig. S1) preferentially excited PBSs whereas B-type pulses (in the 510–540 nm) region excites Chl, carotenoids and PBSs.

	Pulse A	Pulse B	$F_i/F_0$	$F_M/F_0$
<b>A</b>				
Control + nigericin (30 $\mu M$ )	4.91	5.45	1	3.81
DCMU + HA + nigericin (30 $\mu M$ )	9.88	8.81	3.75	3.82
Ratio	2.01	1.61		
<b>B</b>				
Control dark-adapted	3.6	5.87	1	3.48
Control pre-illuminated	10.0	10.8	1.2	1.53
Ratio	2.78	1.83		

pulse with respect to 1.61 fold for the B type ones), as expected if the additional excitonic input observed when  $Q_A$  is reduced, stems from the excitonic coupling between PBSs and PSI. Altogether these data support the occurrence of a noticeable spill-over from PSII photochemical units to PSI ones, in good agreement with the conclusions reached by Ley and Butler with the unicellular red alga *P. cruentum* [22,23]. As put forward by these authors such an energy transfer may account for the rather unusually low  $F_M/F_0$  ratio, 3.8, observed despite the addition of nigericin which, as shown above, relieves part of the quenching (even though the  $F_M/F_0$  ratio observed in *P. cruentum* by Ley and Butler [22] was, 1.85, i.e. significantly lower than reported here for *C. crispus*). The present conclusion that in dark-adapted *C. crispus*, a significant energy spill-over may occur from PSII to PSI and that it likely accounts for the fluorescence quenching is also consistent with the conclusion drawn by Yokono et al. [26] who studied the fluorescence emission spectra and fluorescence lifetimes in several macroscopic red algae. They also suggested that the extent of this energy spill-over would be linearly correlated to the relative absorbance of PBS with respect to that of Chl binding protein and this may account for the interspecies variability of the  $F_M/F_0$  ratio mentioned above.

Although it makes this hypothesis more plausible, the occurrence of spill-over in the dark-adapted state does not necessarily imply that the light-induced quenching stems from an increased spill-over. To address this issue we compared the PSI absorption cross section in conditions of potentially minimal spill-over, i.e. oxidized  $Q_A$  and dark-adapted thalli and of potentially maximal spill-over, i.e. after induction of a strong quenching ( $F_M/F_0$ ) by a train of 40 saturating pulses (Table 1B). In the latter conditions, the extent of cytochrome *f* oxidation was assessed 500 ms after the last saturating pulse in the series in order to allow for the complete relaxation of the primary and secondary PSI electron donors and acceptors. Importantly, 500 ms was not enough to allow for the complete reoxidation of  $Q_A$  (we estimate from the fluorescence yield changes that about 40% of  $Q_A^-$  was still reduced at this time so that 60% of PSII units were not promoting spill-over). The preillumination induced an increase of the PSI antenna size 2.78 and 1.76 fold with A and B type pulses, respectively. Thus, as described above, the increase in the PSI absorption cross section is more pronounced when exciting more specifically PBSs. In addition and even more importantly, the increase in the PSI absorption cross section is larger than in dark-adapted conditions (2.78 versus 2, and this even though  $Q_A$  is not fully reduced in the preilluminated case), as expected if indeed fluorescence quenching stems from spill-over. Thus, a significant fraction of the excitonic energy absorbed by PBSs is efficiently transferred to PSI and this efficiency increases when i)  $Q_A$  is reduced and ii) when fluorescence is quenched. These various characteristics are those expected in the spill-over model and, taking into account the fact that  $Q_A$  was reduced in only ~40% of PSII, one would expect that, when quenching is maximum and  $Q_A$  is fully reduced, the relative contribution of spill-over in the excitonic flux to PSI is  $1.78/0.4 = 4.45$  fold larger than that of the PSI antenna (we note that, using the same relative units, when quenching is minimal and  $Q_A$  is fully reduced, the relative contribution of energy transfer from PSII to PSI would be equal to one). Our conclusions regarding *C. crispus* are thus consistent with those reached earlier by Ley and Butler with *P. cruentum*. Indeed, not only did they evidence the occurrence of a significant energy transfer from PSII to PSI photochemical units in dark adapted samples [22,23] but they showed that the extent of this energy spill-over can be modulated by exciting preferentially PSII (state II inducing light) or PSI (state I inducing light) [25]. The parameters that trigger these changes yet remain to be identified.

### 3.2. A fluorescence quenching controlled by the redox state of the plastoquinone pool

We have observed that, in *C. crispus*, a single saturating pulse is sufficient to promote a significant quenching as previously observed

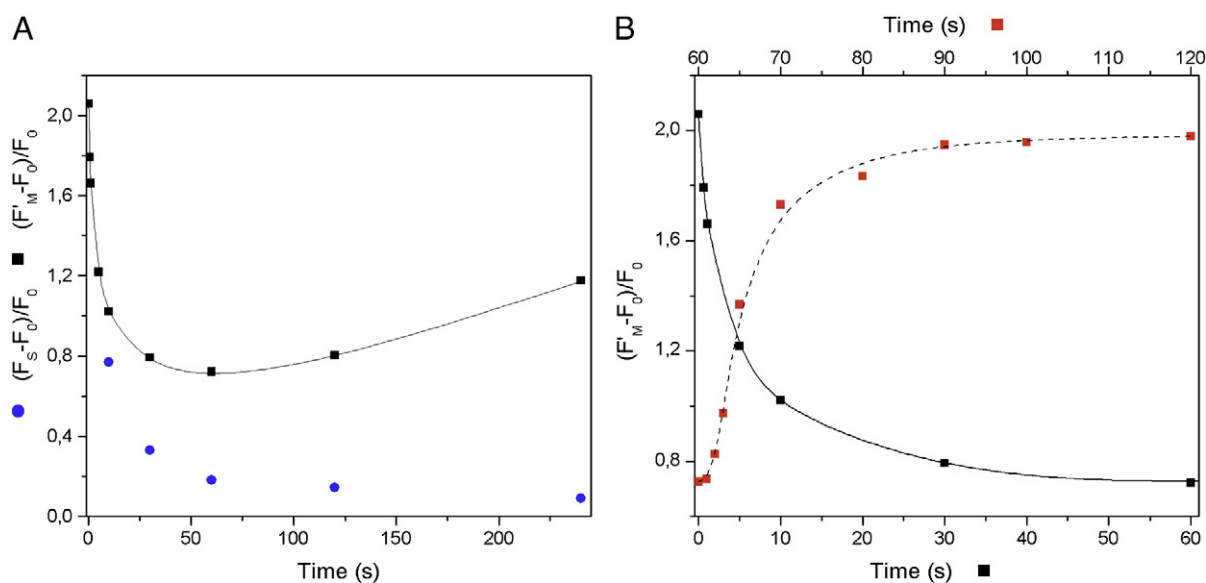
by Delphin et al. in *R. violacea* [20]. To gain further insights into the mechanism behind this process, we focused on the characterization of the quenching induced by a single saturating pulse. In the experiments described below (Fig. 5A and B) we applied a 200 ms light-pulse and probed its effect with several observables described below. One methodological difficulty in such an experiment is that the complete relaxation of the quenching induced by a pulse has to be reached before applying another pulse. Ley and Butler [25] reported that, in *P. cruentum*, energy spill-over, which, according to the results presented above, is responsible for most of the fluorescence quenching, decreases upon illumination with light exciting specifically PSI. We checked that this applies to *C. crispus* as well and that full relaxation of the quenching could be achieved by applying a weak continuous illumination of FR light ( $\lambda \geq 720$  nm, 2 min), as illustrated for example in Fig. 5B. As pointed out in [25], this mere finding hints at the redox state of the electron carrier upstream of PSI, i.e., under all likelihood, the plastoquinones, including  $Q_A$ , being responsible for the modulation of the fluorescence yield.

We first analyzed the time course of the fluorescence quenching after the pulse. Fig. 5A shows the relative fluorescence yield measured just before ( $(F_S - F_0) / F_0$ , blue symbols) and after ( $(F'_M - F_0) / F_0$ , black symbols) a second saturating pulse applied at various delay time following the first one. As can be seen, these two parameters underwent a marked decrease in the first tens of seconds following the triggering pulse. The maximum of the quenching was observed at ~60 s and decreased as the time in the dark was further increased. As regard to  $(F_S - F_0) / F_0$ , it continuously decreased with increasing delays between pulses.  $(F'_M - F_S)$  increased similarly indicating the progressive, but sluggish, reoxidation of  $Q_A$ . Yet, 4 min was not sufficient for  $F_S$  to reach the  $F_0$  level, suggesting that the reoxidation of the plastoquinone pool in the dark is a very slow process.

As noted above, the finding that FR illumination promotes the reversion of the quenching, suggests that the extent of the latter is related to the redox state of the electron transfer chain upstream of PSI. To further characterize the kinetics of this FR induced recovery, we first submitted the sample to one saturating pulse. It was then kept in the dark for 60 s to allow the quenching to develop completely and then submitted to FR illumination of various durations. The variation of  $(F'_M - F_0) / F_0$  as a function of this duration is shown in Fig. 5B (red symbols) and is compared to the evolution of the quenching induced by a single pulse (black symbols). The increase of  $(F'_M - F_0) / F_0$  upon FR illumination first displayed a lag-phase, ~1 s, and then rose to reach the  $(F_M - F_0) / F_0$  value with a half-time of ~5 s. A 5 fold increase in the FR intensity expectedly shortened the lag phase but did not significantly accelerate the relaxation of the quenching thus showing that it is kinetically limited by a non-photochemical process (not shown). As regard to the parameter(s) that trigger(s) the quenching, the observation that FR illumination enhances the relaxation of the quenching does not argue for a contribution of the luminal pH. Indeed, the FR-induced turnovers of PSI likely contribute to building up – rather than dissipating – a trans-membrane proton motive force. On the contrary it supports the hypothesis that the quenching is controlled by the redox state of plastoquinol or  $Q_A$ .

The critical role of the redox state of the plastoquinone pool is further illustrated in Fig. 6 that shows the kinetics of absorption changes associated with the redox changes of the chlorophyll dimer  $P_{700}$ , borne by PSI. These absorption changes were induced by FR illumination and the maximum signal was assessed by using a pulse of saturating light (see [36] and Materials & methods). Expectedly, we observed a bleaching, indicating the oxidation of  $P_{700}$ , which reached a steady state level of ~90% showing that the PSII to PSI excitation ratio is smaller than ~1/10 under these illumination conditions. We compared the kinetics of oxidation of  $P_{700}$  in the dark-adapted and quenched states. As above, the latter were induced by a single pulse and the oxidation kinetics of  $P_{700}$  were measured 1 min, a time where the quenching is

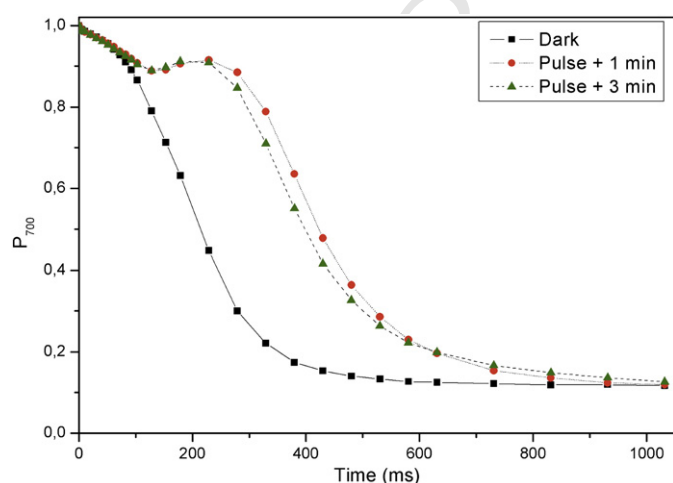




**Fig. 5.** Fluorescence yield measured as the function of the time of darkness following one saturating pulse. The thallus has been illuminated by a first pulse (200 ms duration) followed by a second pulse given at various time. A:  $(F_s - F_0)/F_0$  (blue symbols) and  $(F_M - F_0)/F_0$  (black symbols) are plotted as a function of the dark time between the two pulses. At the end of each experiment, the thallus was illuminated by Far Red ( $100 \mu\text{E m}^{-2} \text{s}^{-1}$  for 1 min) in order to restore the  $F_M$  level (see text). B: black symbols  $(F_M - F_0)/F_0$ : same as in Fig. 5A on an expanded time scale. Red symbols: the thallus has been first illuminated by one pulse. After 1 min of darkness the thallus was submitted to a FR illumination ( $100 \mu\text{E m}^{-2} \text{s}^{-1}$ ) followed by a second pulse.  $(F_M - F_0)/F_0$  measured after the second pulse is plotted as a function of the duration of the FR illumination.

maximal, and 3 min after the pulse, to probe the redox state of the photosynthetic chain when the relaxation has started to develop.

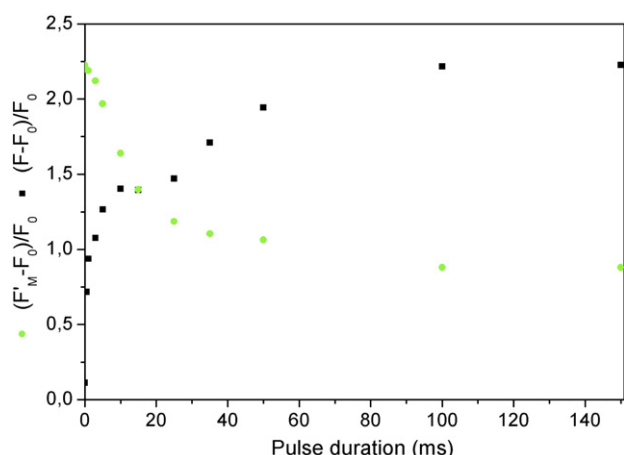
The most insightful feature in this comparison is the lag-phase that precedes the oxidation of  $P_{700}$  1 min after the pulse. As such, this lag shows that more PSI turnovers are required than in the dark-adapted case to reach a given oxidation state. This may either stem from the electron carrier upstream of  $P_{700}$  being more reduced or from a more efficient cyclic electron flow. We favor the former hypothesis for the two following main reasons: i) we have shown above that the fluorescence yield  $F_s$  measured 1 min after the pulse is larger than  $F_0$  suggesting that a fraction of  $Q_A$ , and hence of the plastoquinone pool, is still reduced, ii) we observed here that, after the lag-phase, the oxidation time-course of  $P_{700}$  at 1 min had a similar slope than in the dark-adapted case. Contrary to this, a more efficient



**Fig. 6.** The  $P_{700}$  redox changes upon a FR illumination. Intensity of far red illumination:  $500 \mu\text{E m}^{-2} \text{s}^{-1}$ . Black symbols: dark adapted thallus. Red symbols: the thallus has been submitted to a 200 ms saturating pulse and, after 1 min of dark, submitted to FR illumination. Green symbols: the thallus has been submitted to a 200 ms saturating pulse and, after 3 min of darkness, submitted to FR illumination.

cyclic electron flow would, everything being the same, result in more efficient reduction competing with the oxidation and this would be manifested by a gentler slope in the oxidation phase (see [4,37] for a discussion). If the various absolute rates involved remain constant, the time required to reach the half-oxidation of  $P_{700}$  is proportional to the number of turnovers undergone by one PSI during this time laps. Roughly, it is two-fold larger 1 min after the pulse than in the dark adapted state. This shows that the overall pool of electron donor to PSI, including plastoquinone, is larger in the former case than in the latter. In addition, the fact that this results in the appearance of a lag-phase shows that the additional electron carriers have a more negative redox potential than those available in the dark-adapted state making plastoquinol good candidates. Assuming 6 plastoquinols, or 12 electron equivalents, in the pool, the 200 ms duration of the lag-phase translates into a maximal PSI turnover-rate of  $\sim 60 \text{s}^{-1}$ . The observation that the oxidation kinetics of  $P_{700}$  was similar 1 min and 3 min after the pulse shows that the reoxidation of the plastoquinone pool in the dark is very slow as already discussed above on the basis of the data shown in Figs. 1 and 2. This, again, contrasts with plants [38] or unicellular algae [39] where the Plastid Terminal Oxidase PTOX drives the reoxidation of the plastoquinones much more quickly.

At this stage, we can tentatively propose that the quenching is directly controlled by the redox state of the plastoquinone pool and/or  $Q_A$ . To discriminate between these two candidates, we studied the extent of quenching as a function of the duration of the saturating pulse. Fig. 7 shows the variation of the fluorescence yield measured after one pulse of varying duration (black symbols). This corresponds to the so-called OJIP curve [40] and displays its characteristic feature: i) a fast initial rise developing in the 1–2 ms time range, which is assigned to the reduction of  $Q_A$ , ii) following this fast component, a further progressive fluorescence increase developing in the tens of ms time range and assigned to the reduction of the plastoquinone pool, and, lastly, iii) a slower component, in the 40–100 ms time range associated with the reduction of the PSI electron acceptors [41,42]. In Fig. 7, in green symbols, is also shown the maximum fluorescence yield measured 1 min after the first pulse, again as a function of the duration of this pulse. Although 2 ms of saturating light



**Fig. 7.** Fluorescence changes as a function of the duration of the pulse of light. The sample was submitted to a pulse of varying duration. The fluorescence yield changes resulting from this illumination were probed immediately after this pulse and are depicted by the black symbols,  $(F' - F_0) / F_0$ , as a function of the duration of the pulse. A second pulse of 250 ms duration of saturating light was applied 60 s afterwards, in order to probe the changes in  $F'_M$ , as a function of the duration of the first pulse (Green symbols:  $(F'_M - F_0) / F_0$ ).

intensity is sufficiently long to reduce  $Q_A$  in all the PSII reaction centers (see [43] for example), we observed hardly any quenching after a 1 or 2 ms long pulse, showing that the reduction of  $Q_A$  alone is not sufficient to promote the quenching. Notably, the quenching had almost fully developed (~85%) after a 30 ms pulse and its increase paralleled in time the reduction of the plastoquinone pool. This suggests that the parameter triggering the quenching is the redox state of the plastoquinone pool rather than that of  $Q_A$ .

To further assess this conclusion, we attempted to tune the redox state of the plastoquinone pool in the dark. To this end we relied on the chlororespiratory chain [44–46] which involves a NADPH:plastoquinone oxidoreductase (NDH) on the one hand and a plastoquinol oxidase (PTOX) on the other hand. Under anoxic conditions, one expects the latter to be inhibited and thus the plastoquinone pool to be reduced in the dark owing to the activity of the former. We noted above that the long-lasting high fluorescence level,  $F_S$ , observed after a pulse was indicative of a sluggish PTOX activity. Despite this sluggishness the plastoquinone pool is mostly oxidized in dark-adapted thalli from *C. crispus* suggesting that the NDH activity is even lower than its PTOX chlororespiratory accomplice. Consistent with this, anoxic conditions did not induce any change in  $F_0$  nor in  $F_M$  that would reflect the reduction of  $Q_A$  or of the plastoquinone pool in the dark. We thus resorted to a different strategy to achieve the reduction of the plastoquinone pool without soliciting PSII. It has been reported that, in land plants, illumination enhances the chlororespiratory influx of electrons into the plastoquinone pool [47,48]. Even though the molecular rationale for this observation is presently unknown, we reasoned that the same may apply to *C. crispus*. We thus first imposed anoxic conditions by flushing argon and then illuminated the sample by a weak FR light. Expectedly, this protocol maintained the plastoquinone pool in its oxidized state in the light as checked by the time course of  $P_{700}$  redox changes. Accordingly, when probed 1 s after the end of the FR preillumination,  $F'_M$  was equal to  $F_M$  in aerobic conditions and only slightly lower in anoxic ones (see Table 2). Consistent with the empirical assumption that the chlororespiratory influx is enhanced by the FR illumination, the  $F'_M$  value was lower when probed 60 s after the preillumination and this applied both in aerobic and anoxic conditions, although to a smaller extent in the former case than in the latter. In aerobic conditions this decrease in the  $F'_M$  value reversed 240 s after the end of the preillumination, likely owing to the PTOX activity. Contrary to this, the quenching continued to develop in anoxic conditions. In addition, 240 s after the end of the preillumination, the oxidation

**Table 2**

$(F'_M - F_0) / F_0$  measured in aerobiosis and anaerobiosis. For anoxic condition, the thallus has been incubated for more than 1 h in anaerobiosis.

$F'_M - F_0 / F_0$ in <i>Chondrus crispus</i> in aerobiosis and anaerobiosis.		
Time after 10 min of Far Red illumination	Aerobiosis	Anoxic
1 s	2.18	1.92
60 s	1.88	0.99
240 s	2.15	0.77

time course of  $P_{700}$  showed, in anoxic conditions, the signature described above of a fully reduced plastoquinone pool. Hence the quenching develops upon the reduction of the plastoquinone pool even when neither PSII nor PSI turns over. Moreover the fact that, in anoxic conditions, it develops in the dark (or at least in the dark period following the FR illumination) argues against any triggering event that would relax rather than to develop in the dark, such as a transmembrane  $\Delta pH$ .

Even though the present results point to the redox state as being prominent among the parameters that control the fluorescence quenching, one cannot ignore the simple fact that the addition of nigericin decreases the extent of the quenching. This delicate duet played by the redox state of the intersystem electron carriers and the lumenal pH is further illustrated in Fig. S3 that displays again the experiment presented in Fig. 2 now followed by an FR illumination. Dark-adapted thalli were first exposed to a series of 6 pulses in the presence of nigericin (7  $\mu M$ ). As described above, this induced a quenching of smaller extent than in the absence of uncoupler. As discussed above, Far Red (180 s) efficiently enhanced the relaxation of the quenching but an additional series of 6 pulses given to this far red preilluminated thallus induced a quenching of larger extent than in the dark-adapted case. Thus far red seems to have opposite effects on the quenching since it accelerates not only its relaxation but also its onset when pulses are superimposed. These two opposite effects can be reconciled assuming that Far red illumination induces on the one hand the reoxidation of the plastoquinone pool and this is sufficient to relieve the quenching, but in addition it acidifies the lumen by promoting PSI turnovers and hence facilitates the onset of the quenching when the plastoquinone pool is reduced by the first pulse in the second series. In summary, spill-over would require both a reduced plastoquinone pool and a low lumenal pH but the former would prevail over the latter.

#### 4. Conclusion

At present, there is no consensus for the mechanism of state transitions in phycobilisome-containing cyanobacteria (for review, [49,50]) but the physical displacement of PBS through lateral diffusion is considered as the main mechanism accounting for the light-induced decrease in the PSII fluorescence yield [51]. In contrast, several studies conducted with red algae rather support energy spill-over from PSII to PSI, as a major source of fluorescence quenching [19,23]. This is in agreement with Fluorescence Recovery After Photobleaching studies conducted with *P. cruentum* that did not support the occurrence of a large scale diffusion of PBSs in red algae [52]. As will be discussed below, in an effort to build a unified mechanistic view from several experimental observations that arose from our study of the light-induced quenching in dark-adapted thalli of *C. crispus*, we acknowledge that the extent of quenching shows a dependency on the lumenal pH but is caused by an increased sensitization of PSI by PBS – mostly through changes in spill-over – under the redox control of the PQ pool.

We argued that the two processes for regulating the extent of PSI sensitization by a PSII peripheral antenna – that is spill-over and direct change in PSI sensitization by PBS in phycobilisome-containing organisms – are commonly discriminated by comparing the extent

of the excitation quenching when probed with wavelengths preferentially absorbed by PBS or the chlorophyll antennae. This rationale has been followed as well in [14], in their study of the regulation of energy transfer in cyanobacteria. Importantly, we observed, as they did, that the quenching induced by a series of pulses is systematically slightly larger when probed in the PBS absorption wavelength range than in the Chl one (see Fig. 4). As proposed by McConnell et al. this finding supports the view that, under conditions promoting spill-over, a small fraction of the energy absorbed by PBS would by-pass PSII and be directly transferred to PSI. We thus confirmed that this direct change in PBS/PSI connection contributes for a minor part to the light-induced quenching in *C. crispus*. However based on the dependence of the PSI turnover rate on the redox state of  $Q_A$  we have argued that, in line with the conclusions drawn by Ley and Butler from their studies of *P. cruentum*, and more recently from studies with multicellular red algae [26], there is already an excitonic coupling between PBS and PSI due to significant energy spill-over in dark-adapted thalli. We further demonstrated that this spill-over increases with the light-induced quenching in *C. crispus*. The increase of this excitonic connectivity allowing direct and indirect energy transfer between PBS and PSI also has been proposed to account for light-induced changes in fluorescence yield in unicellular red algae [25,53] and in cyanobacteria such as *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 [14].

