UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CARACTÉRISATION MOLÉCULAIRE ET ÉTUDE DE FONCTION DES LIPOCALINES DE PLANTES AU COURS DE L'ACCLIMATATION AU FROID.

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE

PAR
JEAN-BENOIT FRENETTE CHARRON

JUIN 2007

REMERCIEMENTS

Je tiens à remercier le Dr **Fathey Sarhan** de m'avoir accueilli dans son laboratoire et conseillé dans la réalisation de mon travail de doctorat, tout en me laissant une grande liberté. Je le remercie de la confiance qu'il m'a accordée.

Je tiens à remercier tous les membres du jury d'avoir accepté d'y participer. Je remercie le Dr **Armand Séguin** et le Dr **Éric Rassart** d'avoir bien voulu prêter intérêt à ce travail malgré leur emploi du temps bien rempli. Merci d'avoir accepté de vous déplacer pour venir examiner ce travail. Je remercie le Dr **Normand Chevrier** de présider le jury lors de ma soutenance de thèse.

Je remercie le Dr **François Ouellet** pour sa gentillesse, son écoute et son soutien. Merci d'avoir cru en mon projet. Je tiens aussi à le remercier pour les nombreux conseils qu'il m'a donnés tout au long de cette étude et pour l'aide qu'il m'a apportée dans l'analyse des lipocalines.

Je tiens également à remercier toutes les personnes avec qui j'ai collaboré et sans qui ce travail ne serait pas ce qu'il est. Je remercie plus particulièrement le Dr **Ghislain Breton,** avec qui j'ai eu plaisir à travailler à nouveau, pour toute son aide technique et les conseils qu'il m'a donnés. Merci pour son amitié.

Je remercie tous les membres du laboratoire pour leur sympathie et leur aide, en particulier : **Ndjido Ardo Kane, Karine Tremblay, Marie Champoux, Guillaume Brault et Geneviève Major** piliers de ce laboratoire. Je vous adresse un grand merci pour toute votre aide, votre bonne humeur et tous les moments agréables passés ensemble.

Je remercie également le Fonds PAFARC de l'Université du Québec à Montréal pour m'avoir accordé une bourse. Je remercie également les organismes qui

ont financé le projet: Génome Québec, Génome Canada, et le Conseil de recherches en sciences naturelles et en génie du Canada.

Un grand merci à mes amis qui, par leur soutien et leur amitié, m'ont beaucoup aidé. Qu'ils me pardonnent de ne pas les citer mais je ne les oublie pas.

Merci à ma famille **Huguette Frenette, Denis Charron et Dave Charron** pour leur soutien et leur amour sans faille. Sans vous, je n'aurais pu y arriver.

Mes derniers remerciements vont à **Sonia Do Carmo**, pour l'équilibre qu'elle m'apporte, pour ses encouragements et pour son écoute toujours attentive de mes histoires de plantes. Tu es mon inspiration.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