We have described above that the quenching, which we interpret as resulting from an energy transfer from the closed PSII traps to PSI, displays highly multiphasic kinetics. It involves first a quickly developing component that accounts for half of the total quenching and is followed by slower components that are not fully completed after 30 min of illumination. At least two conditions must be fulfilled for spill-over to occur efficiently: firstly, the PSII photochemical traps must be closed and, secondly, a physical contact between the PSII and PSI photochemical units must exist to allow the excitonic coupling of neighboring pigments. According to Förster theory [54,55], one of the parameters that determine the efficiency of the energy transfer is their distance. Thus, if one assumes that the distance separating a given PSII unit from a given PSI unit is the main parameter determining the time required for them to be in physical contact, then the multiple kinetic components would merely reflect the large distribution of distances between PSII and PSI photochemical units. Alternatively, one may assume that the distance between PSII and PSI units is homogeneously distributed, in which case the multiple kinetic components would reflect the heterogeneity of the rate of spill-over promotion. Imaging techniques suggests that PBSs are aligned along the membrane [56] thereby providing support to the spatially homogenous but kinetically heterogeneous hypothesis.

Our data support the idea that the light-induced quenching is mainly controlled by the redox state of plastoquinone, as proposed previously by Ley and Butler in *P. cruentum* [25]. It remains however that nigericin and other proton carriers relieve part of this quenching, an observation that is consistent with those made by Delphin et al. that led them to interpret this quenching as being controlled by the luminal pH [20]. These apparently conflicting conclusions can be reconciled assuming that the physical contact that, as just discussed, is required for spill-over can be tuned by electrostatics effects involving protonable residues borne by the PSII and PSI protein complexes: at high pH, the unprotonated state(s) would loosen the interactions between PSI and PSII units whereas at lower pH, these would be promoted. This would account for the observation that the extent of quenching induced by a series of saturating pulses is decreased upon the addition of nigericin (Fig. 2). It would also account for the fact that addition of nigericin results in an increase of  $(F_M - F_0) / F_0$  ratio. Indeed in the dark, the hydrolysis of mitochondrial ATP generates a  $\Delta pH$  [29]. In *Chlamydomonas*, the luminal pH thereby sustained in the dark has been estimated to be  $\sim 5.6$  [28] and it is, expectedly, larger in the presence of nigericin ( $\sim 7.2$ ). In the present model, the triggering parameter would thus be the reduction of the PQ pool

but the extent of the quenching would depend on the luminal pH as this would contribute to loosening or tightening the connectivity between PSII and PSI units. Finally we emphasize that, according to the present model, the event that triggers the onset of spill-over is the reduction of the PQ pool indeed. This is reminiscent of the lateral redistribution of the LHCII antenna during the state 1–state 2 transition seen in green algae [57–59]. Yet, state transitions imply a cascade of biochemical reactions, such as antenna phosphorylation, which requires several minutes to reach completion and kinetically control the quenching [58]. On the contrary, upon oxidation of the PQ pool in red algae induced by a strong FR illumination the spill-over efficiency decreases much more quickly which argues for the spill-over efficiency being controlled by a totally different mechanism that yet remains to be elucidated.

## 5. Uncited reference

[60]

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbabo.2013.04.004>.

## References

- [1] A.W.D. Larkum, M. Vesik, Algal plastids: their fine structure and properties, *Photosynth. Algae* (2003) 11–28.
- [2] A.A. Arteni, L.N. Liu, T.J. Aartsma, Y.Z. Zhang, B.C. Zhou, E.J. Boekema, Structure and organization of phycobilisomes on membranes of the red alga *Porphyridium cruentum*, *Photosynth. Res.* 95 (2008) 169–174.
- [3] H.N. Su, B.B. Xie, X.Y. Zhang, B.C. Zhou, Y.Z. Zhang, The supramolecular architecture, function, and regulation of thylakoid membranes in red algae: an overview, *Photosynth. Res.* 104 (2010) 73–87.
- [4] C.A. Wraight, A.R. Crofts, Energy-dependent quenching of chlorophyll  $\alpha$  fluorescence in isolated chloroplasts, *Eur. J. Biochem.* 17 (1970) 319–327.
- [5] B. Demmig-Adams, Carotenoids and photoprotection in plants: a role for the xanthophyll zeaxanthin, *Biochim. Biophys. Acta* 1020 (1990) 1–24.
- [6] A.M. Gilmore, H.Y. Yamamoto, Zeaxanthin formation and energy-dependent fluorescence quenching in pea chloroplasts under artificially mediated linear and cyclic electron transport, *Plant Physiol.* 96 (1991) 635–643.
- [7] K.K. Niyogi, X.-P. Li, V. Rosenberg, H.-S. Jung, Is PsbS the site of non-photochemical quenching in photosynthesis? *J. Exp. Bot.* 56 (2005) 375–382.
- [8] M. Iwai, M. Yokono, N. Inada, J. Minagawa, Live-cell imaging of photosystem II antenna dissociation during state transitions, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 2337–2342.
- [9] P. Muller, X. Li, K.K. Niyogi, Non-photochemical quenching. A response to excess light energy, *Plant Physiol.* 125 (2001) 1558–1566.
- [10] S. Eberhard, G. Finazzi, F.A. Wollman, The dynamics of photosynthesis, *Annu. Rev. Genet.* 42 (2008) 463–515.
- [11] C.W. Mullineaux, D. Emlin-Jones, State transitions: an example of acclimation to low-light stress, *J. Exp. Bot.* 56 (2005) 389–393.
- [12] D. Kirilovsky, C.A. Kerfeld, The orange carotenoid protein in photoprotection of photosystem II in cyanobacteria, *Biochim. Biophys. Acta* 1817 (2012) 158–166.
- [13] C. Huang, X. Yuana, J. Zhao, D.A. Bryant, Kinetic analyses of state transitions of the cyanobacterium *Synechococcus* sp. PCC 7002 and its mutant strains impaired in electron transport, *Biochim. Biophys. Acta* 1607 (2003) 121–130.
- [14] M.D. McConnell, R. Koop, S. Vasil'ev, D. Bruce, Regulation of the distribution of chlorophyll and phycobilin-absorbed excitation energy in cyanobacteria. A structure-based model for the light state transition, *Plant Physiol.* 130 (2002) 1201–1212.
- [15] D.C. Fork, K. Satoh, The control by state transitions of the distribution of excitation energy in photosynthesis, *Annu. Rev. Plant Physiol.* 37 (1986) 335–361.
- [16] J. Biggins, D. Bruce, Regulation of excitation energy transfer in organisms containing phycobilins, *Photosynth. Res.* 20 (1989) 1–34.
- [17] M. Matsuzaki, O. Misumi, T. Shin-I, S. Maruyama, M. Takahara, S.Y. Miyagishima, T. Mori, K. Nishida, F. Yagisawa, K. Nishida, Y. Yoshida, Y. Nishimura, S. Nakao, T. Kobayashi, Y. Momoyama, T. Higashiyama, A. Minoda, M. Sano, H. Nomoto, K. Oishi, H. Hayashi, F. Ohta, S. Nishizaka, S. Haga, S. Miura, T. Morishita, Y. Kabeya, K. Terasawa, Y. Suzuki, Y. Ishii, S. Asakawa, H. Takano, N. Ohta, H. Kuroiwa, K. Tanaka, N. Shimizu, S. Sugano, N. Sato, H. Nozaki, N. Ogasawara, Y. Kohara, T. Kuroiwa, Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D, *Nature* 428 (2004) 653–657.
- [18] J. Collén, et al., Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida, *PNAS* (2013) in press.
- [19] E. Delphin, J.-C. Duval, A.-L. Etienne, D. Kirilovsky, State transitions or  $\Delta pH$ -dependent quenching of photosystem II fluorescence in red algae, *Biochemistry* 35 (1996) 9435–9445.



- [20] E. Delphin, J.-C. Duval, A.-L. Etienne, D. Kirilovsky, DeltapH-dependent photosystem II fluorescence quenching induced by saturating, multiturnover pulses in red algae, *Plant Physiol.* 118 (1998) 103–113.
- [21] E. Delphin, J.-C. Duval, D. Kirilovsky, Comparison of state 1-state 2 transitions in the green alga *Chlamydomonas reinhardtii* and in the red alga *Rhodella violacea*: effect of kinase and phosphatase inhibitors, *Biochim. Biophys. Acta* 1232 (1995) 91–95.
- [22] A.C. Ley, W.L. Butler, Efficiency of energy transfer from photosystem II to photosystem I in *Porphyridium cruentum*, *Proc. Natl. Acad. Sci. U. S. A.* 73 (1976) 3957–3960.
- [23] A.C. Ley, W.L. Butler, Energy transfer from photosystem II to photosystem I in *Porphyridium cruentum*, *Biochim. Biophys. Acta* 462 (1977) 290–294.
- [24] A.C. Ley, W.L. Butler, D.A. Bryant, A.N. Glazer, Isolation and function of allophycocyanin B of *Porphyridium cruentum*, *Plant Physiol.* 59 (1977) 974–980.
- [25] A.C. Ley, W.L. Butler, Energy distribution in the photochemical apparatus of *Porphyridium cruentum* in state I and state II, *Biochim. Biophys. Acta* 592 (1980) 349–363.
- [26] M. Yokono, A. Murakami, S. Akimoto, Excitation energy transfer between photosystem II and photosystem I in red algae: larger amounts of phycobilisome enhance spillover, *Biochim. Biophys. Acta* 1807 (2011) 847–853.
- [27] H. Kautsky, A. Hirsch, Neue Versuche zur Kohlensäureassimilation, *Naturwissenschaften* 19 (1931) 964–964.
- [28] G. Finazzi, F. Rappaport, In vivo characterization of the electrochemical proton gradient generated in darkness in green algae and its kinetic effects on cytochrome b6f turnover, *Biochemistry* 37 (1998) 9999–10005.
- [29] F. Rappaport, G. Finazzi, Y. Pierre, P. Bennis, A new electrochemical gradient generator in thylakoid membranes of green algae, *Biochemistry* 38 (1999) 2040–2047.
- [30] P. Joliot, G. Finazzi, Proton equilibration in the chloroplast modulates multiphasic kinetics of non-photochemical quenching of fluorescence in plants, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 12728–12733.
- [31] P. Gräber, H.T. Witt, Relations between the electrical potential, pH gradient, proton flux and phosphorylation in the photosynthetic membrane, *Biochim. Biophys. Acta* 423 (1976) 141–163.
- [32] W. Junge, The critical electric potential difference for photophosphorylation. Its relation to the chemiosmotic hypothesis and to the triggering requirements of the ATPase system, *Eur. J. Biochem.* 14 (1970) 582–592.
- [33] P. Joliot, A. Joliot, Quantification of the electrochemical proton gradient and activation of ATP synthase in leaves, *Biochim. Biophys. Acta* 1777 (2008) 676–683.
- [34] G.H. Krause, E. Weis, Chlorophyll fluorescence and photosynthesis: the basics, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42 (1991) 313–349.
- [35] C.W. Mullineaux, J.F. Allen, State 1-state 2 transitions in the cyanobacterium *Synechococcus* 6301 are controlled by the redox state of electron carriers between photosystems I and II, *Photosynth. Res.* 23 (1990) 297–311.
- [36] C. Klughammer, U. Schreiber, An improved method, using saturating light pulses, for the determination of photosystem I quantum yield via P700+–absorbance changes at 830 nm, *Planta* 192 (1994) 261–268.
- [37] P. Joliot, G.N. Johnson, Regulation of cyclic and linear electron flow in higher plants, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 13317–13322.
- [38] P. Joliot, A. Joliot, Excitation transfer between photosynthetic units: the 1964 experiment, *Photosynth. Res.* 76 (2003) 241–245.
- [39] L. Houille-Vernes, F. Rappaport, F.-A. Wollman, J. Alric, X. Johnson, Plastid terminal oxidase 2 (PTOX2) is the major oxidase involved in chlororespiration in *Chlamydomonas*, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 20820–20825.
- [40] R.J. Strasser, Govindjee, The Fo and the O-J-I-P fluorescence rise in higher plants and algae, in: J.H. Argyroudi-Akoyunoglou (Ed.), *Regulation of Chloroplast Biogenesis*, Plenum Press, New York, 1991, pp. 423–426.
- [41] R. Delosme, Etude de l'induction de fluorescence des algues vertes et des chloroplastes au début d'une illumination intense, *Biochim. Biophys. Acta* 143 (1967) 108–128.
- [42] G. Schansker, S.Z. Tóth, R.J. Strasser, Methylviologen and dibromothymoquinone treatments of pea leaves reveal the role of photosystem I in the Chl a fluorescence rise OJIP, *Biochim. Biophys. Acta* 1706 (2005) 250–261.
- [43] F. Rappaport, D. Beal, A. Joliot, P. Joliot, On the advantages of using green light to study fluorescence yield changes in leaves, *Biochim. Biophys. Acta* 1767 (2007) 56–65.
- [44] P. Bennis, Evidence for a respiratory chain in the chloroplast, *Proc. Natl. Acad. Sci. U. S. A.* 79 (1982) 4352–4356.
- [45] P. Bennis, Chlororespiration revisited – mitochondrial-plastid interactions in *Chlamydomonas*, *Biochim. Biophys. Acta* 1186 (1994) 59–66.
- [46] G. Peltier, L. Cournac, Chlororespiration, *Annu. Rev. Plant Biol.* 53 (2002) 523–550.
- [47] T. Shikanai, Cyclic electron transport around photosystem I: genetic approaches, *Annu. Rev. Plant Biol.* 58 (2007) 199–217.
- [48] E. Gotoh, Y. Kobayashi, M. Tsuyama, The post-illumination chlorophyll fluorescence transient indicates the RuBP regeneration limitation of photosynthesis in low light in *Arabidopsis*, *FEBS Lett.* 584 (2010) 3061–3064.
- [49] J.J. Van Thor, C.W. Mullineaux, H.C.P. Matthijs, K.J. Hellingwerf, Light harvesting and state transitions in cyanobacteria, *Bot. Acta* 111 (1998) 430–443.
- [50] C.W. Mullineaux, The thylakoid membranes of cyanobacteria: structure, dynamics and function, *Aust. J. Plant Physiol.* 26 (1999) 671–677.
- [51] S. Joshua, C.W. Mullineaux, Phycobilisome diffusion is required for light-state transitions in cyanobacteria, *Plant Physiol.* 135 (2004) 2112–2119.
- [52] L.-N. Liu, T.J. Aartsma, J.-C. Thomas, B.-C. Zhou, Y.-Z. Zhang, FRAP analysis on red algae reveals the fluorescence recovery is ascribed to intrinsic photoprocesses of phycobilisomes than large-scale diffusion, *PLoS One* 4 (2009) e5295.
- [53] A.C. Ley, W.L. Butler, Effects of chromatic adaptation on the photochemical apparatus of photosynthesis in *Porphyridium cruentum*, *Plant Physiol.* 65 (1980) 714–722.
- [54] T. Förster, Experimentelle und theoretische Untersuchungen des zwischenmolekularen Übergangs von Elektronenanregungsenergie, *Z. Naturforsch.* 4 (1949) 321–327.
- [55] L. Stryer, Fluorescence energy transfer as a spectroscopic ruler, *Annu. Rev. Biochem.* 47 (1978) 819–846.
- [56] L.-N. Liu, T.J. Aartsma, J.-C. Thomas, G.E.M. Lamers, B.-C. Zhou, Y.-Z. Zhang, Watching the native supramolecular architecture of photosynthetic membrane in red algae, *J. Biol. Chem.* 283 (2008) 34946–34953.
- [57] C. Bonaventura, J. Myers, Fluorescence and oxygen evolution from *Chlorella pyrenoidosa*, *Biochim. Biophys. Acta* 189 (1969) 366–383.
- [58] F.-A. Wollman, P. Delepelaire, Correlation between changes in light energy distribution and changes in thylakoid membrane polypeptide phosphorylation in *Chlamydomonas reinhardtii*, *J. Cell Biol.* 98 (1984) 1–7.
- [59] F.-A. Wollman, State transitions reveal the dynamics and flexibility of the photosynthetic apparatus, *EMBO J.* 20 (2001) 3623–3630.
- [60] K. Erixon, W.L. Butler, The relationship between Q<sub>A</sub>-C-550 and cytochrome b559 in photoreactions at –196 degrees in chloroplasts, *Biochim. Biophys. Acta* 234 (1971) 381–389.

## 4 Conclusion & Perspectives

Ces travaux de thèse ont permis de poser quelques pierres et ainsi de contribuer à "l'édifice *Chondrus*". Autour d'une thématique d'ordre physiologique, qui portait sur la réponse au stress abiotique de *C. crispus*, j'ai abordé la question de manière pluridisciplinaire et à différentes échelles :

- une échelle globale, la transcriptomique, à l'aide d'outils bioinformatiques,
- une échelle très ciblée, l'étude de la photosynthèse, avec des techniques de biophysique.

Avant d'aborder la notion de stress, il a fallu apprendre à utiliser les techniques sus-citées et établir une ligne de base, c'est à dire caractériser le modèle *C. crispus* dans son état non-stressé. Pour cela, j'ai exploité les données du génome, analysé un premier transcriptome pilote, avant d'entreprendre l'analyse d'un plus grand jeu de données, et enfin, j'ai établi un protocole d'étude de la photosynthèse adapté à *C. crispus*.

Pour en arriver là, le parcours a été jalonné d'obstacles qu'il a fallût contourner. Le premier obstacle a concerné le tout premier objectif de ma thèse, qui était le calibrage des conditions de stress précédant les expériences de transcriptomique. Pour cela j'avais deux possibilités : quantifier la production de ROS résultant d'un stress abiotique et mesurer l'efficacité de la photosynthèse.

**Du DCF et des ROS.** La quantification de la production de ROS est effectuée en mesurant la fluorescence émise par le DCF au contact des ROS. La difficulté était de réussir à isoler le pic d'émission de fluorescence du DCF trop proche de ceux des PBS pour que le spectrofluoromètre utilisé puisse les discriminer. J'y suis parvenu en mettant en place un protocole permettant de précipiter l'ensemble des phycobiliprotéines. Je n'ai hélas pas pu quantifier le signal de manière reproductible, pour des raisons probables de fuite du DCF hors des cellules. Je me suis tournée alors vers la microscopie à fluorescence afin de suivre la production de ROS dans les cellules de *C. crispus in vivo*. Le capteur du microscope était suffisamment sensible pour discriminer l'émission de fluorescence du DCF de celles des PBS. Bien qu'il ne fut toujours pas possible de mesurer cette production de manière quantitative, nous avons suivi la production de ROS de manière qualitative et observé qu'elle était la plus prononcée au niveau des chloroplastes, ce qui a appuyé l'hypothèse que l'appareil photosynthétique était particulièrement sensible aux conditions de stress et était un producteur majeur de ROS.

**Lumière bleue contre algues rouges.** Une deuxième possibilité pour calibrer les conditions de stress était donc de mesurer des paramètres liés à la photosynthèse à l'aide de la fluorescence de la chlorophylle. J'avais commencé ces mesures avec un fluoromètre PAM, qui utilise de la lumière bleue pour exciter la chlorophylle. J'ai très vite observé que cette longueur d'onde n'était pas la plus appropriée pour une algue rouge dont les antennes possèdent des pigments surnuméraires qui n'absorbent pas dans le bleu, mais plutôt dans le vert. J'ai également constaté que je n'arrivais pas à provoquer de phénomènes semblables à ceux observés chez les plantes vertes, notamment le NPQ classique. Cela a été le départ de la collaboration avec l'IBPC qui a conduit à mener des travaux originaux de fluorescence et spectrométrie sur la régulation de l'acquisition de la lumière par *C. crispus* et mettre en lumière le phénomène de *spill-over*. Ce phénomène se traduit par un transfert d'énergie du PSII vers le PSI, par un contact physique entre les deux photosystèmes, permettant d'une part d'éviter des dommages sur le PSII dus à une lumière excédentaire et très probablement aussi de réguler le flux d'électron, entre flux linéaire et cyclique et produire ainsi plus ou moins d'ATP ou de force réductrice.

Il n'a pas encore été possible d'utiliser ces mesures de fluorescence et de spectrométrie pour parvenir au but initial de calibrage des conditions de stress. En revanche, ces travaux auront permis de progresser sur la compréhension d'un phénomène qui au départ était complètement inattendu. Nous avons ainsi caractérisé le *spill-over*, défini les paramètres qui le déclenchent, ses cinétiques d'induction et de relaxation, et établi un modèle hypothétique de fonctionnement.

Appréhender les réponses de l'appareil photosynthétique de *C. crispus* dans des conditions "normales" était un pré-requis avant d'entamer des études où l'on fait intervenir de nouvelles variables telles que des facteurs de stress. Il sera maintenant peut-être possible de soumettre les algues à des conditions contrôlées et identifier ainsi un fonctionnement altéré de la photosynthèse en réponse à ces conditions.

**Génome, bases de données et bioinformatique.** Une partie importante de cette thèse s'inscrit dans une entreprise de grande envergure, puisque j'ai fait partie intégrante du consortium "Génome *Chondrus*" et pu ainsi participer à un travail collaboratif à l'échelle mondiale, en contribuant d'une part à l'annotation du génome, et donc la "construction" de *C. crispus* comme modèle génomique, et d'autre part à utiliser et exploiter les données génomiques dans le cadre d'analyses d'expression à grande échelle.

L'annotation du génome de *C. crispus* aura mené à beaucoup de réflexion, essentiellement sur

la gestion des données massives et les limites de la bioinformatique. Pour mes travaux, la bioinformatique a été un outil qui avait pour objectif d'automatiser le traitement des séquences, afin par exemple d'annoter les gènes et de les classer dans des catégories fonctionnelles. Seulement cet outil, qui peut être très puissant, dépend invariablement de la qualité des informations que l'on lui fournit et des bases de données dans lesquelles sont stockées ces informations. Les premiers problèmes ont été issus de l'annotation manuelle des gènes de *C. crispus*. Un atelier a été consacré à établir des règles pour harmoniser les informations que devaient fournir le consortium d'annotateurs. Chacun avait donc pour objectif de fournir les informations suivantes : code du modèle de gène, nom du gène, code EC (pour les enzymes) et une case "commentaire".

Trois versions des codes de modèles de gènes sont sorties au court des trois années où j'ai travaillé sur le génome et le transcriptome de *C. crispus*. Ces trois versions ne comportaient pas les mêmes nombre de gènes, et certains modèles de gènes avaient été modifiés. L'annotation manuelle a été initiée pendant la version 1, et s'est poursuivie pendant la version 2. Cela a généré une "désynchronisation" des données, puisque deux types de codes étaient possibles. La version 3 du génome a été générée lorsque l'annotation manuelle a été fusionnée avec l'annotation automatique, et suite aux différences entre la version 1 et la version 2 de nombreuses informations ont été perdues.

Lorsque j'ai eu à manipuler ces données dans la suite de mon travail, j'ai pu constater que, pour certains gènes que j'avais annotés, les informations n'étaient pas présentes dans la base de données finale et c'est probablement le cas pour d'autres catégories de gènes. Un travail d'inventaire sera nécessaire pour compléter ces données et valoriser au mieux l'annotation manuelle.