CHAPITRE III

CHAPITRE IV

CHAPITRE V

ANNEXE I

LISTE DES FIGURES

CHAPITRE II

CHAPITRE III

CHAPITRE IV

CHAPITRE V

ANNEXE I

RÉSUMÉ GÉNÉRAL

L'acclimatation au froid est le processus par lequel certaines plantes, suite à une exposition à des températures légèrement au-dessus de 0°C, acquièrent une tolérance au gel qui leur permet de survivre aux conditions rigoureuses de l'hiver. Ce processus est associé à l'induction de plusieurs gènes et à l'accumulation de plusieurs protéines. Chez le blé, cette accumulation est plus importante chez les variétés d'hiver, plus tolérantes au gel, que chez les variétés de printemps. La comparaison des blés d'hiver et de printemps a, jusqu'à maintenant, permis l'identification de plusieurs gènes potentiellement importants lors de l'acquisition de la tolérance au gel. Mes travaux de doctorat portent précisément sur un groupe de gènes régulés par les basses températures et qui codent pour des protéines de la famille des lipocalines. Ces protéines se retrouvent chez les animaux vertébrés et invertébrés, plantes et bactéries. Elles sont principalement connues comme des protéines de transport et sont impliquées dans plusieurs fonctions cellulaires telles que modulation de la croissance cellulaire et du développement, liaison à des récepteurs et signalisation, et réponse aux stress environnementaux. Dans un premier temps, mon travail a permis de découvrir et caractériser les premières lipocalines de plantes, identifiées chez le blé et *Arabidopsis* et désignées TIL pour « Temperature-Induced Lipocalins ». L'analyse des séquences protéiques a démontré la présence de trois régions conservées qui caractérisent les lipocalines. De plus, cette analyse a permis de montrer une similitude entre ces protéines végétales et l'Apolipoprotéine D humaine (ApoD), la lipocaline bactérienne Blc et la lipocaline d'insecte Lazarillo. Une comparaison de la structure tertiaire des lipocalines TIL avec celle de l'ApoD suggère que ces protéines diffèrent au niveau de leur site d'attachement à la membrane et de leur site de liaison au ligand. De plus, des analyses ont démontré que les ARN messagers *TaTIL* et *AtTIL* s'accumulent au cours de l'acclimatation au froid et suite à un choc thermique. Par la suite, mes études ont permis de classifier les lipocalines de plantes et de suggérer des fonctions possibles. Pour ce faire, une approche intégrée de compilation de données de séquence, profil d'expression, analyses phylogénétiques, et prédictions bioinformatiques a été utilisée. Ceci a permis l'identification de deux nouveaux groupes de lipocalines de plantes, désignés TIL-2 et CHL (Chloroplastic Lipocalins). Deux autres groupes de protéines, soit les violaxanthine dé-époxidases (VDE) et les zéaxanthine époxidases (ZEP), possèdent une légère similarité avec les lipocalines TILs et ont été classifiées sous le nom de protéines « lipocalin-like ». L'expression des lipocalines et des protéines « lipocalin-like » de blé est régulée par différents stress abiotiques. De plus, chez le blé, cette expression est corrélée avec la capacité de la plante à développer une tolérance au gel, ce qui suggère une fonction possible au niveau de la protection contre les dommages provoqués par le gel. Ces résultats, combinés à l'analyse phylogénétique, supportent l'hypothèse que l'évolution des lipocalines est reliée à une fonction de protection cellulaire. Ceci est supporté par l'expression de lipocalines dans des organismes tolérants comme *Debaryomyce*

hansenii, *Porphyra yezoensis* et les céréales d'hiver. Finalement, mes travaux ont contribué à élucider la fonction cellulaire de la lipocaline TIL chez *Arabidopsis*. Différentes approches de type gain ou perte de fonction ainsi que des analyses de biopuces d'ADN et des mesures de capacité photosynthétique ont été réalisées. Les résultats ont démontré que des plantes n'accumulant pas *At*TIL sont très sensibles aux baisses soudaines de température et au stress oxydatif, et que ce phénotype peut être renversé lorsque l'accumulation de cette protéine est rétablie. De plus, la surexpression de *At*TIL augmente la tolérance des plantes à ces deux stress en plus de retarder la floraison et la sénescence de la plante. L'analyse de biopuces d'ADN a indiqué que l'absence de *At*TIL affecte l'expression de 66 gènes. Parmi ceux-ci, de nombreux gènes impliqués dans le contrôle de l'horloge circadienne et de la balance énergétique de la plante. Ces données suggèrent que *At*TIL affecte une voie métabolique alternative qui module le niveau d'énergie cellulaire dans le but d'accroitre la tolérance au stress oxydatif. Les travaux présentés dans cette thèse nous aident donc à mieux comprendre le rôle des lipocalines chez les plantes et à orienter la recherche sur les mécanismes moléculaires, physiologiques et biochimiques associés à l'acclimatation au froid. L'ensemble des informations recueillies nous amènent à croire que les lipocalines végétales possèdent un potentiel important dans notre stratégie globale d'amélioration des céréales dans le but de leur conférer une plus grande tolérance au gel et autres stress abiotiques.

Mots clés : acclimatation au froid; *Arabidopsis thaliana*; lipocaline; stress oxydatif; *Triticum aestivum* L.

CHAPITRE I

REVUE DE LITTÉRATURE

1.0 Introduction

Les plantes contribuent énormément à la vie sur terre. Elles affectent l'environnement global en prévenant l'érosion des sols, en étant la principale source de carbone, en conservant l'humidité et en produisant l'oxygène que l'on respire. Les plantes sont source de fibres, carburants, médicaments et aliments. L'explosion démographique que nous connaissons actuellement engendre des besoins toujours croissants envers cette ressource. Dans le dernier siècle, la modernisation des méthodes agricoles et de croisements a permis d'augmenter considérablement le niveau de production des cultures. Cependant, ces méthodes coûteuses requièrent beaucoup de ressources, de temps et d'espèces ayant des bagages génétiques très variés. La transformation génétique est depuis peu utilisée dans le but de contourner ces barrières. Cette approche est notamment utilisée dans le but de diminuer les dommages causés par le gel chez les espèces économiquement importantes.

Afin d'utiliser cette approche, il est important de comprendre les mécanismes physiologiques, biochimiques et moléculaires impliqués dans la résistance au froid et au gel. Les prochaines sections témoignent des progrès accomplis dans le but d'augmenter nos connaissances sur ces mécanismes.

1.1 Les plantes et les basses températures

Le stress au froid survient lorsque les températures se situent entre 0 et 10°C, tandis que le gel survient lorsque les températures descendent au-dessous de zéro. Le degré de tolérance au froid et au gel est très variable chez les végétaux. On retrouve

donc des plantes tolérantes au froid et au gel, sensibles au froid et au gel, et tolérantes au froid, mais sensibles au gel. Certaines plantes peuvent acquérir un degré de tolérance supérieur au froid ou au gel après une période d'exposition à des températures légèrement supérieures au point critique de sensibilité. Par exemple, le blé d'hiver tolère normalement des températures de -5°C mais peut survivre à -18°C après une période d'exposition au froid (Limin et al., 1997). Certaines espèces comme le maïs sont sensibles au froid. Malgré le fait que celui-ci peut acquérir une certaine tolérance au froid après un passage à 14°C, il lui sera impossible de devenir tolérant au gel (Prasad et al., 1994).

1.2 Effets du gel

La formation de cristaux de glace dans le milieu extracellulaire suite au gel est très dommageable pour les cellules des plantes. Ces dommages s'expliquent de la façon suivante. Le point de congélation de l'eau apoplastique (dans l'espace entre la membrane plasmique et la paroi cellulaire) est plus élevé que celui de l'eau intracellulaire. Par conséquent, l'abaissement de la température cause l'apparition de cristaux de glace d'abord à l'extérieur de la cellule. La pression de vapeur de la glace étant plus faible que celle de l'eau liquide, il en résulte un gradient de potentiel qui déplace l'eau liquide intracellulaire vers les cristaux de glace. On observe alors une déshydratation du protoplasme et l'augmentation de la concentration en solutés ainsi qu'un affaissement de la membrane plasmique (Guy, 1990). Comme la congélation mène à une diminution du volume du cytoplasme, il s'ensuit un réajustement de la dimension de la membrane plasmique par endocytose. C'est lors du dégel que surviennent la majorité des dommages à la membrane. À ce moment, l'augmentation de la surface de la membrane plasmique n'est pas aussi rapide que l'augmentation du volume cytoplasmique, l'expansion subite du cytoplasme provoque une lyse cellulaire (Steponkus, 1984).