Lors de mes analyses de transcriptomique j'ai tenté de classer les gènes dans des groupes fonctionnels grâce à un outil qui pourrait être très puissant, l'annotation GO (gene ontology). Cependant, lors de l'annotation manuelle du génome, nous n'avions pas prévu de case "GO" et cette information manque donc pour tous les gènes annotés manuellement. Le seul moyen de parvenir à assigner des termes GO a été alors de refaire une annotation automatique, ce qui prend beaucoup de temps et génère des inexactitudes.

L'annotation d'un génome est un travail extrêmement chronophage, qui nécessite des moyens humains importants, et est essentielle pour pouvoir travailler de manière efficace sur les données génomiques. Cela demande une communication réactive entre les nombreuses personnes impliquées dans le projet et une anticipation des problèmes liés particulièrement à la gestion de la bases de données. Selon moi, le point de départ d'une étude "-omic" est la réflexion autour de

la conception de la base de données, aussi bien en contenu qu'en visualisation. Un recensement des besoins des utilisateurs amenés à exploiter le génome permettrait de définir les informations utiles et indispensables à importer dans la base de données. Ceci permettrait d'optimiser le travail d'annotation qui suivrait.

**Le transcriptome et son potentiel.** Lorsque j'ai effectué les analyses transcriptomiques, une part importante du temps a été consacrée au travail de classement (et de ré-annotation) des gènes qui, à défaut de pouvoir être automatisé, a donc été fait à la main. C'était un travail indispensable avant de pouvoir chercher un sens biologique dans les données. L'exploitation d'un transcriptome est source d'une infinité d'informations et la progression se fait petits pas par petits pas.

Il est possible de faire des expériences de RNAseq sur *C. crispus* et d'en tirer des informations exploitables. Jusqu'à présent, ces analyses nous auront permis de souligner que les changements qui se produisent durant les différents moments de la journée sont particulièrement importants. Nous avons vu que les HSP semblent avoir un rôle majeur au cours du cycle diurnal, davantage que les enzymes antioxydantes. La réalisation de ces expériences sur le terrain aura donné des informations différentes de celles réalisées précédemment en conditions de laboratoire.

De plus, cela aura permis de mettre en place des outils de biologie moléculaire, avec l'identification de trois gènes dont l'expression est spécifique des tetrasporophytes, permettant d'appuyer l'identification basée sur des critères morphologiques. Ces travaux auront également été le point de départ pour l'identification des gènes de normalisation en RT-qPCR. Cette technique sera en effet amenée à être utilisée de manière ciblée sur les gènes candidats issus de l'étude globale.

Ces expériences auront aussi été une opportunité de découvrir et d'appréhender des données haut-débit et identifier certaines difficultés qui y sont associées. Tirer du sens biologique n'est pas évident lorsque l'on est confronté à autant de données. Il serait intéressant de replacer les gènes dont l'expression est différentielle entre les conditions, au sein des familles auxquelles ils appartiennent et de les grouper par fonctions (j'en reviens aux termes GO). J'ai pu effectuer une part de ce travail avec les gènes relatifs au stress que j'ai annotés, j'avais ainsi une liste la plus exhaustive possible et j'ai pu ainsi estimer la part de gènes de "stress" qui avait une importance significative dans l'expérience, par rapport aux gènes de stress totaux. Pour compléter cette étude, il sera nécessaire d'estimer les rapports de gènes différentiellement exprimés sur l'ensemble des gènes de chaque catégorie fonctionnelle ou famille.



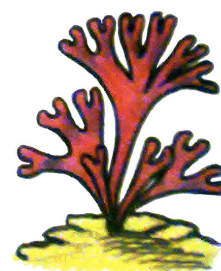
**Pour aller plus loin...** À présent, il reste près de deux-tiers de données qui n'ont pas encore été exploitées. Elles peuvent contenir des informations sur l'écosystème de *C. crispus*, mais également servir à améliorer davantage l'annotation du génome et préciser les modèles de gènes.

Dans l'immédiat, ces données pourront être utilisées dans le cadre de l'étude des voies de biosynthèses des carraghénanes, une thématique majeure dans le laboratoire. Il serait intéressant de surexprimer les sulfurylases qui semblent spécifiques des tetrasporophytes afin de caractériser leur fonction et vérifier si elles participent réellement à la conversion du  $\lambda$ -carraghénane en  $\theta$ -carraghénane. Ceci permettrait également de confirmer l'existence de ce dernier et d'analyser ces propriétés rhéologiques.

L'analyse des séquences brutes des expériences de transcriptomique est potentiellement riche en informations, comme cela avait déjà été le cas pour reconstruire le génome plastidial. À partir de l'expérience "stress", il sera notamment possible de comparer les séquences de quelques gènes choisis et d'évaluer s'il existe des différences au sein des séquences d'ARN entre les différents échantillons, pouvant générer d'éventuels biais dans l'alignement des lectures sur le génome de référence. Sous un aspect plus écologique, il serait également intéressant d'estimer les variations naturelles au sein des transcrits dans des échantillons environnementaux.

Pour des perspectives à plus long terme, une expérience de transcriptomique (avec réplicats biologiques) sur des tissus enrichis en structures reproductives, pourrait révéler de nouveaux gènes différentiellement exprimés entre les stades de vie, incluant les carposporophytes.

Enfin, il reste une myriade de gènes aux fonctions inconnues qui contiennent sans doute des informations inédites pour comprendre le fonctionnement des algues rouges.



## Références

### Introduction

- [1] M Ashraf (2010) Inducing drought tolerance in plants : recent advances. *Biotechnol Adv.* 28 :169-83.
- [2] SM Adl *et al.* (2012) The Revised Classification of Eukaryotes. *J. Eukaryot. Microbiol.* 59 :429-493.
- [3] PJ Keeling (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci* 365 :729-48.
- [4] SB Hedges, JE Blair, ML Venturi, JL Shoe (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* 4 :2.
- [5] HS Yoon, JD Hackett, C Ciniglia, G Pinto, D Bhattacharya (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* 21 : 809-818.
- [6] NJ Butterfield (2000) *Bangiomorpha pubescens* n. gen., n. sp.; implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26 :386-404.
- [7] MJ Harvey, J McLachlan (1973) *Chondrus crispus*. *Proc. Trans. Nova Scotian Inst. Sci.* 27 :1-155.
- [8] S Genicot-Joncour, A Poinas, O Richard, P Potin, B Rudolph, B Kloareg, W Helbert (2009) The Cyclization of the 3,6-Anhydro-Galactose Ring of i-Carrageenan Is Catalyzed by Two D-Galactose-2,6-Sulfurylases in the Red Alga *Chondrus crispus*. *Plant Physiology* 151 :1609-1616,
- [9] P Nyvall Collén, M Lemoine, R Daniellou, J-P Guégan, S Paoletti, W Helbert (2009) Enzymatic degradation of  $\kappa$ -carrageenan in aqueous solution. *Biomacromolecules* 10 :1757-1767.
- [10] C Leblanc *et al.* (1995) Complete sequence of the mitochondrial DNA of the rhodophyte *Chondrus crispus* (Gigartinales). Gene content and genome organization. *Journal of molecular biology* 250 484-95.
- [11] WM Woelkerling (1990) An introduction. *Biology of the Red Algae* Chapter 1, pp. 1-6.
- [12] K Bouarab, P Potin, F Weinberger, J Correa, B Kloareg (2001) The *Chondrus crispus*-*Acrochaete operculata* host-pathogen association, a novel model in glycobiology and applied phycopathology. *Journal of Applied Phycology* 13 :185-193.
- [13] C Hervé, T Tonon, J Collén, E Corre, C Boyen (2006) NADPH oxidases in Eukaryotes : red algae provide new hints! *Current Genetics* 49 :190-204.
- [14] J Collén, C Hervé, I Guisle-Marsollier, JJ Léger, C Boyen (2006) Expression profiling of *Chondrus crispus* (Rhodophyta) after exposure to methyl jasmonate. *J Exp Bot.* 57 :3869-81.
- [15] E Gaquerel, C Hervé, C Labrière, C Boyen, P Potin, JP Salaün (2007) Evidence for oxylipin synthesis and induction of a new polyunsaturated fatty acid hydroxylase activity in *Chondrus crispus* in response to methyljasmonate. *Biochim. Biophys. Acta* 1771 :565-75.
- [16] J Collén, I Guisle-Marsollier, JJ Léger, C Boyen (2007) Response of the transcriptome of the intertidal red seaweed *Chondrus crispus* to controlled and natural stresses. *New Phytol.* 176 :45-55.
- [17] J Collen, V Roeder, S Rousvoal *et al.* (2006) An expressed sequence tag analysis of thallus and regenerating protoplasts of *Chondrus crispus* (Gigartinales, Rhodophyceae) *J Phycol* 42 :104-112.

- [18] H Verbruggen, CA Maggs, GW Saunders, L Le Gall, HS Yoon, O De Clerck (2010) Data mining approach identifies research priorities and data requirements for resolving the red algal tree of life. *BMC Evolutionary Biology* 10 :16.
- [19] C Rhee, WR Briggs (1977) Some Responses of *Chondrus crispus* to Light. I. Pigmentation Changes in the Natural Habitat. *Botanical Gazette* 138 :123-128.
- [20] LC-M Chen, J McLachlan (1972) The life history of *Chondrus crispus* in culture. *Can. J. Bot.* 50, 1055-1060.
- [21] MW Hawkes (1990) Reproductive strategies. *Biology of the Red Algae* Chapter 17, pp. 455-476.
- [22] I Fournet, E Deslandes, J-Y Floc'h (1993) Iridescence : a useful criterion to sort gametophyte from sporophytes in the red alga *Chondrus crispus*. *Journal of Applied Phycology* 5 :535-537.
- [23] EM Gordon-Mills, EL McCandless (1975) Carrageenans in the cell walls of *Chondrus crispus* Stack. (Rhodophyceae, Gigartinales). I. Localization with fluorescent antibody. *Phycologia* 14 :275-281.
- [24] D Rees, E Morris, D Thom, J Madden (1982) Shape and interaction of carbohydrate chains. In *The Polysaccharides (Aspinall)* pp. 195-290.
- [25] H Bixler (1996) Recent development in manufacturing and marketing carrageenan. *Hydrobiologia* 326/327 :35-37.
- [26] E Morris, D Rees, G Robinson (1980) Cation-specific aggregation of carrageenan helices : domain model of polymer gel structure. *J. Mol. Biol.* 138 :349-362.
- [27] B Kloareg, RS Quatrano (1988) Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanography and marine biology* 26 :259-315.
- [28] EI Ask, A Batibasaga, JA Zertuche-Gonzalés, M de San (2001) Three decades of *Kappaphycus alvarezii* (Rhodophyta) introduction to non-endemic locations. In : Chapman ARO, Anderson RJ, Vreeland VJ, Davison IR (eds) 17th Int Seaweed Symposium. Oxford University Press, Oxford, pp. 49-57.
- [29] T Chopin, G Sharp, E Belyea, R Semple, D Jones (1999) Open-water aquaculture of the red alga *Chondrus crispus* in Prince Edward Island, Canada. *Hydrobiologia* 398-399 :417-425.
- [30] AC Neish, PF Shacklock, CH Fox, FJ Simpson (1977) The cultivation of *Chondrus crispus*. Factors affecting growth under greenhouse conditions. *Canadian Journal of Botany* 55 : 2263-2271.
- [31] MG Tasende, MI Fraga (1999) Growth of *Chondrus crispus* Stackhouse (Rhodophyta, Gigartinaceae) in laboratory culture. *Ophelia* 51 :202-213.
- [32] RL Burns, AC Mathieson (1972) Ecological studies of economic red algae. III. Growth and reproduction of natural and harvested populations of *Gigartina stellata* (Stackhouse) batters in New Hampshire. *Journal of Experimental Marine Biology and Ecology* 9 :77-95.
- [33] J Farineau, J-F Morot-Gaudry (2011) La photosynthèse : processus physiques, moléculaires et physiologiques. Collection Synthèse, INRA, Paris.
- [34] T Tonon, D Eveillard, S Prigent, J Bourdon, P Potin, C Boyen, A Siegel (2011) Toward systems biology in brown algae to explore acclimation and adaptation to the shore environment. *OMICS* 15 :883-92.
- [35] JM Cock et al. (2010) The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* 465 :617-621.

- [36] JD Rochaix (1995) *Chlamydomonas reinhardtii* as the photosynthetic yeast. *Annu Rev Genet.* 29 :209-230.
- [37] SS Merchant *et al.* (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318 :245-50.
- [38] E Derelle, C Ferraz, S Rombauts, P Rouzé, AZ Worden, S Robbens, F Partensky, S Degroove, S Echeynié, R Cooke, Y Saeys, J Wuyts, K Jabbari, CBowler, OPanaud, BPiégu, SG Ball, JP Ral, FY Bouget, G Piganeau, B De Baets, A Picard, M Delseny, J Demaille, Y Van de Peer, H Moreau (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci U S A* 103 :11647-11652.
- [39] M Matsuzaki, O Misumi, T Shin-I, S Maruyama, M Takahara, SY Miyagishima, T Mori, K Nishida, F Yagisawa, K Nishida, Y Yoshida, Y Nishimura, S Nakao, T Kobayashi, Y Momoyama, T Higashiyama, A Minoda, M Sano, H Nomoto, K Oishi, H Hayashi, F Ohta, S Nishizaka, S Haga, S Miura, T Morishita, Y Kabeya, K Terasawa, Y Suzuki, Y Ishii, S Asakawa, H Takano, N Ohta, H Kuroiwa, K Tanaka, N Shimizu, S Sugano, N Sato, H Nozaki, N Ogasawara, Y Kohara, T Kuroiwa (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428 :653-657.
- [40] DC Price, *et al.* *Cyanophora paradoxa* Genome Elucidates Origin of Photosynthesis in Algae and Plants. *Science* 335 :843-847.
- [41] NA Blouin, JA Brodie, AC Grossman, P Xu, SH Brawley (2011) *Porphyra* : a marine crop shaped by stress. *Trends Plant Sci.* 16 :29-37.
- [42] J Collén, B Porcel, W Carré, C Chaparro, T Tonon, T Barbeyron, G Michel, B Noel, K Valentin, M Elias, N Kowalczyk, *et al.*, C Boyen. "Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida". Submitted to PNAS.
- [43] G Blanca, G Duncan, I Agarkova, M Borodovsky, J Gurnon, A Kuo, E Lindquist, S Lucas, J Pangilinan, J Polle, A Salamov, A Terry, T Yamada, DD Dunigan, IV Grigoriev, J-M Claverie, JL Van Ettenc (2010) The *Chlorella variabilis* NC64A Genome Reveals Adaptation to Photosymbiosis, Coevolution with Viruses, and Cryptic Sex. *Plant Cell* 22 :2943-2955.
- [44] F Sanger, AR Coulson (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94(3) :441-448.
- [45] AM Maxam, W Gilbert (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74(1977) :560-564.
- [46] M Ronaghi, S Karamohamed, B Pettersson, M Uhlén, P Nyrén (1996) Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry* 242 :84-89.
- [47] JM Rothberg, JH Leamon (2008) The development and impact of 454 sequencing. *Nat Biotechnol.* 26 :1117-24.
- [48] M Džunková, G D'Auria, D Pérez-Villarroya, A Moya (2012) Hybrid sequencing approach applied to human fecal metagenomic clone libraries revealed clones with potential biotechnological applications. *PLoS One* 7 :e47654.
- [49] S Diguistini, NY Liao, D Platt, G Robertson, M Seidel, SK Chan, TR Docking, I Birol, RA Holt, M Hirst, E Mardis, MA Marra, RC Hamelin, J Bohlmann, C Breuil, SJ Jones (2009) De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biol* 10 :R94.
- [50] JH Malone, B Oliver (2011) Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* 9 :34.

- [51] P Mitchell, D Tollervey (2001) mRNA turnover. *Curr Opin Cell Biol.* 13 :320-325.
- [52] A Moya, L Huisman, EE Ball, et al. (2012) Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO<sub>2</sub>-driven acidification during the initiation of calcification. *Mol Ecol*
- [53] S Zenoni, A Ferrarini, E Giacomelli, L Xumerle, M Fasoli, G Malerba, D Bellin, M Pezzotti, M Delledonne (2010) Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq. *Plant Physiol.* 152 :1787-1795.
- [54] LC Carvalhais, PG Dennis, GW Tyson, PM Schenk (2012) Application of metatranscriptomics to soil environments. *J Microbiol Methods* 91 :246-251
- [55] J Kennedy, B Flemer, SA Jackson, DP Lejon, JP Morrissey, F O’Gara, AD Dobson (2010) Marine metagenomics : new tools for the study and exploitation of marine microbial metabolism. *Mar Drugs* 8 :608-28.
- [56] J Handelsman (2004) Metagenomics : application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 68 :669-685.
- [57] SP Place, BA Menge, GE Hofmann (2012) Transcriptome profiles link environmental variation and physiological response of *Mytilus californianus* between Pacific tides. *Functional Ecology* 26 :144-155.
- [58] TG Evans, E Hammill, K Kaukinen, AD Schulze, DA Patterson, KK English, JMR Curtis, KM Miller (2011) Transcriptomics of environmental acclimatization and survival in wild adult Pacific sockeye salmon (*Oncorhynchus nerka*) during spawning migration. *Molecular Ecology* 20 :4472-4489.
- [59] IR Davison, GA Pearson (1996) Stress tolerance in intertidal seaweeds. *J. Phycol.* 32 :197-211.
- [60] J Collén, IR Davison (1999) Reactive oxygen metabolism in intertidal *Fucus* Spp. (Phaeophyceae). *J. Phycol.* 35 :62-69.
- [61] T Hirayama, K Shinozaki (2010) Research on plant abiotic stress responses in the post-genome era : past, present and future. *Plant J.* 61 :1041-52.
- [62] A Joshi, HQ Dang, N Vaid, N Tuteja (2009) Isolation of high salinity stress tolerant genes from *Pisum sativum* by random overexpression in *Escherichia coli* and their functional validation. *Plant Signal Behav.* 4 :400-12.
- [63] CC Chen, CS Liang, AL Kao, CC Yang (2010) HHP1, a novel signalling component in the cross-talk between the cold and osmotic signalling pathways in *Arabidopsis*. *J Exp Bot.* 61 :3305-20.
- [64] MC Morais, H Freitas (2012) The acclimation potential of *Acacia longifolia* to water stress : Implications for invasiveness. *Plant Sci.* 196 :77-84.
- [65] B Vinocur, A Altman (2005) Recent advances in engineering plant tolerance to abiotic stress : achievements and limitations. *Current Opinion in Biotechnology* 16 :123-132.
- [66] GT Huang, SL Ma, LP Bai, L Zhang, H Ma, P Jia, J Liu, M Zhong, ZF Guo (2012) Signal transduction during cold, salt, and drought stresses in plants. *Mol Biol Rep.* 39 :969-987.
- [67] A Conde, MM Chaves, H Gerós (2011) Membrane Transport, Sensing and Signaling in Plant Adaptation to Environmental Stress. *Plant Cell Physiol* 52 :1583-1602.
- [68] S Mahajan, N Tuteja (2005) Cold, salinity and drought stresses : an overview. *Arch Biochem Biophys.* 444 :139-58.

- [69] R Munns, M Tester (2008) Mechanisms of salinity tolerance. *Annu Rev Plant Biol.* 59 :651-81.
- [70] Flowers, T. J. and Colmer, T. D. (2008), Salinity tolerance in halophytes. *New Phytologist*, 179 : 945-963.
- [71] M Ashraf, MR Foolad (2007) Roles of glycine betaine and proline in improving plant abiotic stress resistance. *Environmental and Experimental Botany* 59 :206-216.
- [72] W Wang, B Vinocur, O Shoseyov, A Altman (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9 :244-52.
- [73] N Valdivia, RA Scrosati, M Molis, AS Knox (2011) Variation in Community Structure across Vertical Intertidal Stress Gradients : How Does It Compare with Horizontal Variation at Different Scales ? *PLoS ONE* 6 :e24062.
- [74] CS Lobban, PJ Harrison (1994) Seaweed ecology and physiology. Cambridge university press.
- [75] K Hoëfler (1930) Das Plasmolyse - Verhalten der Rotalgen. *Z. Bot* 23 :570-588.
- [76] RH Reed (1990) Solute accumulation and osmotic adjustment. *Biology of the Red Algae.* Cambridge University Press pp. 147-165.
- [77] RH Reed, IR Davison, JA Chudek, R Foster (1985) The osmotic role of mannitol in the Phaeophyta - an appraisal. *Phycologia* 24 :35-47.
- [78] A Eggert, S Raimund , D Michalik, J West, U Karsten (2007) Ecophysiological performance of the primitive red alga *Dixoniella grisea* (Rhodellophyceae) to irradiance, temperature and salinity stress : growth responses and the osmotic role of mannitol. *Phycologia* 46 :22-28.
- [79] M Kakinuma, DA Coury, Y Kuno, S Itoh, Y Kozawa, E Inagaki, Y Yoshiura, H Amano (2006) Physiological and biochemical responses to thermal and salinity stresses in a sterile mutant of *Ulva pertusa* (Ulvales, Chlorophyta). *Marine Biology* 149 :97-106.
- [80] E Parre, MA Ghars, AS Leprince, L Thiery, D Lefebvre, M Bordenave, L Richard, C Mazars, C Abdelly, A Savouré (2007) Calcium signaling via phospholipase C is essential for proline accumulation upon ionic but not nonionic hyperosmotic stresses in *Arabidopsis*. *Plant Physiol.* 144 :503-512.
- [81] HB Shao, LY Chu, CA Jaleel, CX Zhao (2008) Water-deficit stress-induced anatomical changes in higher plants. *C R Biol.* 331 :215-25.
- [82] KHM Cardozo, T Guaratini, MP Barros, VR Falcão, AP Tonon, NP Lopes, S Campos, MA Torres, AO Souza, P Colepicolo, E Pinto (2007) Metabolites from algae with economical impact. *Comparative Biochemistry and Physiology Part C : Toxicology & Pharmacology* 146 :60-78.
- [83] MY Roleda, C Wiencke, D Hanelt, K Bischof (2007) Sensitivity of the Early Life Stages of Macroalgae from the Northern Hemisphere to Ultraviolet Radiation. *Photochemistry and Photobiology* 83 :851-862.
- [84] K Lao, AN Glazer (1996) Ultraviolet-B photodestruction of a light-harvesting complex. *Proc. Natl Acad. Sci. U S A* 93 :5258-5263.
- [85] AA Solov'ev, ZK Makhneva, YU Erokhin (2001) UV-Induced destruction of light-harvesting complexes from purple bacterium *Chromatium minutissimum*. *Membr Cell Biol.* 14 :463-74.
- [86] M Tedetti, R Sempéré (2006) Penetration of Ultraviolet Radiation in the Marine Environment. A Review. *Photochemistry and Photobiology* 82 : 389-397.