Autre conséquence du gel, la diminution du volume cytoplasmique et la déshydratation à la surface des membranes survenant de -4°C à -10°C provoquent un rapprochement des membranes plasmiques et chloroplastiques. Étant donné les compositions lipidiques différentes de ces membranes, un mélange se produit et on assiste à la formation de vésicules lipidiques mixtes de type phase hexagonale II. Au retour des températures clémentes, ces vésicules mixtes fusionnent à nouveau avec les différentes membranes plasmiques changeant ainsi la composition et la fluidité de ces dernières. Les membranes ainsi modifiées deviennent alors très fragiles et dysfonctionnelles (Steponkus, 1984).

Lorsque les températures descendent en dessous de -10°C, des fractures de la surface cellulaire (*fracture jump lesion*) surviennent chez les plantes sensibles. L'effet de ces fractures n'est pas connu, mais comme pour les deux phénomènes précédents, ces lésions ne sont pas observables chez les plantes capables de tolérer le gel. La tolérance au gel acquise par certaines plantes permet à celles-ci de survivre à ces stress hydriques, thermiques et mécaniques tandis que les plantes sensibles meurent (Burke et al., 1976).

Pour tolérer le gel, la plante fait appel à des mécanismes inductibles qui peuvent apparaître ou disparaître au besoin. Dans la nature, c'est au cours de l'automne, alors que la plante est exposée au froid, que débute l'ensemble des changements conduisant à l'acquisition de ces mécanismes de résistance.

1.3 Effets du froid

Dans un cycle annuel, la prolongation des périodes froides et la diminution de la photopériode sont perçues comme l'avènement de l'hiver. La combinaison de ces deux phénomènes produit de multiples effets observables dans nos régions tempérées tels que la coloration et la chute des feuilles et l'accumulation de réserves sous forme de sucres. La capacité à fleurir de certaines plantes est aussi affectée par le froid. Par exemple, pour le blé d'hiver, une longue exposition à de basses températures accélère

le processus d'induction de la floraison (Simpson et al., 1999). Ce phénomène est nommé vernalisation et intègre le froid comme signal important lors de la transition de la phase végétative à la phase reproductive (Kane et al., 2006). Ces changements perceptibles témoignent de certains ajustements moléculaires et métaboliques chez la plante suite à son exposition au froid. Voici quelques-un de ces ajustements.

1.3.1 Effet du froid sur l'horloge circadienne

À l'intérieur d'un cycle de 24 heures, des baisses de température et de luminosité sont perçues comme la venue de la nuit. Certaines études ont démontré que les plantes sont capables d'intégrer le signal des basses températures comme agent synchronisateur de leur horloge circadienne (McClung, 2001). Par conséquent, les plantes ajustent le fonctionnement de leur horloge afin d'adapter de façon précise leur métabolisme aux différentes conditions conférées par leur environnement (Millar, 2004; Salomé et McClung, 2005). Il a d'ailleurs été démontré que la capacité d'ajuster le fonctionnement de l'horloge circadienne est un caractère qui confère un avantage important lors de l'évolution des espèces végétales (Dodd et al., 2005). Contrairement à la photopériode, la façon dont le froid et les différents stress sont perçus et intégrés dans le fonctionnement de l'horloge circadienne est encore méconnue. Avant d'aborder le phénomène de compensation pour la température et autres stress environnementaux de l'horloge circadienne, il est important de bien comprendre le fonctionnement de cette dernière.

1.3.1.1 Fonctionnement de l'horloge circadienne

Chez les plantes, l'horloge circadienne régule l'expression de plusieurs gènes. Ceci a pour effet d'influencer de nombreux processus biologiques notamment l'élongation de l'hypocotyle, la position des feuilles, le temps de floraison, l'expression des gènes de la machinerie photosynthétique et le rythme de fixation du

CO_2 (McClung, 2006). Les premières expériences utilisant la technologie des puces à ADN montrent qu'environ 10% de tous les ARN d'*Arabidopsis* présentent une expression qui varie significativement au cours d'une période de 24 heures (Harmer., 2000; Schaffer et al., 2001). Le gène CHLOROPHYLL A/B BINDING PROTEIN (CAB), dont l'expression est maximale au milieu de la matinée, est l'un des principaux marqueurs du cycle circadien et est largement utilisé pour caractériser les différents mutants arythmiques. L'utilisation de ce marqueur a permis l'identification de plusieurs gènes composant l'horloge circadienne chez *Arabidopsis*.