- [87] T Soule, K Palmer, Q Gao, RM Potrafka, V Stout, F Garcia-Pichel (2009) A comparative genomics approach to understanding the biosynthesis of the sunscreen scytonemin in cyanobacteria. *BMC Genomics* 10 :336.
- [88] G Kräbs, K Bischof, D Hanelt, U Karsten, C Wiencke (2002) Wavelength-dependent induction of UV-absorbing mycosporine-like amino acids in the red alga *Chondrus crispus* under natural solar radiation. *Journal of Experimental Marine Biology and Ecology* 268 :69-82.
- [89] HA Kratsch, RR Wise (2000) The ultrastructure of chilling stress. *Plant Cell & Environment* 23 :337-350.
- [90] P Mazur (1984) Freezing of living cells : mechanisms and implications. *Am J Physiol Cell Physiol* 247 :C125-C142.
- [91] PJ Lillford, CB Holt (2002) In vitro uses of biological cryoprotectants. *Philos Trans. R. Soc. Lond. B* 357 : 945-951.
- [92] C Guy, F Kaplan, J Kopka, J Selbig, DK Hinch (2008) Metabolomics of temperature stress. *Physiologia Plantarum* 132 :220-235.
- [93] NL Lohrmann , BA Logan, AS Johnson (2004) Seasonal acclimatization of antioxidants and photosynthesis in *Chondrus crispus* and *Mastocarpus stellatus*, two co-occurring red algae with differing stress tolerances. *Biol Bull.* 207 :225-32.
- [94] S Nath, B Deb (2011) Abiotic stress tolerance mechanism of plants by heat shock proteins : A review. *An Annual Journal of Dept. of Biotechnology* 3 : 45-49.
- [95] PG Falkowski, JA Raven (1997) Aquatic photosynthesis. Blackwell Science.
- [96] F Brechignac, M Andre (1985) Continuous Measurements of the Free Dissolved CO<sub>2</sub> Concentration during Photosynthesis of Marine Plants : Evidence for HCO<sub>3</sub><sup>-</sup> Use in *Chondrus crispus*. *Plant Physiol.* 78 :551-554.
- [97] F Brechignac, M Andre, A Gerbaud (1986) Preferential Photosynthetic Uptake of Exogenous HCO<sub>3</sub><sup>-</sup> in the Marine Macroalga *Chondrus crispus*. *Plant Physiol.* 80 :1059-1062.
- [98] N Murata, S Takahashi, Y Nishiyama, SI Allakhverdiev (2007) Photoinhibition of photosystem II under environmental stress. *Biochim Biophys Acta* 1767 :414-21.
- [99] J Collén, MJ Rio, G García-Reina, M Pedersén (1995) Photosynthetic production of hydrogen peroxide by *Ulva rigida* C. Ag.(Chlorophyta). *Planta* 31 :265-271.
- [100] J Collén, IR Davison (1999) Stress tolerance and reactive oxygen metabolism in the intertidal red seaweeds *Mastocarpus stellatus* and *Chondrus crispus*. *Plant, Cell & Environment* 22 :1143-1151.
- [101] G Miller, V Shulaev, R Mittler (2008) Reactive oxygen signaling and abiotic stress. *Physiologia Plantarum* 133 :481-489.
- [102] PS Brooks (2005) Mitochondrial H<sup>+</sup> leak and ROS generation : an odd couple. *Free Radical Biology and Medicine* 38 :12-23.
- [103] A Edreva (2005) Generation and scavenging of reactive oxygen species in chloroplasts : a submolecular approach. *Agric. Ecosyst. Environ.* 106 :119-123.
- [104] IM Møller, PE Jensen, A Hansson (2007) Oxidative Modifications to Cellular Components in Plants. *Annu. Rev. Plant Biol.* 58 :459-481.
- [105] K Apel, H Hirt (2004) Reactive oxygen species : metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* 55 :373-99.

- [106] MA Torres, JDG Jones, JL Dangel (2005) Pathogen-induced, NADPH oxidase-derived reactive oxygen intermediates suppress spread of cell death in *Arabidopsis thaliana*. *Nature Genetics* 37 :1130-1134.
- [107] B Halliwell, JMC Gutteridge (1990) Role of free-radicals and catalytic metal-ions in human disease - an overview. *Methods in Enzymology* 186 :1-85.
- [108] K Asada (2000) The water-water cycle as alternative photon and electron sinks. *Philos Trans R Soc Lond B Biol Sci* 355 :1419-1431.
- [109] VA Gerard (1997) Environmental stress during early development of kelp sporophytes (*Laminaria saccharina*) : how long do effects persist? *Journal of Applied Phycology* 9 :5-9.
- [110] CT Shiu, TM Lee (2005) Ultraviolet-B-induced oxidative stress and responses of the ascorbate-glutathione cycle in a marine macroalga *Ulva fasciata*. *J Exp Bot.* 56 :2851-65.
- [111] L Li, J Zhao, X Tang (2010) Ultraviolet irradiation induced oxidative stress and response of antioxidant system in an intertidal macroalgae *Corallina officinalis* L. *J Environ Sci (China)* 22 :716-22.
- [112] A Cabello-Pasini, E Aguirre-von-Wobeser, FL Figueroa (2000) Photoinhibition of photosynthesis in *Macrocystis pyrifera* (Phaeophyceae), *Chondrus crispus* (Rhodophyceae) and *Ulva lactuca* (Chlorophyceae) in outdoor culture systems. *J Photochem Photobiol B.* 57 :169-78.
- [113] SR Dudgeon, IR Davison, RL Vadas (1990) Freezing tolerance in the intertidal red algae *Chondrus crispus* and *Mastocarpus stellatus* : relative importance of acclimation and adaptation. *Marine Biology* 101 :107-114.
- [114] SR Dudgeon, IR Davison, RL Vadas (1989) Effect of freezing on photosynthesis of intertidal macroalgae : relative tolerance of *Chondrus crispus* and *Mastocarpus stellatus* (Rhodophyta). *Marine Biology* 106 :427-436.
- [115] M Calvin, A Benson (1948) The Path of Carbon in Photosynthesis *Science* 107 :476-480.
- [116] S Eberhard, G Finazzi, F-A Wollman (2008) The dynamics of photosynthesis. *Annu. Rev. Genet.* 42 :463-515.
- [117] GG Hammes (1982) Unifying concept for the coupling between ion pumping and ATP hydrolysis or synthesis. *Proc Natl Acad Sci U S A* 79 :6881-8884.
- [118] D Bhaya, AR Grossman (1993) Characterization of gene clusters encoding the fucoxanthin chlorophyll proteins of the diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res* 21 :4458-66.
- [119] A De Martino, D Douady, M Quinet-Szely, B Rousseau, F Crépineau, K Apt, L Caron (2000) The light-harvesting antenna of brown algae. *Eur. J. Biochem.* 267 :5540-5549.
- [120] GR Wolfe, FX Cunningham, D Durnfordt, BR Green and E Gantt (1994) Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. *Nature* 367 :566-568.
- [121] E Gantt (1975) Phycobilisomes : light-harvesting pigment complexes. *Bioscience* 25 :781-788.
- [122] AN Glazer (1985) Light harvesting by phycobilisomes. *Annu. Rev. Bioph.* 14 :47-77.
- [123] WA Sidler (1994) Phycobilisome and phycobiliprotein structure. In Bryant DA (ed) *The Molecular Biology of Cyanobacteria*. Kluwer Academic Publishers, Netherlands, 139-216.
- [124] LN Liu, TJ Aartsma, JC Thomas, GE Lamers, BC Zhou, YZ Zhang (2008) Watching the native supramolecular architecture of photosynthetic membrane in red algae : topography of phycobilisomes and their crowding, diverse distribution patterns. *J Biol Chem.* 283 :34946-34953.



- [125] RE Blankenship, RC Prince (1985) Excited states redox potentials and the Z scheme of photosynthesis. *Trens Biochem. Sci.* 10 :382-383.
- [126] P Joliot, A Joliot (2002) Cyclic electron transfer in plant leaf. *Proc Natl Acad Sci U S A* 99 :10209-10214.
- [127] B Demmig-Adams, WW Adams (1996) The role of xanthophyll cycle carotenoids in the protection of photosynthesis. *Trends Plant Sci.* 1 :21-26.
- [128] KK Niyogi, X-P Li, V Rosenberg, H-S Jung (2005) Is PsbS the site of non-photochemical quenching in photosynthesis? *J. Exp. Bot.* 56 :375-382.
- [129] E García-Mendoza, MF Colombo-Pallotta (2007) The giant kelp *Macrocystis pyrifera* presents a different nonphotochemical quenching control than higher plants. *New Phytol* 173 :526-36.
- [130] B Bailleul, A Rogato, A de Martino, S Coesel, P Cardol, C Bowler, A Falciatore, G Finazzi (2010) An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc Natl Acad Sci U S A* 107 :18214-9.
- [131] G Finazzi, F Rappaport, A Furia, M Fleischmann, J-D Rochaix, F Zito, and G Forti (2002) Involvement of state transitions in the switch between linear and cyclic electron flow in *Chlamydomonas reinhardtii*. *EMBO Rep* 3 :280-285.
- [132] AN Melkozernov, RE Blankenship (2005) Structural and functional organization of the peripheral light-harvesting system in photosystem I. *Photosynth Res.* 85 :33-50.
- [133] A Busch, M Hippler (2011) The structure and function of eukaryotic photosystem I. *Biochim. Biophys. Acta* 1807 :864-877.
- [134] F-A Wollman (2001) State transitions reveal the dynamics and flexibility of the photosynthetic apparatus. *EMBO J.* 20 :3623-3630.
- [135] R Delosme, J Olive, F-A Wollman (1996) Changes in light energy distribution upon state transitions : an in vivo photoacoustic study of the wild type and photosynthesis mutants from *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta* 1273 :150-158.
- [136] D Kirilovsky (2010) The photoactive orange carotenoid protein and photoprotection in cyanobacteria. *Advances in Experimental Medicine and Biology* 675 :139-159.
- [137] S Takaichi (2011) Carotenoids in Algae : Distributions, Biosyntheses and Functions. *Mar. Drugs* 9 :1101-1118.
- [138] E Delphin, J-C Duval, A-L Etienne, D Kirilovsky (1996) State transitions or  $\Delta$ pH-dependant quenching of photosystem II fluorescence in red algae. *Biochem.* 35 :9435-9445.
- [139] E Delphin, J-C Duval, A-L Etienne, D Kirilovsky (1998) DeltapH-dependent photosystem II fluorescence quenching induced by saturating, multiturnover pulses in red algae. *Plant Physiol.* 118 :103-113.
- [140] E Delphin, J-C Duval, D Kirilovsky (1995) Comparison of state 1-state 2 transitions in the green alga *Chlamydomonas reinhardtii* and in the red alga *Rhodella violacea* : effect of kinase and phosphatase inhibitors. *Biochim. Biophys. Acta* 1232 :91-95.
- [141] R Goss, T Jakob (2010) Regulation and function of xanthophyll cycle-dependent photoprotection in algae. *Photosynthesis Research* 106 :103-122.
- [142] H-W Trissl (2003) Modeling the excitation energy capture in thylakoid membranes. *Advances in photosynthesis and respiration* 14 :245-276.
- [143] K Maxwell, GN Johnson (2000) Chlorophyll fluorescence? A practical guide. *J. Exp. Bot.* 51 :659-668.

- [144] NR Baker (2008) Chlorophyll fluorescence : a probe of photosynthesis in vivo. *Annu Rev Plant Biol.* 59 :89-113.
- [145] SM Ditami, D Scornet, J-L Petit, B Séguens, C Da Silva, E Corre, M Dontrup, K-H Glatting, R König, L Sterck, P Rouzé, Y Van de Peer, JM Cock, C Boyen, T Tonon (2009) Global expression analysis of the brown alga *Ectocarpus siliculosus* (Phaeophyceae) reveals large-scale reprogramming of the transcriptome in response to abiotic stress. *Genome Biology* 10 :66.

## Matériels & Méthodes

- [146] KE Apt, SK Clendennen, DA Powers, AR Grossman (1995) The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. *Mol. Gen. Genet.* 246 :455-464.
- [147] Y Kong (2011) Btrim : a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* 98 :152-153.
- [148] M Martin (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBNet.journal* 17 :10-12.

## Résultats & Discussion

- [149] SE Lewis, SMJ Searle, N Harris, M Gibson, V Iyer, J Richter, C Wiel, L Bayraktaroglu, E Birney, MA Crosby, S Kaminker, B Matthews, SE Prochnik, CD Smith, JL Tupy, GM Rubin, S Misra, CJ Mungall, ME Clamp (2002) Apollo : a sequence annotation editor. *Genome Biology* 3 :research0082.
- [150] Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408 :796-815.
- [151] <http://blast.ncbi.nlm.nih.gov/>
- [152] MH Al-Whaibi (2010) Plant heat shock proteins : a mini review. *KSU Journal (Science)* in press.
- [153] A de Marco (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. *Nature Protocols* 2 :2632-2639.
- [154] MB Yaffe, GW Farr, D Miklos, AL Horwich, ML Sternlicht and H Sternlicht (1992) TCP1 complex is a molecular chaperone in tubulin biogenesis. *Nature* 358 :245-248.
- [155] VA Lewis, GM Hynes, D Zheng, H Saibil and K Willison (1992) The T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol. *Nature* 358 :249-252.
- [156] ER Waters, I Rioflorido (2007) Evolutionary analysis of the small heat shock proteins in five complete algal genomes. *J Mol Evol.* 65 :162-74.
- [157] X Zhang, R Yang, Z Cheng (2008) Cloning and sequence analysis of Mn-superoxide dismutase gene from *Porphyra haitanensis*. EMBL/GenBank/DDBJ databases.
- [158] [www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/)
- [159] P Algarra, JC Thomas, A Mousseau (1990) Phycobilisome heterogeneity in the red alga *Porphyra umbilicalis*. *Plant Physiology* 92 :570-576.
- [160] D Redecker, W Wehrmeyer, W Reuter (1993) Core substructure of the hemiellipsoidal phycobilisome from the red alga *Porphyridium cruentum*. *Eur J Cell Biol* 62 :442-450.

- [161] C Theiss, F-J Schmitt, J Pieper, C Nganou, M Grehn, M Vitali, R Olliges, HJ Eichler, H-J Eckert (2011) Excitation energy transfer in intact cells and in the phycobiliprotein antennae of the chlorophyll d containing cyanobacterium *Acaryochloris marina*. *J Plant Physiol.* 168 :1473-1487.
- [162] D Bhaya, A Dufresne, D Vaultot, AR Grossman (2002) The *hli* gene family in marine and freshwater cyanobacteria. *FEMS Microbiol Lett* 215 : 209-219.
- [163] P Storm, MA Hernandez-Prieto, LL Eggink, JK Hooper and C Funk (2008) The small CAB-like proteins of *Synechocystis* sp. PCC 6803 bind chlorophyll. *In vitro* pigment reconstitution studies on one-helix light-harvesting-like proteins. *Photosynth. Res.* 98 :479-488.
- [164] M Heddad, I Adamska (2002) The evolution of light stress proteins in photosynthetic organisms. *Comp Funct Genomics* 3 :504-510.
- [165] JAD Neilson, G Dion and DG Durnford (2010) Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. *Photosynth Res* 106 :57-71.
- [166] KM Reinisch, SL Wolin (2007) Emerging themes in non-coding RNA quality control. *Current Opinion in Structural Biology* 17 :209-214.
- [167] M-A Dillies, A Rau, J Aubert, C Hennequet-Antier, M Jeanmougin, N Servant, C Keime, G Marot, D Castel, J Estelle, G Guernec, B Jagla, L Jouneau, D Laloë, C Le Gall, B Schaëffer, S Le Crom, M Guedj, F Jaffrézic (2012) A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform*
- [168] DJ McCarthy, Y Chen, GK Smyth (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* 40 :4288-4297.
- [169] Y Benjamini, Y Hochberg (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B - Methodology* 57 :289-300.
- [170] A Le Bail, SM Dittami, PO de Franco, S Rousvoal, MJ Cock, T Tonon, B Charrier (2008) Normalisation genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. *BMC Mol Biol.* 9 :75.
- [171] F Thomas, T Barbeyron, G Michel (2011) Evaluation of reference genes for real-time quantitative PCR in the marine flavobacterium *Zobellia galactanivorans*. *J Microbiol Methods.* 84 :61-6.
- [172] P Chen, HB Shao, D Xu (2011) Cloning and characterization of a female gametophyte-specific gene in *Gracilaria lemaneiformis* (Gracilariales, Rhodophytes). *African Journal of Microbiology Research* 5 :2590-2595.
- [173] XY Ren, XC Zhang (2006). Identification of a putative tetrasporophytespecific gene in *Gracilaria lemaneiformis* (Gracilariales, Rhodophyte). *J. Ocean, Univ. Chin.* 3 :299-303.
- [174] M-L Guillemain, OR Huanel, EA Martinez (2012) Characterization of genetic markers linked to sex determination in the haploid-diploid red alga *Gracilaria chilensis*. *J. Phycol.* 48 :365-372.
- [175] XY Ren, XC Zhang, YX Mao, ZH Sui, D Xu, XN Zang (2008) Cloning and characterization of a Rab11 homologue in *Gracilariopsis lemaneiformis*. *J. Appl. Phycol.* 20 :1103-1109.
- [176] X Ren, Z Sui, Y Mao, X Zang, D Xu, X Zhang (2009) Cloning and characterization of two types of ubiquitin genes from *Gracilariopsis lemaneiformis* (Gracilariales, Rhodophyta). *Journal of Applied Phycology* 21 :273-278.

- [177] YK Lee, SH Kim, JB Hong, O-K Chah, GH Kim, IK Lee (1998) Heat-shock protein 90 may be involved in differentiation of the female gametophytes in *Griffithsia japonica* (Ceramiales, Rhodophyta). *J. Phycol.* 34 :1017-1023.
- [178] E Asamizu, M Nakajima, Y Kitade (2003) Comparison of RNA expression profiles between the two generations of *Porphyra yezoensis* (Rhodophyta), based on expressed sequence tag frequency analysis. *J Phycol.* 930 :923-930.
- [179] M Kakinuma, I Kaneko, D Coury, T Suzuki, H Amano (2006) Isolation and Identification of Gametogenesis-Related Genes in *Porphyra Yezoensis* (Rhodophyta) Using Subtracted cDNA Libraries. *Journal of Applied Phycology* 18 :489-496.
- [180] S Shen, G Zhang, Y Li, L Wang, P Xu, L Yi (2011) Comparison of RNA expression profiles on generations of *Porphyra yezoensis* (Rhodophyta), based on suppression subtractive hybridization (SSH). *BMC research notes* 4 :428.
- [181] X Ren, X Zhang, Z Sui (2006) Identification of phase relative genes in tetrasporophytes and female gametophytes of *Gracilaria/Gracilariopsis lemaneiformis* (Gracilariales, Rhodophyta). *Electron J Biotechn* 9 :127-132.
- [182] M Kamiya, H Kawai, D Moon, LJ Goff (2011) Isolation and characterization of phasespecific cDNAs from carposporophytes of *Gracilariopsis andersonii* (Gracilariales, Rhodophyta). *Eur J Phycol.* 46 :27-35.
- [183] <http://www.uniprot.org/>
- [184] <http://www.geneontology.org/>
- [185] M Ashburner, CA Ball, JA Blake, D Botstein, H Butler, JM Cherry, AP Davis, K Dolinski, SS Dwight, JT Eppig, MA Harris, DP Hill, L Issel-Tarver, A Kasarskis, S Lewis, JC Matese, JE Richardson, M Ringwald, GM Rubin, G Sherlock (2000) Gene ontology : tool for the unification of biology. *Nat Genet.* 25 :25-9.
- [186] <http://www.blast2go.org/>
- [187] JS Craigie, KF Wong (1979) Carrageenan biosynthesis. *Proc Int Seaweed symposium* 9 :369-377.
- [188] AI Saeed, NK Bhagabati, JC Braisted, W Liang, V Sharov, EA Howe, J Li, M Thiagarajan, JA White, J Quackenbush (2006) TM4 microarray software suite. *Methods Enzymol.* 411 :134-93.
- [189] G Noctor, A Mhamdi, S Chaouch, Y Han, J Neukermans, B Marquez-Garcia, G Queval, CH Foyer (2012) Glutathione in plants : an integrated overview. *Plant Cell Environ.* 35 :454-84.
- [190] J Collén, IR Davison (2003) Diurnal changes in photosynthesis and reactive oxygen metabolism in *Gracilaria tikvahiae* (Rhodophyta). *Proceedings International Seaweed Symposium* Oxford University Press. pp 395-401.
- [191] J Ramus, G Rosenberg (1980) Diurnal photosynthetic performance of seaweeds measured under natural conditions. *Marine Biology* 56 :21-28.
- [192] S Sagert, RM Forste, P Feuerpfeil, H Schubert (2000) Daily course of photosynthesis and photoinhibition in *Chondrus crispus* (Rhodophyta) from different shore levels. *European Journal of Phycology* 32 :363-371.
- [193] T Strömberg, MV Nielsen (1986) Effect of diurnal variations in natural irradiance on the apical length growth and light saturation of growth in five species of benthic macroalgae. *Marine Biology* 90 :467-472.

- [194] CH Lillig, C Bernd, A Holmgren (2008) Glutaredoxin systems. *Biochimica et Biophysica Acta* 1780 :1304-1317.
- [195] F Ursini, M Maiorino, R Brigeliusflohe KD Aumann, A Roveri, D Schomburg, L Flohe (1995) Diversity of glutathione peroxidases. *Methods in enzymology* 252 :38-53.
- [196] LH Fu, XF Wang, Y Eyal, YM She, LJ Donald, KG Standing, G Ben-Hayyim (2002) A Selenoprotein in the Plant Kingdom, Mass spectrometry confirms that an opal codon (uga) encodes selenocysteine in *Chlamydomonas reinhardtii* glutathione peroxidase. *Journal of Biological Chemistry* 277 :25983-25991.
- [197] CCC Chang, I Ślesak, L Jordá, A Sotnikov, MI Melzer, Z Miszalski, PM Mullineaux, JE Parker, B Karpińska, S Karpiński (2009) *Arabidopsis* Chloroplastic Glutathione Peroxidases Play a Role in Cross Talk between Photooxidative Stress and Immune Responses. *Plant Physiology* 150 :670-683.
- [198] C Herve, PO De Franco, A Groisillier, T Tonon, C Boyen (2008) New members of the glutathione transferase family discovered in red and brown algae. *Biochem. J.* 412 :535-544.
- [199] C Gachon, A Mingam, B Charrier (2004) Real-time PCR : What relevance to plant studies? *J Exp Bot* 55 :1445-1454.
- [200] SA Bustin, V Benes, T Nolan, MW Pfaffl (2005) Quantitative real-time RT-PCR? a perspective. *J Mol Endocrinol* 34 :597-601.
- [201] J Vandesompele, K De Preter, F Pattyn, B Poppe, N Van Roy, A De Paepe, F Speleman (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3 :R34.1-34.11.
- [202] CL Andersen, JL Jensen, TF Ørntoft (2004) Normalization of real-time quantitative reverse transcription-PCR data : a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64 :5245-5250.
- [203] MW Pfaffl, A Tichopád, C Prgomet, TP Neuvians (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity : BestKeeper? Excel-based tool using pair-wise correlations. *Biotechnol Lett* 26 :509-515.
- [204] G Spinsanti, C Panti, E Lazzeri, L Marsili, S Casini, F Frati, CM Fossi (2006) Selection of reference genes for quantitative RT-PCR studies in striped dolphin (*Stenella coeruleoalba*) skin biopsies. *BMC Mol Biol* 7 :32.
- [205] HC Ingerslev, EF Pettersen, RA Jakobsen, CB Petersen, HI Wergeland (2006) Expression profiling and validation of reference gene candidates in immune relevant tissues and cells from Atlantic salmon (*Salmo salar* L.). *Mol Immunol* 43 :1194-1201.
- [206] D Hoogewijs, K Houthoofd, F Matthijssens, J Vandesompele, JR Vanfleteren (2008) Selection and validation of a set of reliable reference genes for quantitative sod gene expression analysis in *C. elegans*. *BMC Mol Biol* 9 :9.
- [207] M Jain, A Nijhawan, AK Tyagi, JP Khurana (2006) Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochem Biophys Res Commun* 345 :646-651.
- [208] KE Reid, N Olsson, J Schlosser, F Peng, ST Lund (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. *BMC Plant Biol* 6 :27.

- [209] L Chen, HY Zhong, JF Kuang, JG Li, WJ Lu, JY Chen (2011) Validation of reference genes for RT-qPCR studies of gene expression in banana fruit under different experimental conditions. *Planta* 234 :377-390.
- [210] N Lu, X Zang, X Zhang, H Chen, X Feng, L Zhang (2012) Gene cloning, expression and activity analysis of manganese superoxide dismutase from two strains of *Gracilaria lemaneiformis* (Gracilariaceae, Rhodophyta) under heat stress. *Molecules* 17 :4522-32.
- [211] CL Ho, S Teoh, SS Teo, RA Rahim, SM Phang (2009) Profiling the transcriptome of *Gracilaria changii* (Rhodophyta) in response to light deprivation. *Mar Biotechnol* (NY) 11 :513-9.
- [212] SS Teo, CL Ho, S Teoh, RA Rahim, SM Phang (2009) Transcriptomic analysis of *Gracilaria changii* (rhodophyta) in response to hyper- and hypoosmotic stresses *J. Phycol* 45 :1093-1099.
- [213] XJ Wu, JF Niu, AY Huang, ML Xu, GC Wang (2012) Selection of internal control gene for expression studies in porphyra haitanensis (rhodophyta) at different life-history stages. *J Phycol* 48 :1040-1044.
- [214] M Yokono, A Murakami, S Akimoto (2011) Excitation energy transfer between photosystem II and photosystem I in red algae : larger amounts of phycobilisome enhance spillover. *Biochim. Biophys. Acta* 1807 :847-853.



## Annexes

TABLE 20: Le consortium du génome de *Chondrus crispus*. Les laboratoires marqués d'une étoile sont les porteurs du projet.

The *Chondrus crispus* genome consortium. The laboratories bearing the project marked with an asterisk.

Group leader	Location
<b>Europe</b>	
Catherine Boyen*	UMR7139, Station Biologique, Roscoff, France
Mark Cock*	UMR7139, Station Biologique, Roscoff, France
Mirjam Czjzek*	UMR7139, Station Biologique, Roscoff, France
Philip Potin*	UMR7139, Station Biologique, Roscoff, France
William Helbert*	UMR7139, Station Biologique, Roscoff, France
Myriam Valero *	UMR7144, Station Biologique, Roscoff, France
Christine Maggs*	Queens University, Belfast, UK
Richard Reinhardt*	Max-Planck-Institute for Molecular Genetics, Berlin, Germany
Florian Weinberger*	IFM-GEOMAR Kiel, Germany
Klaus Valentin*	AWI, Bremerhaven, Germany
Yves Van de Peer*	Ghent University, Ghent, Belgium
Alvaro Israel	Oceanographic Institute, Haifa, Israel
Georg Pohnert	École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
Felix Figueroa	Universidad de Málaga, Málaga, Spain
Isabel Sousa Pinto	CIMAR Porto, Portugal
John Raven	University of Dundee, Dundee, UK
Mike Guiry	University of Ireland, Galway, Ireland
Ulf Karsten	AWI, Rostock, Germany
Wiebe Kooistra	Stazione Zoologica Naples, Italy
<b>America</b>	
Mariana de Oliveira*	University of Sao Paulo, Brazil
Debashish Bhattacharya *	University of Iowa, Iowa City, USA
Juan Correa*	Pontificia Universidad Catolica de Chile, Santiago, Chile
Thierry Chopin*	University of New Brunswick, New Brunswick, Canada
Alberto S. Cerezo	University of Buenos Aires, Buenos Aires Argentina
Gary Saunders	University of New Brunswick Fredericton, Canada
Lynda Goff	University of California Santa Cruz
Linda Franklin	Franklin and Marshall College, Lancaster, Pennsylvania, USA
<b>Asia</b>	
Duan Delin*	Institute of Oceanology, Qingdao, China
Sung Min Boo	Chungnam National University, Daejeon, Korea
Phang Siew Moi	University of Malaya Kuala Lumpur, Malaysia
Yukihiro Kitade	Hokkaido University, Japan
<b>Oceania</b>	
Cheong Xin Chan	The University of Queensland, Brisbane, Australia
Joe Zuccarello	University of Wellington Wellington, New Zealand
Judy Broom	University of Otago, Dunedin, New Zealand
<b>Africa</b>	
Linekela Kandjengo	University of Namibia, Windhoek, Namibia
<b>Industrial</b>	
Alan Critchley*	Acadian Seaplants Limited, Dartmouth, Canada
Erick Ask	FMC BioPolymer, Philadelphia, USA
Rudolph, Brian	Cp Kelco ApS, Lille Skensved, Denmark

FIGURE 31: Prédiction de modèles de gènes du *scaffold* 62. En fond gris : aucune correspondance dans la base de donnée NCBI, en violet : annotés parmi les gènes de stress. Le code couleur de la colonne "best match" se trouve en bas de la figure.

Prediction of gene models of the scaffold 62. Grey background : no match in NCBI database, in purple : annotated among the stress genes. The color code in the "best match" column is at the bottom of the table.

Gene ID	aa	Prediction	EC	E value	Best match
GSCHCT00012987001	404	Glycosyl transferase, group 1	NA	6,00E-083	<i>Tetrahymena thermophila</i>
GSCHCT00012988001	583	Calreticulin/calnexin	NA	1,00E-107	<i>Arabidopsis thaliana</i>
GSCHCT00012989001	35	hypothetical protein	NA	NA	NA
GSCHCT00012990001	952	trehalose-6-phosphate synthase	2.4.1.15	0	<i>Undaria pinnatifida</i>
GSCHCT00012991001	186	hypothetical protein	NA	NA	NA
GSCHCT00012992001	139	hypothetical protein	NA	NA	NA
GSCHCT00012993001	379	hypothetical protein	NA	1,00E-005	<i>Bacteroides</i> sp.
GSCHCT00012994001	118	hypothetical protein	NA	NA	NA
GSCHCT00012995001	56	hypothetical protein	NA	NA	NA
GSCHCT00012996001	62	hypothetical protein	NA	NA	NA
GSCHCT00012999001	275	unknown protein	NA	1,00E-032	<i>Thalassiosira pseudonana</i>
GSCHCT00013000001	1157	Ubiquitin-protein ligase	6.3.2.19	2,00E-053	<i>Ricinus communis</i>
GSCHCT00013001001	122	hypothetical protein	NA	NA	NA
GSCHCT00013002001	601	Oxidoreductase	1.14.12.17	6,00E-106	<i>Marinomonas</i> sp.
GSCHCT00013004001	340	Prostaglandin-E synthase	5.3.99.3	1,00E-055	<i>Vitis vinifera</i>
GSCHCT00013007001	265	hypothetical protein	NA	NA	NA
GSCHCT00013009001	167	HSP20	NA	NA	NA
GSCHCT00013010001	232	hypothetical protein	NA	NA	NA
GSCHCT00013011001	534	ATP-binding protein of ABC transporter	3.6.3.44	6,00E-119	<i>Oscillatoria</i>
GSCHCT00013012001	269	C-4 methyl sterol oxidase	1.14.13.72	3,00E-063	<i>Dictyostelium discoideum</i>
GSCHCT00013013001	168	HSP20	NA	NA	NA
GSCHCT00013015001	200	hypothetical protein	NA	NA	NA
GSCHCT00013016001	301	WD40-repeat-containing protein	NA	4,00E-042	<i>Salmo salar</i>
GSCHCT00013017001	96	hypothetical protein	NA	NA	NA
GSCHCT00013018001	221	hypothetical protein	NA	NA	NA
GSCHCT00013020001	1399	C2 domain containing protein	NA	9,00E-018	<i>Tetrahymena thermophila</i>
GSCHCT00013021001	617	? Histone-lysine N-methyltransferase	2.1.1.43	1,00E-050	<i>Volvox carteri</i>
GSCHCT00013022001	59	hypothetical protein	NA	NA	NA
GSCHCT00013023001	34	hypothetical protein	NA	NA	NA
GSCHCT00013024001	312	hypothetical protein	NA	NA	NA
GSCHCT00013026001	39	hypothetical protein	NA	NA	NA
GSCHCT00013028001	79	hypothetical protein	NA	NA	NA
GSCHCT00015332001	569	Amidinotransferase	2.1.4.X	7,00E-081	<i>Chlamydomonas reinhardtii</i>
GSCHCT00013031001	1460	hypothetical protein	NA	7,00E-025	NA
GSCHCT00013033001	399	glycosyltransferase family 14 protein	2.4.1.X	7,00E-025	<i>Oriza sativa</i>
GSCHCT00013035001	653	glutathione S-transferase	NA	1,00E-028	<i>Taeniopygia guttata</i>
GSCHCT00013036001	304	Glutaredoxin	1.20.4.1	NA	NA
GSCHCT00013037001	508	PUFA elongation enzyme	NA	2,00E-032	<i>Thraustochytrium</i>
GSCHCT00013038001	880	hypothetical protein	NA	NA	NA
GSCHCT00013044001	580	conserved hypothetical protein	NA	1,00E-008	<i>Citrobacter</i>
GSCHCT00013048001	269	Incomplete AP-3 complex subunit beta-1	NA	5,00E-033	<i>Ixodes scapularis</i>
GSCHCT00013049001	188	hypothetical protein	NA	NA	NA
GSCHCT00013050001	169	hypothetical protein	NA	NA	NA
GSCHCT00013051001	1105	AP-3 complex subunit beta-1	NA	4,00E-108	<i>Xenopus laevis</i>
GSCHCT00013052001	546	Pentatricopeptide repeat containing protein	NA	2,00E-026	<i>Selaginella</i>
GSCHCT00013053001	62	hypothetical protein	NA	NA	NA
GSCHCT00013054001	51	hypothetical protein	NA	NA	NA
GSCHCT00013056001	477	acetyl-CoA carboxylase	6.3.4.14	4,00E-175	<i>Synechococcus</i>
GSCHCT00013057001	1455	ubiquitin thiolesterase UBP13	3.1.2.15	2,00E-122	<i>Arabidopsis thaliana</i>
GSCHCT00013059001	600	transcriptional repressor TUP1-like	2.7.11.7	3,00E-106	<i>Dictyostelium discoideum</i>
GSCHCT00013060001	785	transcriptional repressor TUP1-like	2.7.11.7	1,00E-105	<i>Dictyostelium discoideum</i>
GSCHCT00013061001	54	hypothetical protein	NA	NA	NA

Fungi & Animals	Stramenopiles	Cyanobacteria	Eukaryota
Viridiplanta	Bacteria	Rhodophyta	



FIGURE 32: Listes des gènes impliqués dans la réponse au stress oxydatif, identifiés au sein du génome de *C. crispus*. EC : numéro d'enzyme, aa : nombre d'acide aminés, %ID : pourcentage d'identité avec le meilleur score de la base de donnée NCBI, E value : probabilité d'obtenir l'alignement par hasard, SP : peptide signal.

List of genes implicated in the response to oxidative stress, identified in the genome of *C. crispus*. EC : enzyme number, aa ; number of aminoacids, %ID : percentage of identity with the best match in NCBI database, E value : probability to obtain the alignment by chance, SP : signal peptide.

Name	Full Name	EC	aa	Best match	% ID	E value	SP
<b>Oxidoreductases</b>							
APX1	Ascorbate Peroxidase	1.11.1.11	237	<i>Galdieria sulphuraria</i>	62	1,00E-085	no
APX2		1.11.1.11	297	<i>Galdieria sulphuraria</i>	66	5,00E-092	no
SOD Cu/Zn	Superoxide Dismutase	1.15.1.1	202	<i>Rhodothermus marinus</i>	43	7,00E-023	yes
SOD Cu/Zn		1.15.1.1	295	<i>Rhodothermus marinus</i>	45	1,00E-020	yes
SOD Fe/Mn	Mn Superoxide Dismutase	1.15.1.1	224	<i>Porphyra haitanensis</i>	66	4,00E-088	no
SOD Fe/Mn		1.15.1.1	311	<i>Porphyra yezoensis</i>	55	1,00E-073	yes
CAT	Catalase	1.11.1.6	494	<i>Ulva fasciata</i>	57	3,00E-164	no
CAT		1.11.1.6	513	<i>Alteromonadales bacterium</i>	55	8,00E-151	yes
GR	GR Glutathione Reductase	1.8.1.7	515	<i>Parvibaculum lavamentivorans</i>	56	9,00E-142	yes
MDHAR	Monodehydroascorbate Reductase	1.6.5.4	476	<i>Selaginella moellendorffii</i>	43	6,00E-095	no
DHAR/GDH	Dehydroascorbate Reductase	1.8.5.1	271	<i>Ectocarpus siliculosus</i>	35	1,00E-032	no
DHAR/GDH	GST Thioredoxin-fold	NA	256	<i>Selaginella moellendorffii</i>	30	4,00E-022	no
DHAR/GDH	GST Thioredoxin-fold	NA	256	<i>Selaginella moellendorffii</i>	30	4,00E-022	no
GPX	Glutathione Peroxidase	1.11.1.9	265	<i>Alteromonas macleodii</i>	49	1,00E-034	no
GPX		1.11.1.9	164	<i>Griffithsia japonica</i>	35	1,00E-018	no
PRX2	Peroxiredoxin	1.11.1.15	227	<i>Physcomitrella patens</i>	56	3,00E-075	no
PRX3		1.11.1.15	212	<i>Acinetobacter baumannii</i>	73	1,00E-088	no
PRX1		1.11.1.15	249	<i>Phaeodactylum tricornutum</i>	57	1,00E-077	no
GRX1	Glutaredoxin	1.20.4.1	313	<i>Phaeodactylum tricornutum</i>	41	1,00E-062	no
GRX2		1.20.4.1	304	<i>Ricinus communis</i>	38	2,00E-038	no
GRX3		1.20.4.1	156	<i>Ostreococcus tauri</i>	46	1,00E-028	no
GRX4		1.20.4.1	152	<i>Fritillaria agrestis</i>	55	3,00E-025	yes
PMSR	Peptide Methionine Sulfoxide Reductase	1.8.4.6	556	<i>Gracilaria gracilis</i>	51	9,00E-113	no
PMSR		1.8.4.6	537	<i>Volvox carteri</i>	50	8,00E-079	yes
TXR	Thioredoxine Reductase	1.8.1.9	364	<i>Physcomitrella patens</i>	61	2,00E-117	no
PRXase	Peroxidase	1.11.1.-	589	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	536	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	no
PRXase		1.11.1.-	555	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	563	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	581	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	561	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	566	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	574	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	703	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	614	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	604	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	560	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	540	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	546	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	564	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	631	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	571	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
PRXase		1.11.1.-	534	<i>Planctomycetes</i>	≈ 60	≈ 1E-070	yes
RDH	Retinol Dehydrogenase	1.1.1.300	353	<i>Xenopus laevis</i>	35	3,00E-040	no

FIGURE 33: Listes des gènes correspondants aux Heat Shock Proteins, identifiés au sein du génome de *C. crispus*. EC : numéro d'enzyme, aa : nombre d'acide aminés, %ID : pourcentage d'identité avec le meilleur score de la base de donnée NCBI, E value : probabilité d'obtenir l'alignement par hasard, SP : peptide signal.  
List of genes corresponding to Heat SHock Proteins, identified in the genome of *C. crispus*. EC : enzyme number, aa ; number of aminoacids, %ID : percentage of identity with the best match in NCBI database, E value : probability to obtain the alignment by chance, SP : signal peptide.

#### Heat Shock Proteins

HSP90	HSP90	NA	838	<i>Thalassiosira pseudonana</i>	47	0	no
HSP90	HSP90	NA	739	<i>Porphyra yezoensis</i>	78	0	no
HSP90	HSP90	NA	781	<i>Populus trichocarpa</i>	50	0	no
iHSP90	incomplete hsp90	NA	298	<i>Griffithsia japonica</i>	60	5,00E-065	no
ClpA/B		NA	786	<i>Arthrospira maxima</i>	64	0	no
ClpA/B	HSP100	NA	745	<i>Gracilaria tenuistipitata</i>	95	0	no
HSP20		NA	153	<i>Salinibacter ruber</i>	31	4,00E-009	yes
HSP20		NA	168	<i>Meiothermus ruber</i>	29	4,00E-006	no
HSP20	Heat Shock Protein 20	NA	166	<i>Griffithsia japonica</i>	30	3,00E-006	no
HSP20		NA	149	<i>Griffithsia japonica</i>	31	6,00E-005	no
HSP20		NA	167	<i>Griffithsia japonica</i>	35	9,00E-008	no
HSP20		NA	162	<i>Salinibacter ruber</i>	29	4,00E-008	? P=0,54
PDIA	Protein Disulfide Isomerase	5.3.4.1	492	<i>Neurospora crassa</i>	40	3,00E-080	yes
PDIA		5.3.4.1	539	<i>Tetrahymena thermophila</i>	32	9,00E-058	yes
PPIA		5.2.1.8	207	<i>Ciona intestinalis</i>	49	1,00E-044	no
PPIA		5.2.1.8	170	<i>Ostreococcus tauri</i>	50	2,00E-037	no
PPIA		5.2.1.8	109	<i>Physcomitrella patens</i>	66	4,00E-035	no
PPIA		5.2.1.8	430	<i>Naegleria gruberi</i>	33	9,00E-045	no
PPIA	Peptidyl-Prolyl cis-trans Isomerase / Calreticulin	5.2.1.8	114	<i>Griffithsia japonica</i>	60	2,00E-022	yes
PPIA		5.2.1.8	165	<i>Planctomyces limnophilus</i>	80	8,00E-071	no
PPIA		5.2.1.8	191	<i>Planctomyces maris</i>	53	6,00E-051	no
PPIA		5.2.1.8	418	<i>Arabidopsis thaliana</i>	47	9,00E-044	no
PPIA		5.2.1.8	407	<i>Taeniopygia guttata</i>	53	1,00E-095	no
PPIA		5.2.1.8	218	<i>Venerupis philippinarum</i>	49	1,00E-038	no
Calnexin	Calnexin	NA	583	<i>Arabidopsis thaliana</i>	43	1,00E-107	no
HSF	Heat Shock Transcription Factor	NA	464	<i>Glycine max</i>	39	2,00E-043	no

#### T Complex

TCP1 alpha	T-Complex Protein alpha	NA	544	<i>Oryza sativa</i>	66	0	no
TCP1 beta	T-Complex Protein beta	NA	538	<i>Branchiostoma floridae</i>	66	0	no
TCP1 gamma	T-Complex Protein gamma	NA	547	<i>Physarum polycephalum</i>	63	0	no
TCP1 delta	T-Complex Protein delta	NA	558	<i>Selaginella moellendorffii</i>	61	0	no
TCP1 epsilon	T-Complex Protein epsilon	NA	537	<i>Micromonas sp.</i>	68	0	no
TCP1 zeta	T-Complex Protein zeta	NA	529	<i>Physcomitrella patens</i>	61	0	no
TCP1 eta	T-Complex Protein eta	NA	520	<i>Physcomitrella patens</i>	61	0	no
TCP1 theta	T-Complex Protein theta	NA	573	<i>Nematostella vectensis</i>	54	2,00E-173	no

#### DnaK DnaJ GrpE

HSP70	DnaK Chaperone	NA	656	<i>Porphyra yezoensis</i>	81	0	no
HSP70	DnaK Chaperone	NA	668	<i>Gossypium hirsutum</i>	68	0	yes
HSP70	DnaK Chaperone	NA	647	<i>Gossypium hirsutum</i>	68	0	yes
HSP70	DnaK Chaperone	NA	631	<i>Selaginella moellendorffii</i>	74	0	no
HSP70	DnaK Chaperone	NA	583	<i>Gracilaria tenuistipitata</i>	78	0	no
HSP70	DnaK Chaperone	NA	874	<i>Populus trichocarpa</i>	36	7,00E-137	no
HSP70	DnaK Chaperone	NA	939	<i>Monodelphis domestica</i>	27	8,00E-078	yes
iHSP70	Incomplete HSP70	NA	294	<i>Ricinus communis</i>	30	1,00E-015	no
HSP40/DnaJ		NA	413	<i>Phytophthora infestans</i>	48	5,00E-101	no
HSP40/DnaJ		NA	412	<i>Phytophthora infestans</i>	41	4,00E-068	no
HSP40/DnaJ		NA	444	<i>Acaryochloris marina</i>	57	8,00E-117	yes
HSP40/DnaJ	Heat Shock Protein 40 / DnaJ	NA	331	<i>Picea sitchensis</i>	44	5,00E-075	no
HSP40/DnaJ		NA	423	<i>Phytophthora infestans</i>	38	1,00E-064	no
HSP40/DnaJ		NA	413	<i>Arabidopsis thaliana</i>	32	8,00E-044	no
HSP40/DnaJ		NA	399	<i>Canis familiaris</i>	31	8,00E-035	no
HSP70-BP	HSP70 binding protein	NA	443	<i>Ostreococcus tauri</i>	42	1,00E-091	no
GrpE	HSP GrpE	NA	320	<i>Physcomitrella patens</i>	43	2,00E-030	yes
GrpE		NA	234	<i>Thalassiosira pseudonana</i>	47	1,00E-037	yes

FIGURE 34: Suite de la table 33.  
Table 33, continued.

<b>Tubulin Binding Complex</b>							
TBCA	no found						
TBCB	tubulin folding cofactor B	NA	279	<i>Volvox carteri</i>	48	8,00E-063	no
TBCC	tubulin folding cofactor C	NA	311	<i>Vitis vinifera</i>	31	3,00E-028	no
TBCD	tubulin folding cofactor D	NA	686	<i>Thalassiosira pseudonana</i>	35	2,00E-085	no
TBCE	tubulin folding cofactor E	NA	537	<i>Thalassiosira pseudonana</i>	35	2,00E-085	no
<b>GroES GroEL</b>							
HSP60	CPN60/GroEL	NA	585	<i>Guillardia theta</i>	62	0	no
HSP60	CPN60/GroEL	NA	577	<i>Ostreococcus lucimarinus</i>	70	4,00E-170	no
iGroEL	Incomplete GroEL	NA	75	<i>Populus trichocarpa</i>	87	2,00E-017	no
CPN10		NA	101	<i>Griffithsia japonica</i>	81	4,00E-040	no
CPN10	HSP10/GroES	NA	242	<i>Thalassiosira pseudonana</i>	40	8,00E-035	no

TABLE 21: Annotation manuelle des gènes de biosynthèse du phycobilisome chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible, APC : allophycocyanine, PBS : phycobilisome, PC : phycocyanine, PE : phycoérythrine. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret. Manual annotation of phycobilisome biosynthesis genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data, APC : allophycocyanin, PBS : phycobilisome, PC : phycocyanin, PE : phycoerythrin. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
apcA	allophycocyanin $\alpha$ chain	Ccr_CDS_125	chloroplast
apcB	allophycocyanin $\beta$ chain	Ccr_CDS_126	chloroplast
apcC	phycobilisome APC-associated core linker polypeptide	GSCHCT00009753001	nuclear
apcD	allophycocyanin $\gamma$ chain	Ccr_CDS_8	chloroplast
apcE	phycobilisome core-membrane linker polypeptide	Ccr_CDS_124	chloroplast
apcF	allophycocyanin $\beta$ -18 subunit	Ccr_CDS_161	chloroplast
cpcA	phycocyanin $\alpha$ chain	Ccr_CDS_176	chloroplast
cpcB	phycocyanin $\beta$ chain	Ccr_CDS_175	chloroplast
cpcE	phycocyanobilin Cys-84 $\alpha$ -C-phycocyanin lyase	GSCHCT00013625001	nuclear
cpcF	phycocyanobilin Cys-84 $\alpha$ -C-phycocyanin lyase	-	-
cpcG	phycobilisome rod-core linker polypeptide	Ccr_CDS_163	chloroplast
cpcS	phycocyanobilin :Cys-84 $\beta$ -phycocyanin lyase	GSCHCT00010740001	nuclear
cpcT	phycocyanobilin :Cys-153 $\beta$ -phycocyanin lyase	GSCHCT00002841001	nuclear
mpeU	phycoerythrobilin phycoerythrin lyase	GSCHCT00013138001	nuclear
/	PBS lyase HEAT domain protein repeat-containing protein	GSCHCT00000578001	nuclear
/	PBS lyase HEAT domain protein repeat-containing protein	GSCHCT00010574001	nuclear
cpeA	R-phycoerythrin class I $\alpha$ subunit	Ccr_CDS_166	chloroplast
cpeB	R-phycoerythrin class I $\beta$ subunit	Ccr_CDS_165	chloroplast
/	PBS rod linker polypeptide, possibly C-CP-associated	GSCHCT00007564001	nuclear
/	PBS rod linker polypeptide, possibly R-PE-associated	GSCHCT00007648001	nuclear
/	PBS rod linker polypeptide, possibly R-PE-associated	GSCHCT00012349001	nuclear
/	PBS rod linker polypeptide, possibly R-PE-associated	GSCHCT00003298001	nuclear
/	PBS rod linker polypeptide, possibly R-PE-associated	GSCHCT00005887001	nuclear
/	PBS rod linker polypeptide, possibly R-PE-associated	GSCHCT00012567001	nuclear
rpeG1	R-phycoerythrin $\gamma$ subunit	GSCHCT00006429001	nuclear
rpeG2	R-phycoerythrin $\gamma$ subunit	GSCHCT00000799001	nuclear
rpeG3	R-phycoerythrin $\gamma$ subunit	GSCHCT00005946001	nuclear
cpeR	phycoerythrin operon regulator	-	-
cpeS	phycoerythrobilin-phycoerythrin lyase	NA	nuclear
cpeT	phycoerythrobilin-phycoerythrin lyase	-	-
cpeU	phycoerythrobilin-phycoerythrin lyase	-	-
cpeY	phycoerythrobilin-phycoerythrin lyase	GSCHCT00006389001	nuclear
cpeZ	phycoerythrobilin-phycoerythrin lyase	NA	nuclear
pebA	15,16-dihydrobiliverdin-ferredoxin oxidoreductase	GSCHCT00014694001	nuclear
pebB	phycoerythrobilin-ferredoxin oxidoreductase	GSCHCT00006933001	nuclear
pcyA	phycocyanobilin :ferredoxin oxidoreductase	-	-
bvdR	biliverdin reductase	-	-
hy2	phytochromobilin synthase	-	-
rpaA	winged helix OmpR family protein	-	-
rpoA	DNA-directed RNA polymerase subunit $\alpha$	Ccr_CDS_78	chloroplast
rpoB	winged helix OmpR family protein	Ccr_CDS_131	chloroplast
nblA	phycobilisome degradation protein	Ccr_CDS_164	chloroplast

TABLE 22: Annotation manuelle des gènes d'antenne du photosystème I chez *C. crispus*.  
CDS : séquence codante, NA : donnée indisponible, PS I : photosystème I. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.  
Manual annotation of photosystem I structural genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data, PS I : photosystem I. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
lhcr1	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00000761001	nuclear
lhcr2	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00012812001	nuclear
lhcr3	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00014077001	nuclear
lhcr4	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00014619001	nuclear
lhcr5	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00000400001	nuclear
lhcr6	PSI light-harvesting chlorophyll a/b binding protein	GSCHCT00000796001	nuclear
lhcr7	PSI light-harvesting chlorophyll a/b binding protein	GSCHC2T00014564001	nuclear
lhcb4	CP24		
lhcb4	CP29-like light-harvesting chlorophyll a/b binding protein		
lhcb5	CP26-like light-harvesting chlorophyll a/b binding protein		
li818r-like	light-harvesting chlorophyll a/b binding protein		
hlip1	CAB/ELIP/HLIP-related protein	GIDCcT00025272001	nuclear
hlip2	CAB/ELIP/HLIP-related protein	SNAPCcT00034592001	nuclear
hlip3	CAB/ELIP/HLIP-related protein	GIDCcT00025274001	nuclear
hlip4	CAB/ELIP/HLIP-related protein	GSCHCT00000672001	nuclear
hlip5	CAB/ELIP/HLIP-related protein	GSCHCT00010668001	nuclear
hlip6	CAB/ELIP/HLIP-related protein	GSCHCT00010671001	nuclear
hlip	CAB/ELIP/HLIP-related protein	CCr_CDS_199	chloroplast
elip	CAB/ELIP/HLIP-related protein	SNAPCcT00037586001	nuclear
lil-like	light-induced-like protein		

TABLE 23: Annotation manuelle des gènes de structure du photosystème I chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.

Manual annotation of photosystem I structural genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
psaA	photosystem I P700 chlorophyll a apoprotein A1	Ccr_CDS_106	chloroplast
psaB	photosystem I P700 chlorophyll a apoprotein A2	Ccr_CDS_105	chloroplast
psaC	photosystem I subunit VII	Ccr_CDS_172	chloroplast
psaD	photosystem I subunit II	Ccr_CDS_110	chloroplast
psaE	photosystem I subunit IV	Ccr_CDS_64	chloroplast
psaF	photosystem I subunit III	Ccr_CDS_6	chloroplast
psaG	photosystem I subunit V	-	-
psaH	photosystem I subunit VI	-	-
psaI	photosystem I subunit VIII	Ccr_CDS_184	chloroplast
psaJ	photosystem I subunit IX	Ccr_CDS_7	chloroplast
psaK	photosystem I subunit X	Ccr_CDS_168	chloroplast
psaL	photosystem I subunit XI	Ccr_CDS_148	chloroplast
psaM	photosystem I subunit XII	Ccr_CDS_1	chloroplast
psaN	photosystem I subunit N	-	-
psaO	photosystem I subunit O	GSCHCT00002577001	nuclear

TABLE 24: Annotation manuelle des gènes de structure du photosystème II chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.

Manual annotation of phycobilisome biosynthesis genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
psbA	photosystem II D1 core protein	Ccr_CDS_25	chloroplast
psbB	photosystem II CP47 protein	Ccr_CDS_68	chloroplast
psbC	photosystem II CP43 protein	Ccr_CDS_35	chloroplast
psbD	photosystem II D2 core protein	Ccr_CDS_34	chloroplast
psbE	photosystem II PsbE protein	Ccr_CDS_188	chloroplast
psbF	cytochrome b559 b subunit	Ccr_CDS_187	chloroplast
psbH	photosystem II reaction center protein H	Ccr_CDS_65	chloroplast
psbI	photosystem II PsbI protein	Ccr_CDS_196	chloroplast
psbJ	photosystem II PsbJ protein	Ccr_CDS_185	chloroplast
psbK	photosystem II protein K	Ccr_CDS_102	chloroplast
psbL	photosystem II PsbL protein	Ccr_CDS_186	chloroplast
psbM	photosystem II PsbM protein	NA	nuclear
psbN	photosystem II reaction center protein N	Ccr_CDS_66	chloroplast
psbO	photosystem II Mn-stabilizing polypeptide	GSCHCT00004801001	nuclear
psbP	photosystem II oxygen-evolving complex	GIDCcT00003465001	nuclear
psbQ	oxygen-evolving enhancer protein 3	GSCHCT00002109001	nuclear
psbR	photosystem II PsbR protein	-	-
psbS	photosystem II PsbS protein	-	-
psbT	photosystem II reaction center protein T	Ccr_CDS_67	chloroplast
psbU	photosystem II 12 kD extrinsic protein	GSCHCT00004799001	nuclear
psbV	cytochrome c550	Ccr_CDS_12	chloroplast
psbW	photosystem II 13kD protein	Ccr_CDS_43	chloroplast
psbX	photosystem II PsbX protein	Ccr_CDS_10	chloroplast
psbY	photosystem II PsbY protein	-	-
psbZ	photosystem II protein Z	Ccr_CDS_101	chloroplast
psb27	photosystem II Psb27 protein	GIDCcT00029623001	nuclear
psb28	photosystem II Psb28 protein	-	-

TABLE 25: Annotation manuelle des gènes du complexe cytochrome  $b_6/f$  et des transporteurs solubles d'électrons chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.

Manual annotation of phycobilisome biosynthesis genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
petA	cytochrome f	Ccr_CDS_122	chloroplast
petB	cytochrome $b_6$	Ccr_CDS_49	chloroplast
petC	cytochrome $b_6/f$ Rieske protein	GSCHCT00014586001	nuclear
petD	cytochrome $b_6/f$ complex subunit IV	Ccr_CDS_50	chloroplast
petG	cytochrome $b_6/f$ complex subunit V	Ccr_CDS_103	chloroplast
petL	cytochrome $b_6/f$ complex subunit VI	-	-
petM	cytochrome $b_6/f$ complex subunit VII	-	-
petN	cytochrome $b_6/f$ complex subunit VIII	Ccr_CDS_112	chloroplast
petP	cytochrome $b_6/f$ complex subunit PetP	Ccr_CDS_200	chloroplast
ccsA	c-type cytochrome synthesis protein	Ccr_CDS_167	chloroplast
dsbD	cytochrome c biogenesis protein transmembrane region	Ccr_CDS_27	chloroplast
ccs1	c-type cytochrome biogenesis protein	Ccr_CDS_28	chloroplast
petE	plastocyanin	-	-
petF	ferredoxin (2Fe-2S)	Ccr_CDS_70	chloroplast
petJ	cytochrome $c_{553}$	Ccr_CDS_13	chloroplast
petH	ferredoxin-NADP oxidoreductase	GSCHCT00000556001	nuclear

TABLE 26: Annotation manuelle des gènes de structure de l'ATP synthase chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.

Manual annotation of ATP synthase structural genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
atp1	H <sup>+</sup> -transporting ATP synthase protein I	-	-
atpA	ATP synthase CF1 $\alpha$ chain	Ccr_CDS_140	chloroplast
atpB	ATP synthase CF1 $\beta$ chain	Ccr_CDS_128	chloroplast
atpC	ATP synthase $\gamma$ chain	GSCHCT00013491001	nuclear
atpD	ATP synthase $\delta$ chain	Ccr_CDS_139	chloroplast
atpE	ATP synthase $\epsilon$ chain	Ccr_CDS_127	chloroplast
atpF	ATP synthase CFO B chain subunit I	Ccr_CDS_138	chloroplast
atpG	ATP synthase CFO B' chain subunit II	Ccr_CDS_194	chloroplast
atpH	ATP synthase CFO C chain lipid-binding subunit III	Ccr_CDS_137	chloroplast
atpI	ATP synthase CFO A chain subunit IV	Ccr_CDS_136	chloroplast



TABLE 27: Annotation manuelle des gènes de biosynthèse de la chlorophylle chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.  
Manual annotation of chlorophyll biosynthesis genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
cao	chlorophyll a oxygenase (chlorophyll b synthase)	-	-
pao1	pheophorbide a oxygenase	GSCHCT00007371001	nuclear
pao2	pheophorbide a oxygenase	GSCHCT00012713001	nuclear
pao3	pheophorbide a oxygenase	GSCHCT00000454001	nuclear
rccR	red chlorophyll catabolite reductase	-	-
chlB	light-independent protochlorophyllide reductase subunit B	-	-
chlD	Mg-chelatase subunit D (Mg-protoporphyrin IX chelatase)	GSCHCT00008152001	nuclear
acsF	Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase	Ccr_CDS_111	chloroplast
chlG	chlorophyll a synthase	GSCHCT00001439001	nuclear
chlH	Mg-chelatase subunit H (Mg-protoporphyrin IX chelatase)	GSCHCT00014434001	nuclear
chlI	Mg-chelatase subunit I (Mg-protoporphyrin IX chelatase)	Ccr_CDS_190	chloroplast
chlL	light-independent protochlorophyllide reductase	-	-
chlM	Mg-protoporphyrin IX methyltransferase	none	nuclear
chlN	light-independent protochlorophyllide reductase subunit N	-	-
gluRS1	glutamyl-tRNA synthetase	GSCHCT00004319001	nuclear
gluRS2	glutamyl-tRNA synthetase	GSCHCT00011758001	nuclear
hemA	glutamyl-tRNA synthetase	GSCHCT00012161001	nuclear
hemB	$\delta$ -aminolevulinic acid dehydratase (porphobilinogen synthase)	GSCHCT00006223001	nuclear
hemC	porphobilinogen deaminase (hydroxymethylbilane synthase)	GSCHCT00006756001	nuclear
hemD	uroporphyrin III synthase	GSCHCT00013274001	nuclear
hemE1	uroporphyrinogen decarboxylase	GSCHCT00008159001	nuclear
hemE2	uroporphyrinogen decarboxylase	GSCHCT00006146001	nuclear
hemF	coproporphyrinogen III oxidase	GSCHCT00006179001	nuclear
-	coproporphyrinogen III oxidase-related protein	GSCHCT00003962001	nuclear
hemG	protoporphyrinogen oxidase	GSCHCT00001523001	nuclear
hemG	protoporphyrinogen oxidase	GSCHCT00008928001	nuclear
hemH	ferrochelatase (protoheme ferro-lyase)	GSCHCT00005290001	nuclear
hemL	glutamate-1-semialdehyde 2,1-aminomutase	GSCHCT00008188001	nuclear
ho1	heme oxygenase	GSCHCT00011450001	nuclear
porA	light-dependent NADPH-protochlorophyllide oxido-reductase	GSCHCT00014565001	nuclear
-	short-chain dehydrogenase/reductase superfamily protein	GSCHCT00004307001	nuclear
cox10	protoheme IX farnesyltransferase	GSCHCT00010276001	nuclear
upm1	uroporphyrin III methylase	GSCHCT00001730001	nuclear

TABLE 28: Annotation manuelle des gènes de biosynthèse des caroténoïdes et de fixation du carbone chez *C. crispus*. CDS : séquence codante, NA : donnée indisponible. Les gènes absents des génomes plastidiaux et nucléaires sont signalés par un tiret.

Manual annotation of carotenoid biosynthesis and carbon fixation genes in *C. crispus*. CDS : coding DNA sequence, NA : non-available data. Genes missing from both plastidial et nuclear genomes are represented by a dash.

Gene	Produit/Product	CDS code	Genome
ipp	isopentenyl-diphosphate $\delta$ -isomerase I	GSCHCT00011207001	nuclear
ggps	geranylgeranyl pyrophosphate synthase	GSCHCT00012532001	nuclear
psy	geranylgeranyl pyrophosphate synthase	GIDCcT00001702001	nuclear
zds	$\zeta$ -carotene desaturase	GSCHCT00003462001	nuclear
lcyE	lycopene $\epsilon$ cyclase	GSCHCT00007289001	nuclear
lcyB	lycopene $\beta$ cyclase	GSCHCT00004251001	nuclear
bch	$\beta$ -carotene hydroxylase	P_UA0ABB193YA02FM1	chloroplast
zep	zeaxanthin epoxidase	-	-
vde	violaxanthin deepoxidase	-	-
crtO	$\beta$ carotene ketolase	P_UA0ABA473YP09RM1	chloroplast
nced	9-cis-epoxycarotenoid dioxygenase	GSCHCT00009050001	nuclear
rbcl	ribulose biphosphate carboxylase, small chain	Ccr_CDS_152	chloroplast
rbcS	ribulose biphosphate carboxylase, large chain	Ccr_CDS_153	chloroplast

TABLE 29: Séquences des amorces utilisées en PCR quantitative pour la validation de l'expérience de RNAseq chez *C. crispus*. CDS : séquence codante, GP1-6 : marqueurs potentiels pour les gamétophytes, M1-2 : marqueurs potentiels des gamétophytes mâles, F : marqueur potentiel des gamétophytes femelles, T1-4 : marqueurs potentiels des tetrasporophytes.

Sequences of the quantitative PCR primers used for the validation of the RNA-seq experiment in *C. crispus*. CDS : coding DNA sequence, GP1-6 : potential markers for gametophytes, M1-2 : potential markers for male gametophytes, F : potential markers for female gametophytes, T1-4 : potential markers for tetrasporophytes.

amorce	5' (sens)	3' (antisens)	CDS code
GP1	GGTCGCTCCAATTTCTTTACTC	AGTCGGCTCCTACTTCCAG	GSCHC2T00004033001
GP2	GTCTACTTCCCGAACCTAAC	CTTGCTCTGTTTCACCTTCAC	GSCHC2T00017317001
GP3	TTCTTTCTTCAAAGCTTGGC	CTCGTTACCTTCTTGCTTATCAC	GSCHC2T00004034001
GP4	ATCAGGAGAAACGCCATTATTG	AGCCTTCACCTTCTTCAACG	GSCHC2T00012127001
GP5	GATTGCCGATACCTCAGTCC	TTGTCCCTGTAAACCTTGTCC	GSCHC2T00005488001
GP6	AAGCGTGGACAAGGTTGAG	GTGTCGGACCCATAGAAGAAC	GSCHC2T00012911001
M1	TCTTGTTGAATGGAGATGTCTTC	TTGGTCACAGTCGCCTTTG	GSCHC2T00009103001
M2	TCCACAGACAGGGAACAGG	CTAACTCGCCACGATGACG	GSCHC2T00006517001
F	ACGGTGTGAAATGCGAGAG	TTGTCCAGCCTTCTTGTGC	GSCHC2T00007767001
T1	GCCGTATCAATGACAACTTGG	GCGTGACCGTGTACTTCC	GSCHC2T00016872001
T2	TGATACCGAGAGACACTTTGAC	AGCGAGGACTCCACGAATC	GSCHC2T00003913001
T3	CGACTACACTGCCACCATC	CCTTGTACGACGGTATCC	GSCHC2T00006895001
T4	CCCCGCTGTGATTGTAGTAG	ACCATCATCATTGTCCATGTATC	GSCHC2T00016106001



# Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida

Jonas Collén<sup>a,b,1</sup>, Betina Porcel<sup>c,d,e</sup>, Wilfrid Carré<sup>f</sup>, Steven G. Ball<sup>g</sup>, Cristian Chaparro<sup>h</sup>, Thierry Tonon<sup>a,b</sup>, Tristan Barbeyron<sup>a,b</sup>, Gurvan Michel<sup>a,b</sup>, Benjamin Noel<sup>c</sup>, Klaus Valentin<sup>i</sup>, Marek Elias<sup>j</sup>, François Artiguenave<sup>c,d,e</sup>, Alok Arun<sup>a,b</sup>, Jean-Marc Aury<sup>c</sup>, José F. Barbosa-Neto<sup>h</sup>, John H. Bothwell<sup>k,l</sup>, François-Yves Bouget<sup>m,n</sup>, Loraine Brillet<sup>f</sup>, Francisco Cabello-Hurtado<sup>o</sup>, Salvador Capella-Gutiérrez<sup>p,q</sup>, Bénédicte Charrier<sup>a,b</sup>, Lionel Cladière<sup>a,b</sup>, J. Mark Cock<sup>a,b</sup>, Susana M. Coelho<sup>a,b</sup>, Christophe Colleoni<sup>g</sup>, Mirjam Czejek<sup>a,b</sup>, Corinne Da Silva<sup>c</sup>, Ludovic Delage<sup>a,b</sup>, France Denoeud<sup>c,d,e</sup>, Philippe Deschamps<sup>g</sup>, Simon M. Dittami<sup>a,b,r</sup>, Toni Gabaldón<sup>p,q</sup>, Claire M. M. Gachon<sup>s</sup>, Agnès Groisillier<sup>a,b</sup>, Cécile Hervé<sup>a,b</sup>, Kamel Jabbari<sup>c,d,e</sup>, Michael Katinka<sup>c,d,e</sup>, Bernard Kloareg<sup>a,b</sup>, Nathalie Kowalczyk<sup>a,b</sup>, Karine Labadie<sup>c</sup>, Catherine Leblanc<sup>a,b</sup>, Pascal J. Lopez<sup>t</sup>, Deirdre H. McLachlan<sup>k,l</sup>, Laurence Meslet-Cladière<sup>a,b</sup>, Ahmed Moustafa<sup>u,v</sup>, Zofia Nehr<sup>a,b</sup>, Pi Nyvall Collén<sup>a,b</sup>, Olivier Panaud<sup>h</sup>, Frédéric Partensky<sup>a,w</sup>, Julie Poulain<sup>c</sup>, Stefan A. Rensing<sup>x,y,z,aa</sup>, Sylvie Rousvoal<sup>a,b</sup>, Gaele Samson<sup>c</sup>, Aikaterini Symeonidi<sup>y,aa</sup>, Jean Weissenbach<sup>c,d,e</sup>, Antonios Zambounis<sup>bb,s</sup>, Patrick Wincker<sup>c,d,e</sup>, and Catherine Boyen<sup>a,b</sup>

<sup>a</sup>Université Pierre-et-Marie-Curie University of Paris VI, Station Biologique, 29680 Roscoff, France; <sup>b</sup>Centre National de la Recherche Scientifique, Station Biologique, Unité Mixte de Recherche 7139 Marine Plants and Biomolecules, 29680 Roscoff, France; <sup>c</sup>Commissariat à l'Energie Atomique, Institut de Génétique/Genoscope, 91000 Evry, France; <sup>d</sup>Centre National de la Recherche Scientifique, Unité Mixte de Recherche 8030, CP5706, 91000 Evry, France; <sup>e</sup>Université d'Evry, 91025 Evry, France; <sup>f</sup>Centre National de la Recherche Scientifique, Université Pierre-et-Marie-Curie, FR2424, ABiMS (Analysis and Bioinformatics for Marine Science), Station Biologique, 29680 Roscoff, France; <sup>g</sup>Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche 8576, Centre National de la Recherche Scientifique-Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France; <sup>h</sup>Laboratoire Génome et Développement des Plantes, Unité Mixte de Recherche, Centre National de la Recherche Scientifique/Institut de Recherche pour le Développement, Université de Perpignan Via Domitia, F-66860 Perpignan Cedex, France; <sup>i</sup>Alfred Wegener Institute for Polar and Marine Research, 27570 Bremerhaven, Germany; <sup>j</sup>Life Science Research Center, Department of Biology and Ecology, Faculty of Science, University of Ostrava, 710 00 Ostrava, Czech Republic; <sup>k</sup>School of Biological Sciences, Queen's University Belfast, Belfast BT9 7BL, United Kingdom; <sup>l</sup>Queen's University Marine Laboratory, Portaferry BT22 1PF, United Kingdom; <sup>m</sup>Observatoire Océanologique, Université Pierre-et-Marie-Curie-University of Paris VI, 66651 Banyuls-sur-mer, France; <sup>n</sup>Laboratoire d'Observatoire d'Océanographie Microbienne, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7621, 66651 Banyuls-sur-mer, France; <sup>o</sup>Mechanisms and Origin of Biodiversity Team, Unité Mixte de Recherche 6553-Ecobi, Campus de Beaulieu-Bât14A, University of Rennes1, 35042 Rennes, France; <sup>p</sup>Centre for Genomic Regulation, 08003 Barcelona, Spain; <sup>q</sup>Universitat Pompeu Fabra, 08003 Barcelona, Spain; <sup>r</sup>Program for Marine Biology, Department of Biology, University of Oslo, 0316 Oslo, Norway; <sup>s</sup>Microbial and Molecular Biology Department, Scottish Marine Institute, Scottish Association for Marine Science, Oban PA37 1QA, United Kingdom; <sup>t</sup>Département Milieux et Peuplements Aquatiques, Unité Mixte de Recherche-Biologie des Organismes et Écosystèmes Aquatiques, Centre National de la Recherche Scientifique, Muséum National d'Histoire Naturelle, Université Pierre et Marie Curie, Institut de Recherche pour le Développement 207, 75005 Paris, France; <sup>u</sup>Department of Biology and <sup>v</sup>Biotechnology Graduate Program, American University in Cairo, New Cairo, Egypt; <sup>w</sup>Oceanic Plankton Group, Station Biologique, Centre National de la Recherche Scientifique Unité Mixte de Recherche 7144, 29680 Roscoff, France; <sup>x</sup>Faculty of Biology, University of Freiburg, 79085 Freiburg, Germany; <sup>y</sup>Freiburg Initiative in Systems Biology, University of Freiburg, 79085 Freiburg, Germany, and <sup>z</sup>BIOSS Centre for Biological Signalling Studies, University of Freiburg, 79085 Freiburg, Germany; <sup>aa</sup>Faculty of Biology, University of Marburg, D-35032 Marburg, Germany; and <sup>bb</sup>Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece

Edited by Robert Haselkorn, University of Chicago, Chicago, IL, and approved February 6, 2013 (received for review December 18, 2012)

Red seaweeds are key components of coastal ecosystems and are economically important as food and as a source of gelling agents, but their genes and genomes have received little attention. Here we report the sequencing of the 105-Mbp genome of the florideophyte *Chondrus crispus* (Irish moss) and the annotation of the 9,606 genes. The genome features an unusual structure characterized by gene-dense regions surrounded by repeat-rich regions dominated by transposable elements. Despite its fairly large size, this genome shows features typical of compact genomes, e.g., on average only 0.3 introns per gene, short introns, low median distance between genes, small gene families, and no indication of large-scale genome duplication. The genome also gives insights into the metabolism of marine red algae and adaptations to the marine environment, including genes related to halogen metabolism, oxylipins, and multicellularity (microRNA processing and transcription factors). Particularly interesting are features related to carbohydrate metabolism, which include a minimalistic gene set for starch biosynthesis, the presence of cellulose synthases acquired before the primary endosymbiosis showing the polyphyly of cellulose synthesis in Archaeplastida, and cellulases absent in terrestrial plants as well as the occurrence of a mannosylglycerate synthase potentially originating from a marine bacterium. To explain the observations on genome structure and gene content, we propose an evolutionary scenario involving an ancestral red alga that was driven by early ecological forces to lose genes, introns, and intergenetic DNA; this loss was

followed by an expansion of genome size as a consequence of activity of transposable elements.

The red algae, together with the glaucophytes and the Chloroplastida, are members of the Archaeplastida, the phylogenetic group formed during the primary endosymbiosis event that gave rise to the first photosynthetic eukaryote. Red algal genomes, both plastid and nuclear, also contributed, via secondary endosymbiosis, to several other eukaryotic lineages, including

Author contributions: J.C., B.P., T.T., G.M., B.N., K.V., J.-M.A., J.H.B., F.-Y.B., J.M.C., B.K., C.L., P.W., and C.B. designed research; J.C., B.P., W.C., S.G.B., C. Chaparro, T.T., T.B., G.M., B.N., K.V., M.E., F.A., A.A., J.-M.A., J.F.B.-N., J.H.B., F.-Y.B., L.B., F.C.-H., S.C.-G., B.C., L.C., J.M.C., S.M.C., C. Colleoni, M.C., C.D.S., L.D., F.D., P.D., S.M.D., T.G., C.M.M.G., A.G., C.H., K.J., M.K., N.K., K.L., C.L., P.J.L., A.M., O.P., F.P., J.P., S.A.R., S.R., G.S., J.W., A.Z., P.W., and C.B. performed research; B.P. and W.C. contributed new reagents/analytic tools; J.C., B.P., W.C., S.G.B., C. Chaparro, T.T., T.B., G.M., B.N., K.V., M.E., F.A., A.A., J.-M.A., J.F.B.-N., J.H.B., F.-Y.B., L.B., F.C.-H., S.C.-G., B.C., L.C., J.M.C., S.M.C., C. Colleoni, M.C., C.D.S., L.D., F.D., P.D., S.M.D., T.G., C.M.M.G., A.G., C.H., K.J., M.K., B.K., N.K., K.L., C.L., P.J.L., D.H.M., L.M.-C., A.M., Z.N., P.N.C., O.P., F.P., J.P., S.A.R., G.S., A.S., J.W., A.Z., P.W., and C.B. analyzed data; and J.C., S.G.B., T.T., T.B., G.M., M.E., and C.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the EMBL database (sequence nos. [CAKH01000001](https://www.ebi.ac.uk/ena/record/CAKH01000001)–[CAKH010003241](https://www.ebi.ac.uk/ena/record/CAKH010003241)).

<sup>1</sup>To whom correspondence should be addressed. E-mail: [collen@sb-roscoff.fr](mailto:collen@sb-roscoff.fr).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1221259110/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1221259110/-DCSupplemental).

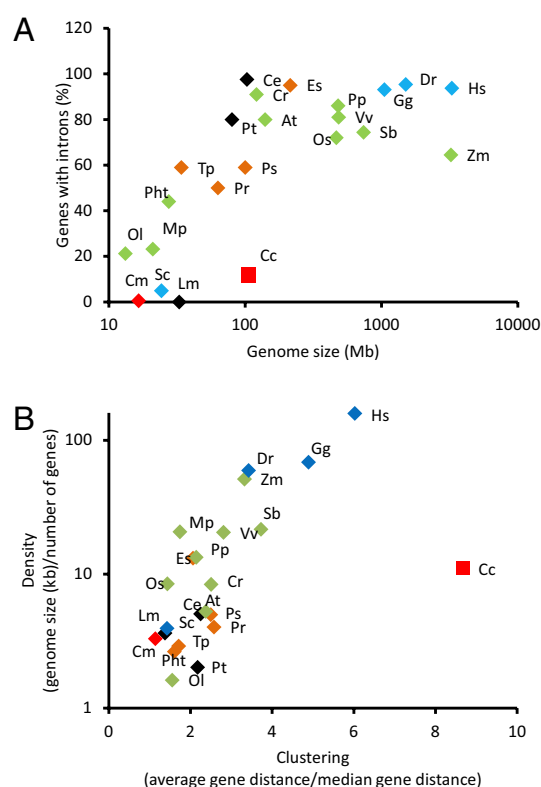
stramenopiles, alveolates, cryptophytes, and haptophytes (1), and thus genes of red algal origin are spread widely among the eukaryotes. Knowledge about red algal genes and genomes therefore is crucial for understanding eukaryote evolution. The red macroalgal fossil record stretches back 1.2 billion years, providing the oldest evidence of morphologically advanced, multicellular, sexually reproducing eukaryotes (2). Ecologically, red algae represent the most species-rich group of marine macrophytes with more than 6,000 described species ([www.algaebase.org](http://www.algaebase.org)). They are important components of many marine ecosystems, including rocky intertidal shores and coral reefs, and also are present in fresh water (3). Red algae also show some unusual physiological traits. Their photosynthetic antennae are built with phycobiliproteins, the thylakoids are unstacked, and they totally lack flagella and centrioles. In contrast to Chloroplastida, which produce starch in their chloroplasts, red algae store carbon as starch granules in their cytosol (floridean starch) (3). Their cell wall is a complex assemblage of cellulose, various hemicelluloses, and unique sulfated galactans (agars and carrageenans) (4). Economically, red macroalgae are important for their polysaccharide content. For example, carrageenans, the main sulfate-containing compounds in many red algae, are used as texturing agents and had a market value of more than US\$500 million in 2010 (5). Red algae, especially nori (*Pyropia* and *Porphyra* species), also are used directly for human consumption with a market value of ~US\$1,300 million/y (6).

A number of transcriptomic studies are available on red algae, including the genera *Porphyra*, *Chondrus*, and *Gracilaria* (see ref. 7 and references therein), which investigate developmental processes and physiological responses and establish the contribution of red algae to diverse evolutionary lineages via secondary endosymbiosis events. However, red macroalgae have been the last group of complex multicellular organisms lacking a high-quality reference genome sequence. The closest fully sequenced relative of the red macroalgae is the unicellular extremophile *Cyanidioschyzon merolae*, which has a reduced genome (8) and belongs to the Cyanidiales, a group that diverged from other red algae about 1.4 billion years ago (1).

In the present study, we analyze the genome of *Chondrus crispus* Stackhouse (Gigartinales), or Irish moss, an intertidal red seaweed, up to 20 cm long, found on rock shores in the northern Atlantic Ocean. *Chondrus* is a member of the florideophytes, the largest group of extant red algae, representing 95% of known species (3). It is a common seaweed with a typical red algal triphasic life history with easy access to all three life cycle phases: the haploid female and male gametophytes, the diploid tetrasporophyte, and the diploid carposporophyte (present on the female gametophyte). The cell wall contains carrageenan, typically with  $\iota$ - and  $\kappa$ -carrageenan in the gametophyte and  $\lambda$ -carrageenan in the sporophyte. In contrast to most other red algae, important scientific background knowledge exists for *Chondrus*, including studies on the mitochondrial genome (9), transcriptomics (10, 11), interactions with pathogens (12), effects of UV radiation (13), stress metabolism (14), and population ecology (15, 16). Thus, the availability of the *C. crispus* genome should help promote this organism as a model species for florideophyte algae and shed light on key aspects of eukaryotic evolution.

## Results and Discussion

**Reduced Genome with Exceptionally Compact Clustered Genes.** The genome sequence was obtained using DNA purified from a clonally growing unialgal culture of a gametophyte of *Chondrus crispus* and was sequenced using the Sanger technology. The assembled nuclear genome of *Chondrus* contains 1,266 scaffolds totaling 105 Mbp. A combination of expert and automatic annotation predicts 9,606 genes. The results of the annotation are described in detail in the *SI Appendix*. Genes are remarkably compact, containing only 1.32 exons on average (i.e., many fewer than other organisms of similar genome size), and most genes (88%) are monoxonic (Fig. 1A). The sparse introns are small, with an average length of 182 nucleotides (Table 1). The intron



**Fig. 1.** Structural features of the *Chondrus crispus* (Cc) genome. (A) Percentage of genes with introns as a function of genome size in selected eukaryotes. (B) Gene density as a function of clustering in selected eukaryotes. Species: *Ostreococcus lucimarinus* (Ol), *Cyanidioschyzon merolae* (Cm), *Micromonas pusilla* (Mp), *Saccharomyces cerevisiae* (Sc), *Phaeodactylum tri-cornutum* (Pht), *Leishmania major* (Lm), *Thalassiosira pseudonana* (Tp), *Phytophthora ramorum* (Pr), *Paramecium tetraurelia* (Pt), *Phytophthora sojae* (Ps), *Caenorhabditis elegans* (Ce), *Chlamydomonas reinhardtii* (Cr), *Arabidopsis thaliana* (At), *Ectocarpus siliculosus* (Es), *Oryza sativa* (Os), *Physcomitrella patens* (Pp), *Vitis vinifera* (Vv), *Sorghum bicolor* (Sb), *Gallus gallus* (Gg), *Danio rerio* (Dr), *Zea mays* (Zm), *Homo sapiens* (Hs). Green symbols indicate chloroplastids; red, rhodophytes; blue, opisthokonts; brown, stramenopiles; and black, others.

content of *Chondrus* and its distant relative *C. merolae*, as well as the limited data available on the gene structure of other red algae (17), suggest that compact genes are typical for this group and thus possibly are an ancestral trait. It is worth noting that the nucleomorphs of red algal origin in cryptomonads also have low intron content (18). Although we cannot exclude the possibility that a massive loss of introns could have occurred after the secondary endosymbiotic event, this observation suggests that the ancestral endosymbiotic red alga, which gave rise to these nucleomorphs, also had few introns. There is increasing evidence that the last eukaryotic common ancestor was intron rich and that there have been both intron losses and intron gains in the evolution of eukaryotes (19). The low number of introns in red algae thus would be a secondary feature that arose after the split between the green and red lineages about 1.5 billion years ago (1). The few introns that are present in *Chondrus* possibly have a regulatory function because, on average, transcripts for intron-containing genes accumulated to higher levels than those of monoxonic genes (*SI Appendix*, Fig. S1.1B). This result is in line with previous observations in other eukaryotes (20, 21).

Genes in *Chondrus* are clustered in gene-dense regions interspersed with sequences containing numerous repetitive elements. As a result, we observed a low median distance (0.8 kbp) between genes compared with the average distance (6.9 kbp). The ratio between average and median intergenic distances in different eukaryotes makes it clear that *Chondrus* presents an exceptionally low gene density and a high degree of clustering



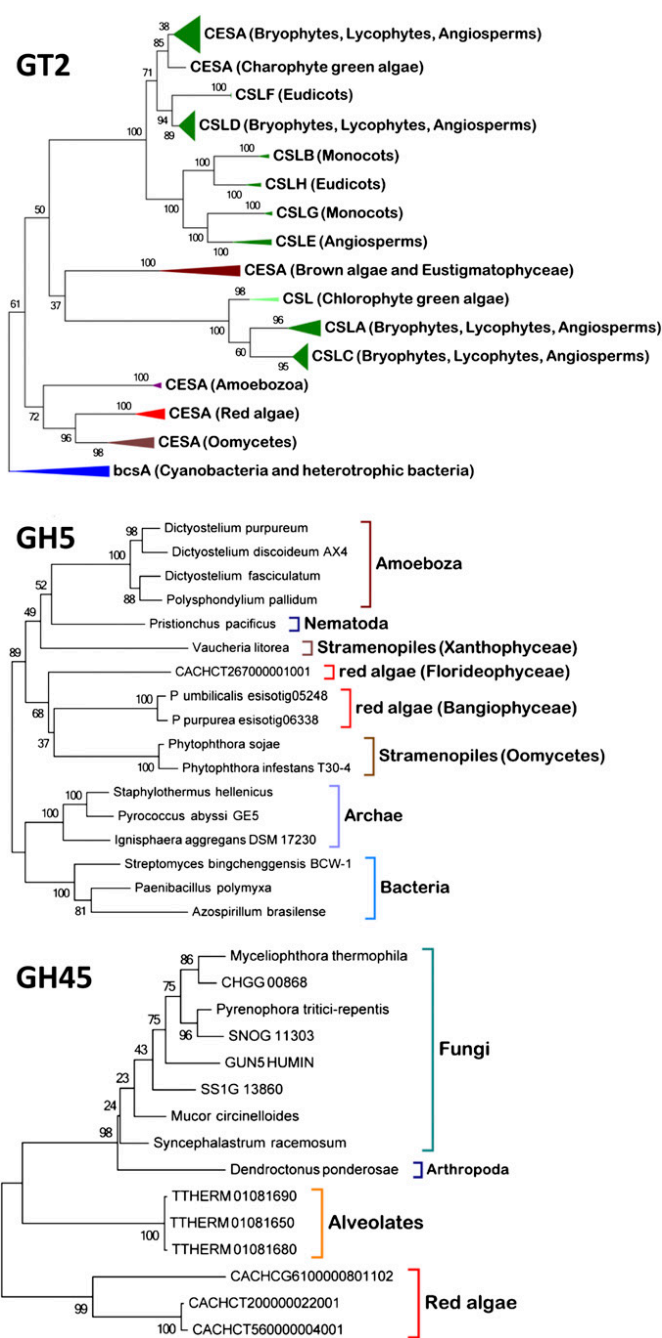


shown in Fig. 2, 57% of *Chondrus* orthology groups were not found in other red algae, demonstrating large gene diversity even within this lineage.

**Unique Carbohydrate Metabolism.** The *Chondrus* genome contains 31 glycoside hydrolases (GH) and 65 glycosyltransferases (GT) belonging to 16 GH and 27 GT families, respectively (SI Appendix, Table S7.13). These enzymes are involved in cell-wall metabolism and in the synthesis of other polysaccharides as well as protein and lipid glycosylation. *Chondrus* features all the genes needed to synthesize and recycle starch (SI Appendix, Table S7.14) but with a surprisingly low redundancy. The finding of only 12 starch-related genes revolutionizes our understanding of the building of this important polymer. Indeed, until very recently it was assumed that the complexity of starch metabolism in the green lineage reflected the complexity of the structure of starch granules. The *Chondrus* genome clearly invalidates this hypothesis.

One gene homologous to family GT7 chondroitin synthase and nine genes similar to carbohydrate sulfotransferases (CSTs) were identified. These enzymes are involved in the biosynthesis of sulfated polysaccharides, glycosaminoglycans, in animals, suggesting that their *Chondrus* homologs are involved in carrageenan biosynthesis. CSTs also are conserved in brown algae but are absent in available genomes of terrestrial plants. The clustering of red algal CSTs with those of animals and brown algae (SI Appendix, Figs. S7.3–S7.5) confirms that the synthesis of sulfated polysaccharides is an ancient eukaryotic capacity which has been lost by plants during the conquest of land (27). In addition, *Chondrus* possesses 12 galactose-6-sulfurylases, which are responsible for the last step of carrageenan biosynthesis and are unique to red algae (28), and three GH16 enzymes related to  $\kappa$ -carrageenases from marine bacteria (29), which putatively are involved in cell-wall expansion and recycling.

*Chondrus* contains two cellulose synthases (CESA) similar to those of the red alga *Porphyra* sp. (55% identity) and *Griffithsia monilis* (62% identity). Like the CESA from *G. monilis*, the *Chondrus* CESAs display a CBM48 in N terminus (30). In a Blast search against the NR database, the closest homologs of red algal sequences are CESA from Oomycetes (~35% identity), from *Dictyostelium* spp. (~30% identity), and from various bacteria (~28% identity). In contrast, CESA from land plants are more distant (~20% identity). A phylogenetic analysis with the bacterial CESAs as outgroup indicates that CESA and cellulose synthase-like proteins (CSL) from Chloroplastida diverge into two unrelated clades (Fig. 3). The first clade encompasses CESA and CLSB, D, E, F, G, and H and likely derived from the single cellulose synthase of charophytes. The second clade, which includes CSLA and CSLC, originates from a CSL from chlorophytes which is not found in the transcriptomes of charophytes (31). Red algal CESA emerge together with CESA from oomycetes in a distinct cluster rooted by CESA from Amoebozoa, confirming the tendency observed in blastp searches. Therefore, the CESAs from red algae and from green algae and embryophytes have different origins. Amoebozoa were not involved in the primary plastid endosymbiosis; thus acquisition of the bacterial cellulose synthase likely occurred before the primary endosymbiosis, and is not necessarily of cyanobacterial origin (32, 33). The nature of the different ancestral bacteria involved in horizontal gene transfer (HGT) with red algae and green algae is difficult to resolve, because all bacterial cellulose synthesis A (bCsA) genes tend to cluster together in an unrooted tree. *Chondrus* lacks family GH9 cellulases, which are found in land plants. In contrast, the genome contains three other families of cellulases (GH5, GH6, and GH45), which are absent in Chloroplastida but are conserved in various bacteria and heterotrophic eukaryotes. Phylogenetic analyses confirm that the GH5 cellulases emerge in a clade encompassing cellulases from oomycetes, Amoebozoa, and Nematoda, whereas red algal GH45 cellulases are related to cellulases from fungi (Fig. 3). GH6 cellulases from *Chondrus* are conserved both in bacteria and fungi but seem closer to bacterial GH6



**Fig. 3.** Phylogenetic trees of the cellulose synthases CESA and cellulose synthase-like proteins CSL (family GT2) and of the cellulases of the GH5 and GH45 families. All phylogenetic trees were constructed using the maximum likelihood (ML) approach with the program MEGA 5.05 ([www.megasoftware.net](http://www.megasoftware.net)). Numbers indicate the bootstrap values in the ML analysis.

cellulases. Because at least the red algal GH5 and GH45 cellulases share common ancestors with cellulases from opisthokonts or Amoebozoa, these proteins are ancient eukaryotic enzymes predating the primary plastid endosymbiosis. Thus, these ancestral cellulases probably were involved initially in the degradation of bacterial cellulose. After the acquisition of the cellulose biosynthetic pathway, these red algal enzymes likely evolved to participate in cell-wall remodeling.

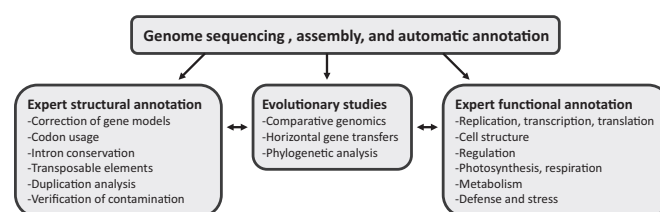
**Unusual Metabolic Features.** Because of their evolutionary history and their habitat, red algae feature some uncommon enzymes related to primary and secondary metabolism. As an illustration,



the *Chondrus* genome contains a gene similar to the mannosylglycerate synthase (MGS) from the marine bacterium *Rhodothermus marinus* (48% identity) and from some Archaea (~29% identity). This family GT78 enzyme synthesizes mannosylglycerate, an osmolyte required for thermal adaptation in thermophilic microorganisms (34, 35). This rare compound is known in red algae as “digeneaside” and accumulates during photosynthesis (36). MGS are not found in the available genomes of glaucophytes, green algae, or land plants, with the exception of *Physcomitrella patens* and *Selaginella moellendorffii*. Nonetheless, we have identified GT78 homologs in transcriptomic data of five other red algae and five streptophyte algae. A phylogenetic analysis indicates that the GT78 sequences from red algae, streptophytes, mosses, and lycophytes constitute two distinct clades, rooted by the MGS from *R. marinus* (*SI Appendix, Fig. S7.2*). Thus, there was a lateral transfer between a common ancestor of green and red algae with a thermophilic marine bacterium. Most extant red algae retained this enzyme; in the green lineage this gene was lost early by chlorophytes, but it was conserved by streptophytes, mosses, and lycophytes. It finally was lost by land plants after the divergence from lycophytes.

Several sets of anabolic and catabolic reactions previously considered specific to plants or animals were found in *Chondrus*. This result raises intriguing questions about the biological roles and the regulation of the related genes, molecules, and metabolic pathways, in particular whether their functions and mechanisms of action are conserved across different lineages. Examples are the C18 (plant-like) and C20 (animal-like) oxylipins and related compounds that have been identified in *Chondrus* in the context of studies on biotic stress response (28, 29). Interestingly, only two genes encoding lipoxygenase (*SI Appendix, Table S7.10*) have been identified, a surprising result given the diversity of oxylipins observed in this alga. The presence of methyl jasmonate, a plant hormone involved in stress signaling, has been detected in vitro after incubation with linolenic acid (37, 38). However, no candidates for allene oxide synthase, allene oxide cyclase, or jasmonic acid carboxyl methyltransferase were found. This outcome indicates that methyl jasmonate and oxylipin synthesis in *Chondrus* may be carried out by enzymes other than the ones characterized so far.

Despite the overall reduced genome, a number of gene families have remained diverse or were subject to recent diversification and expansion. One example of this diversity is the comparatively large set of genes related to halogen metabolism. Halogens play an important role in the metabolism of marine red algae (9), and transcriptomic data indicate that the corresponding genes are highly expressed (*SI Appendix, Table S8.1*). For example, 20 genes encoding animal-like heme peroxidase homologs were identified (*SI Appendix, Fig. S8.4*). In mammals these genes play a major role during pathogen ingress, releasing hypohalous acids (39), but their

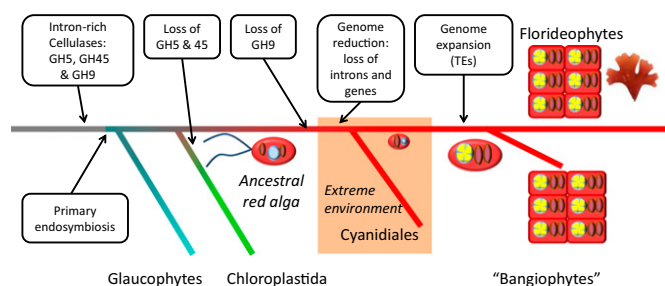


**Fig. 5.** Snapshot of the *C. crispus* genome analysis and an outline of the contents of the *SI Appendix*.

function in red algae is unknown. To our knowledge, animals and marine bacteria are the only groups of organisms in which this type of protein has been found. Their occurrence in *Chondrus* provides additional evidence for the hypothesis that proteins from the peroxidase-cyclooxygenase superfamily (such as heme peroxidases) have an ancient origin (40). In addition, the *Chondrus* genome encodes 15 members of the phosphatidic acid phosphatase type 2-haloperoxidase family. Interestingly, it also harbors a group of haloalkane dehalogenase and haloacid dehalogenase enzymes, which remove halogens from alkanes. This group of enzymes previously has been found only in prokaryotes and in the brown seaweed *E. siliculosus* (41). The large size of these halogen-related gene families likely is a specific evolutionary adaptation to the marine environment, allowing brown and red macroalgae to take benefit of halide chemistry and to modulate finely halogen metabolism, which plays an important role in defense reactions, redox reactions, and the production of secondary metabolites. Supporting this hypothesis are the facts that *E. siliculosus* has a similarly rich repertoire of peroxidases and haloperoxidases, with ~16 representatives (41), and that brown algae in general have an active halogen metabolism (42).

**Evolutionary Scenario.** The *Chondrus* genome sheds lights on the early evolution of Archaeplastida. The presence of cellulase families GH5 and GH45 in *Chondrus* supports the notion that the ancestor of the Archaeplastida was a cellulolytic protist feeding on bacterial exopolysaccharides such as cellulose. This hypothesis is consistent with the ancient eukaryotic origin of family GH9 cellulases (43). After their divergence, red algae only kept GH5 and GH45 cellulases, and green algae and plants lost these genes and conserved GH9 cellulases. Repeated exposition to bacterial genomic DNA also could explain the HGTs of various bacterial origins found in the *Chondrus* genome (e.g., GT2, GT78). Cellulose biosynthesis was acquired independently in red algae and green algae; independent acquisition could partially explain the structural diversity of cellulose-synthesizing enzyme complexes and cellulose microfibrils in Archaeplastida (44).

The compact structure of the nonrepetitive part of the *Chondrus* genome and genes also indicates that the red algal lineage went through an evolutionary bottleneck (Fig. 4). Early in the evolution of red algae, but after their divergence from green algae, selective pressure for small physical size or low nutrient requirements probably caused a reduction of the genome, with loss of introns and intergenic material. This bottleneck also could explain the lack of flagella in all life-cycle stages in red algae, because the corresponding genes may have been lost during the genome compaction. It has been suggested previously (45) that, because of the limited low pH tolerance of cyanobacteria, early eukaryotic algae would have had less competition in acidic environments where fewer photosynthetic organisms were present. The extant red algae *C. merolae* or *Galdieria sulphuraria* live in an environment with high temperature and low pH, and even though it is not obvious why such conditions would reduce genome size, it is clear that these conditions favor compact genomes in red algae and may indicate that the ancestral red alga was an acido- and thermophilic organism. The evolutionary bottleneck also might explain the high number of orphan genes in the genome, because red algae were forced to reinvent gene functions that were lost



**Fig. 4.** Proposed scenario for the evolution of red algae. An ancestor with flagella and an intron-rich genome invaded an extreme environment, possibly acidic and high temperature, with a strong selection pressure toward a reduced genome, where a genome reduction took place. The red algae later recolonized the marine and freshwater environments and experienced an expansion of the genome through the activity of transposable elements. They now are represented by the florideophytes and the bangiaophytes (red algae that are neither Cyanidiales nor florideophytes). Red ovals represent plastids; light blue circles, nucleus with ancestral genes; yellow, transposable elements.

during the genome reduction. If this hypothesis is correct, we predict that the ongoing red algal genome projects on *Porphyra* spp. (6), *P. cruentum* (46), and *C. tuberculatum* (47) will show similar gene and genome organization.

In conclusion, this study presents a reference genome for a multicellular red alga and provides a number of unexpected insights into the origin and evolution of this ancestral plant lineage. It also provides fundamental data on the unique metabolic pathways of this large and economically important group of marine algae. In addition, because of its unique genome characteristics, *C. crispus* constitutes a novel model species for studying the complex evolutionary forces that shape eukaryotic genomes. Finally, as an archive of the gene content of ancestral marine plants, this genome will help comparatively delineate the innovations that were necessary for the emergence of land plants and their adaptation to the terrestrial environments.

## Materials and Methods

A gametophyte of *C. crispus* Stackhouse (Gigartinales) was collected at Peggy's Cove, Nova Scotia, Canada (44°29'31"N, 63°55'11"W) in 1985 by Juan Correa and since then has been growing vegetatively in unialgal culture.

1. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol* 21(5):809–818.
2. Butterfield NJ (2000) *Bangiomorpha pubescens* n. gen., n. sp.: Implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* 26(3):386–404.
3. Woelkerling WJ (1990) *Biology of the Red Algae*, eds Cole KM, Sheath RG (Cambridge Univ Press, Cambridge, UK), pp 1–6.
4. Popper ZA, et al. (2011) Evolution and diversity of plant cell walls: From algae to flowering plants. *Annu Rev Plant Biol* 62:567–590.
5. Bixler HJ, Porse H (2010) A decade of change in the seaweed hydrocolloids industry. *J Appl Phycol* 23(3):321–335.
6. Blouin NA, Brodie JA, Grossman AC, Xu P, Brawley SH (2011) *Porphyra*: A marine crop shaped by stress. *Trends Plant Sci* 16(1):29–37.
7. Chan C, et al. (2012) *Porphyra* (Bangioophyceae) transcriptomes provide insights into red algal development and metabolism. *J Phycol* 48(6):1328–1342.
8. Matsuzaki M, et al. (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* 428(6983):653–657.
9. Leblanc C, et al. (1995) Complete sequence of the mitochondrial DNA of the rhodophyte *Chondrus crispus* (Gigartinales). Gene content and genome organization. *J Mol Biol* 250(4):484–495.
10. Collén J, Hervé C, Guisne-Marsollier I, Léger JJ, Boyen C (2006) Expression profiling of *Chondrus crispus* (Rhodophyta) after exposure to methyl jasmonate. *J Exp Bot* 57(14):3869–3881.
11. Collén J, Guisne-Marsollier I, Léger JJ, Boyen C (2007) Response of the transcriptome of the intertidal red seaweed *Chondrus crispus* to controlled and natural stresses. *New Phytol* 176(1):45–55.
12. Bouarab K, Potin P, Correa J, Kloareg B (1999) Sulfated oligosaccharides mediate the interaction between a marine red alga and its green algal pathogenic endophyte. *Plant Cell* 11(9):1635–1650.
13. Kräbs G, Watanabe M, Wiencke C (2004) A monochromatic action spectrum for the photoinduction of the UV-absorbing mycosporine-like amino acid shinorine in the red alga *Chondrus crispus*. *Photochem Photobiol* 79(6):515–519.
14. Collén J, Davison I (1999) Stress tolerance and reactive oxygen metabolism in the intertidal red seaweeds *Mastocarpus stellatus* and *Chondrus crispus*. *Plant Cell Environ* 22(9):1143–1151.
15. Krueger-Hadfield SA, Collén J, Daguin-Thiébaud C, Valero M (2011) Genetic population structure and mating system in *Chondrus crispus* (Rhodophyta). *J Phycol* 47(3):440–450.
16. Wang X, et al. (2008) Inter-simple sequence repeat (ISSR) analysis of genetic variation of *Chondrus crispus* populations from North Atlantic. *Aquat Bot* 88(2):154–159.
17. Hoef-Emden K, et al. (2005) Actin phylogeny and intron distribution in bangiophyte red algae (Rhodophyta). *J Mol Evol* 61(3):360–371.
18. Moore CE, Archibald JM (2009) Nucleomorph genomes. *Annu Rev Genet* 43:251–264.
19. Csuros M, Rogozin IB, Koonin EV (2011) A detailed history of intron-rich eukaryotic ancestors inferred from a global survey of 100 complete genomes. *PLoS Comput Biol* 7(9):e1002150.
20. Lanier W, Moustafa A, Bhattacharya D, Comeran JM (2008) EST analysis of *Ostreococcus lucimarinus*, the most compact eukaryotic genome, shows an excess of introns in highly expressed genes. *PLoS ONE* 3(5):e2171.
21. Shabalina SA, et al. (2010) Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Mol Biol Evol* 27(8):1745–1749.
22. Kapraun DF (2005) Nuclear DNA content estimates in multicellular green, red and brown algae: Phylogenetic considerations. *Ann Bot (Lond)* 95(1):7–44.
23. Lang D, et al. (2010) Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: A timeline of loss, gain, expansion, and correlation with complexity. *Genome Biol Evol* 2:488–503.
24. Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457(7228):405–412.
25. Nozaki H, et al. (2007) A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*. *BMC Biol* 5:28.
26. Chan CX, et al. (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Curr Biol* 21(4):328–333.
27. Michel G, Tonon T, Scornet D, Cock JM, Kloareg B (2010) The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytol* 188(1):82–97.
28. Genicot-Joncour S, et al. (2009) The cyclization of the 3,6-anhydro-galactose ring of iota-carrageenan is catalyzed by two D-galactose-2,6-sulfurylases in the red alga *Chondrus crispus*. *Plant Physiol* 151(3):1609–1616.
29. Michel G, et al. (2001) The kappa-carrageenase of *P. carrageenovor* features a tunnel-shaped active site: A novel insight in the evolution of Clan-B glycoside hydrolases. *Structure* 9(6):513–525.
30. Matthews PR, Schindler M, Howles P, Arioli T, Williamson RE (2010) A CESA from *Griffithsia monilis* (Rhodophyta, Florideophyceae) has a family 48 carbohydrate-binding module. *J Exp Bot* 61(15):4461–4468.
31. Timme RE, Bachvaroff TR, Delwiche CF (2012) Broad phylogenomic sampling and the sister lineage of land plants. *PLoS ONE* 7(1):e29696.
32. Nobles DR, Romanovicz DK, Brown RM, Jr. (2001) Cellulose in cyanobacteria. Origin of vascular plant cellulose synthase? *Plant Physiol* 127(2):529–542.
33. Nobles DR, Brown RM (2004) The pivotal role of cyanobacteria in the evolution of cellulose synthases and cellulose synthase-like proteins. *Cellulose* 11(3–4):437–448.
34. Martins LO, et al. (1999) Biosynthesis of mannosylglycerate in the thermophilic bacterium *Rhodothermus marinus*. Biochemical and genetic characterization of a mannosylglycerate synthase. *J Biol Chem* 274(50):35407–35414.
35. Empadinhas N, da Costa MS (2011) Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. *Environ Microbiol* 13(8):2056–2077.
36. Kremer B (1980) Taxonomic implications of algal photoassimilate patterns. *Br Phycol J* 15(4):399–409.
37. Bouarab K, et al. (2004) The innate immunity of a marine red alga involves oxylipins from both the eicosanoid and octadecanoid pathways. *Plant Physiol* 135(3):1838–1848.
38. Gaquerel E, et al. (2007) Evidence for oxylipin synthesis and induction of a new polyunsaturated fatty acid hydroxylase activity in *Chondrus crispus* in response to methyljasmonate. *Biochim Biophys Acta* 1771(5):565–575.
39. Davies MJ, Hawkins CL, Pattison DJ, Rees MD (2008) Mammalian heme peroxidases: From molecular mechanisms to health implications. *Antioxid Redox Signal* 10(7):1199–1234.
40. Bernrothner M, Zamocky M, Furtmüller PG, Peschek GA, Obinger C (2009) Occurrence, phylogeny, structure, and function of catalases and peroxidases in cyanobacteria. *J Exp Bot* 60(2):423–440.
41. Cock JM, et al. (2010) The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature* 465(7298):617–621.
42. La Barre S, Potin P, Leblanc C, Delage L (2010) The halogenated metabolism of brown algae (Phaeophyta), its biological importance and its environmental significance. *Mar Drugs* 8(4):988–1010.
43. Davison A, Blaxter M (2005) Ancient origin of glycosyl hydrolase family 9 cellulase genes. *Mol Biol Evol* 22(5):1273–1284.
44. Tsekos I (2002) The sites of cellulose synthesis in algae: Diversity and evolution of cellulose-synthesizing enzyme complexes. *J Phycol* 65(4):635–655.
45. Brock TD (1973) Lower pH limit for the existence of blue-green algae: Evolutionary and ecological implications. *Science* 179(4072):480–483.
46. Bhattacharya D, Price D, Yoon HS, Rajah V, Zaeuner S (2012) Sequencing and analysis of the *Porphyridium cruentum* genome. *J Phycol* 48(s1):55.
47. Chan C, Martone P (2012) Recent advances in the calliarthron genome: Climate responses and cell wall evolution. *J Phycol* 48(s1):51.
48. Batzoglu S, et al. (2002) ARACHNE: A whole-genome shotgun assembler. *Genome Res* 12(1):177–189.

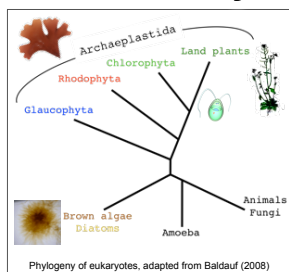


# Photosynthesis in the red alga *Chondrus crispus* back to the spill-over

Nathalie KOWALCZYK<sup>1</sup>, Pierre JOLIOT<sup>2</sup>, Fabrice RAPPAPORT<sup>2</sup>, Jonas COLLEN<sup>1</sup>, Catherine BOYEN<sup>1</sup> & Francis-André WOLLMAN<sup>1</sup>

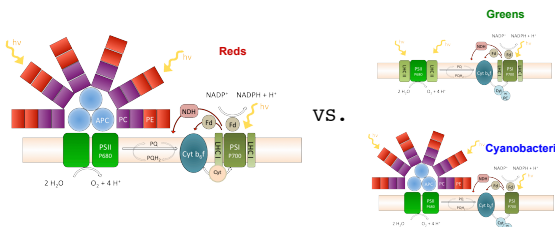
<sup>1</sup>CNRS / UPMC - UMR 7139 - Station Biologique de Roscoff - 29680 Roscoff, Breizh. Team: Stress Abiotique & Génomique Fonctionnelle  
<sup>2</sup>CNRS / UPMC - UMR 7141 - Institut de Biologie Physico-Chimique - 75005 Paris, France. Team: Physiologie Membranaire et Moléculaire du Chloroplaste

## Concerning red algae...



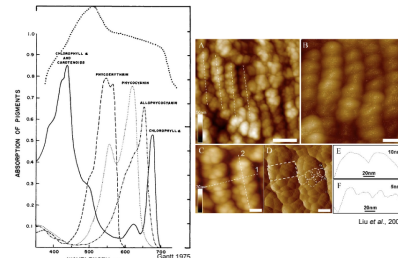
*Chondrus crispus* (Gigartinales) is an intertidal red macroalga, abundant on rocky shores of the North Atlantic. Due to its ecological and economical importance (as a carrageenan source), it has a long research history. Chosen as a model for multicellular rhodophytes, *Chondrus* is the first multicellular red alga with a sequenced genome.

## Photosynthetic machinery



Red algae possess a unique thylakoidal organisation<sup>[1]</sup> : like cyanobacteria, their photosystems (PS) are evenly distributed and they have phycobilisomes (PBS) as PSII antennae. Like green organisms, they have a chlorophyll-containing light harvesting complex as PSI antenna.

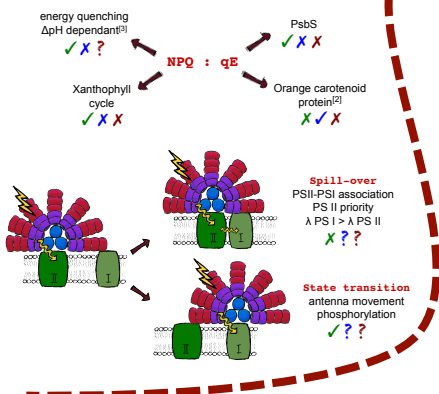
## Phycobilisomes



PBSs are big extrinsic proteic complexes, tightly organized on the membrane, containing red and blue pigments able to absorb green to orange light, filling chlorophyll's gap.

## Managing light

There is a lack of information concerning red algal processes for light distribution.



## BIOLOGICAL QUESTION

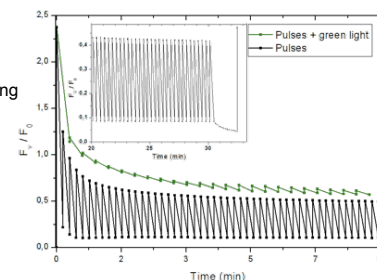
What are the mechanisms of light distribution in red algal photosynthetic chains?

### Methods

Chlorophyll fluorescence and P700 oxidation course using a JTS spectrometer.  
 $F_0$ ,  $F_M$ ,  $F_S$ ,  $F_V$ : initial, maximal, steady and variable fluorescence.

Induction of a fluorescence quenching by 200 ms saturating pulses separated by 20 s of dark = no  $\Delta pH$  build-up.  
 + no stronger effect with continuous light = no qE-type quenching  
 $F_S$  stays higher than  $F_0$  = no antenna movement<sup>[4]</sup>

## 1. a pulse-induced quenching



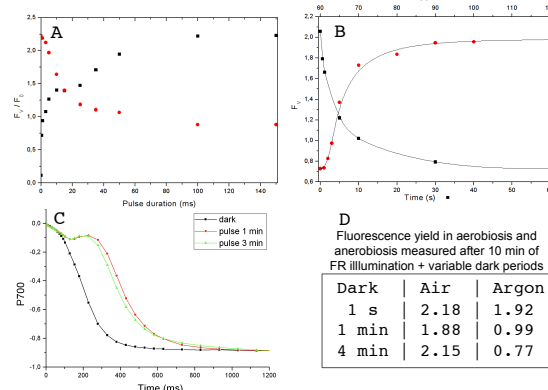
## 2. controlled by the redox state of the Plastocyanin (PQ) pool

A. 80% of the quenching occurs after a 40 ms pulse, which is enough to reduce all the PQ pool.

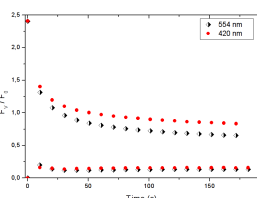
B. a 200 ms pulse + 1 min of dark leads to a maximum effect and is reversible in 1 min of far red (FR) illumination.

C. P700 FR-induced oxidation kinetics is delayed after the pulse, revealing a fully reduced PQ pool.

D. the reduction of the PQ pool by anoxygenic in the dark also triggers the quenching.

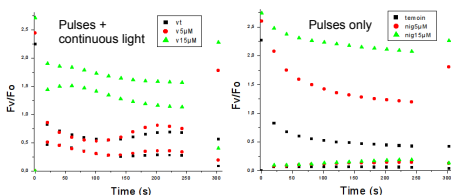


## 3. connecting PSII and PSI : the spill-over



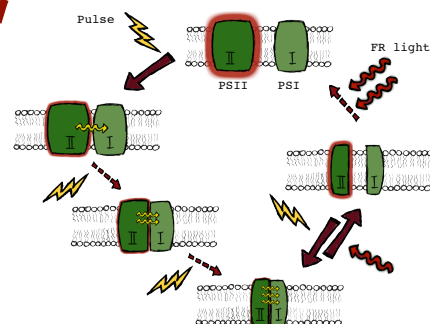
The fluorescence quenching can be seen by detecting through PBS (green light) and chlorophyll (blue light). Inconsistent with a state transition antenna movement, it supports the spill-over model.

Adding the uncoupler nigericin partly inhibits the quenching, without suppressing it. The luminal pH (~6<sup>[5]</sup>) in the dark seems to be required for the spill-over to develop.

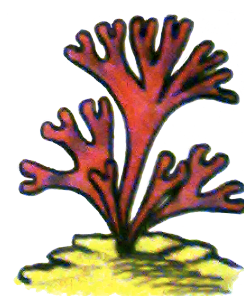


1. Su HN, Xie BB, Zhang XY, Zhou BC, Zhang YZ (2010) The supramolecular architecture, function, and regulation of thylakoid membranes in red algae: an overview. *Photosynth Res* 106:73-87.
2. Kirovsky D & Kerfeld CA (2012) The orange carotenoid protein in photosystem II in cyanobacteria. *Biochim Biophys Acta* 1817:158-168.
3. Delphin E, Duval JC, Etienne AL, Kirovsky D (1998) Delta pH-dependent photosystem II fluorescence quenching induced by saturating, multibeam pulses in red algae. *Plant Physiol* 118:103-113.
4. Fork DC & Ann KS (1988) The control by state transitions of the distribution of excitation energy in photosynthesis. *Rev Plant Physiol* 37:335-361.
5. Finazzi G & Rappaport F (1998) In vivo characterization of the electrochemical proton gradient generated in darkness in green algae and its kinetic effects on cytochrome b6/f turnover. *Biochemistry* 37:9999-10005.

## Conclusion & hypothesis



The spill-over characteristics :  
 - fast and sensitive, triggered by a 200 ms pulse  
 - direct control by the PQ pool redox state, like state transitions, but probably no phosphorylation cascade (too fast).  
 - complex inducing/relaxing kinetics, possibly explained by structural changes  
 - minimum  $\Delta pH$  requirement, but no dependence on the light-induced  $\Delta pH$ .







*Chondrus crispus* est une macroalgue rouge, choisie comme modèle d'étude des Rhodophytes multicellulaires. Il est admis que la réponse de *C. crispus* aux agents stressants d'origine abiotique relatifs à son environnement intertidal ; lumière, dessiccation, chaleur, osmolarité, pluie ; explique en grande partie sa distribution dans l'estran. Le but de ma thèse est de tester cette hypothèse en utilisant des approches multi-échelles. L'accent est mis notamment sur l'étude du métabolisme des espèces réactives de l'oxygène (ROS) considéré comme particulièrement important dans la tolérance au stress, ainsi que la photosynthèse, principal producteur de ROS et indicateur majeur de l'état physiologique des végétaux.

Le premier objectif était une étude génomique globale. D'abord, l'annotation manuelle du génome de *C. crispus*, afin de mieux comprendre le fonctionnement du métabolisme lié au stress et la photosynthèse. Puis, en utilisant les technologies de séquençage à haut débit, nous avons mené des expériences de transcriptomique afin de caractériser l'expression des gènes des trois stades de vie majeurs de *C. crispus* et identifier de nouveaux candidats impliqués dans la réponse au stress, menant à de futures études plus ciblées.

Le second objectif de cette thèse était d'analyser la réponse de l'appareil photosynthétique à la lumière, en utilisant la fluorescence de la chlorophylle et la spectrométrie, des techniques couramment utilisées chez les végétaux verts et les cyanobactéries. Cependant, les algues rouges possèdent un appareil photosynthétique unique, intermédiaire entre les plantes vertes - avec une antenne PSI similaire - et les cyanobactéries, avec des phycobilisomes en guise d'antenne PSII ; les mécanismes de dissipation d'énergie grâce aux caroténoïdes sont absents. Nous avons dû établir un protocole approprié afin d'interpréter les données générées, et ainsi caractériser un phénomène de photoprotection original, le *spill-over*, un transfert d'énergie efficace résultant d'un contact physique entre le PSI et le PSII, contrôlé par l'état redox du pool de plastoquinones.

---

*Chondrus crispus* is a red macroalga, chosen as a model for multicellular rhodophytes. It is believed that the response of *C. crispus* to abiotic stressors related to its tidal environment ; light, desiccation, heat, rainfall ; explains to a large extent its distribution on the shore. The aim of this thesis was to test this hypothesis using a multiscale approach. Two foci are the metabolism of reactive oxygen species (ROS), thought to be important in stress tolerance, and photosynthesis, the principal ROS producer and major indicator of physiological status.

The first goal was a global genomic study. With the expert annotation of the *C. crispus* genome, for a better understanding of the functioning of the stress-related metabolism and photosynthesis. Then using next generation sequencing technologies to carry out transcriptomic experiments in order to characterize the gene expression of the three major life stages of *C. crispus* and identify new candidates implicated in the stress response, leading to future targeted studies.

The second goal of this thesis was to analyze the response of the photosynthetic apparatus to light, using chlorophyll fluorescence and spectrometry. These technics are often used in green organisms and cyanobacteria, however, red algae possess a unique photosynthetic apparatus, intermediate between the green plants - with a similar PS I antenna - and the cyanobacteria, having phycobilisomes as PS II antenna ; they also lack known mechanisms of energy dissipation using carotenoids. We had to establish a suited protocol to interpret the generated data, and thus characterize an original and controverted photoprotection phenomenon, the *spill-over*, an efficient energy transfer resulting from the physical contact between PSII and PSI, controlled by the redox state of the plastoquinone pool.