Le modèle actuel de l'horloge circadienne est majoritairement basé sur plusieurs boucles de rétrocontrôle au centre desquelles se retrouvent invariablement trois protéines : deux facteurs de transcription de type MYB, *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) et *LATE ELONGATED HYPOCOTYL* (*LHY*), et la protéine *TIMING OF CAB1 EXPRESSION* (*TOC1*, ou *PSEUDO RESPONSE REGULATOR1*, *PRR1*) (Alabadí et al., 2001; Alabadí et al., 2002). La boucle de rétrocontrôle principale implique les gènes *LHY* et *CCA1* qui sont exprimés de façon rythmique au lever du jour. Les deux protéines correspondantes s'accumulent fortement durant deux à trois heures, et répriment l'expression de *TOC1* en se fixant à son promoteur. Pendant la journée, l'accumulation des protéines CCA1/LHY diminue peu à peu, favorisant graduellement la transcription de *TOC1*. L'accumulation de la protéine TOC1 durant la nuit active, le matin suivant, la transcription des gènes *CCA1/LHY*. Cette dernière activation nécessite l'expression d'autres gènes majoritairement exprimés en fin de journée: *EARLY FLOWERING 3* (*ELF3*), *EARLY FLOWERING 4* (*ELF4*), *LUX ARRHYTHMO* (*LUX*) et *GIGANTEA* (*GI*) (Schaffer et al., 1998; Fowler et al., 1999; Doyle et al., 2002; Hazen et al., 2005).

Une seconde boucle de rétrocontrôle est proposée pour expliquer le fonctionnement de l'horloge circadienne chez *Arabidopsis*. Au cœur de cette boucle se retrouve une protéine hypothétique s'accumulant en soirée qui viendrait activer l'expression de *TOC1*. L'expression du gène codant pour cette protéine hypothétique

serait elle-même réprimée par la protéine TOC1. Cette protéine demeure inconnue, mais il s'agirait peut-être de GIGANTEA (Locke et al., 2005).

Une troisième boucle implique des gènes apparentés à *TOC1*. Ce dernier, partage de l'homologie avec 4 autres gènes, soit *APRR* (*ARABIDOPSIS PSEUDO-RESPONSE REGULATORS*) *3*, *5*, *7* et *9* (Matsushika et al., 2000). Le regroupement de ces gènes est connu sous le nom de quintette APRR. À l'aube, *APRR9* est le premier gène du quintette APRR à être activé, suivi successivement de *APRR7*, *APRR5*, *APRR3* et de *TOC1*. La durée de l'expression de chacun de ces gènes est de 2 à 3 heures (Makino et al., 2001). Cette cascade d'activation donne justement lieu à cette troisième boucle de rétrocontrôle. CCA1 et LHY régulent positivement l'expression de *APRR5*, *APRR7* et *APRR9* alors qu'à l'inverse les protéines APRR5, APRR7 et APRR9 répriment l'expression de *CCA1* et *LHY* (Farré et al., 2005; Harmer et Kay, 2005; Mizuno et Nakamichi, 2005; Nakamichi et al., 2005). Étant donné la redondance des gènes *APRR*, la perte de la fonction de l'un de ceux-ci n'a que très peu d'effet sur l'horloge. Cependant, lorsque les gènes *APRR5*, *APRR7* et APRR9 sont tous défectueux, il s'ensuit une arythmie sévère chez la plante conférant un rôle central au quintette APRR dans le fonctionnement de l'horloge (Kaczorowski et Quail, 2003; Michael et al., 2003). Cependant, la surexpression des gènes *APRR* n'a que très peu d'effet suggérant ainsi que des facteurs additionnels encore inconnus sont requis (Matsushika et al., 2002; Sato et al., 2002; Murakami et al.,2004).

Plusieurs chercheurs se trouvent déroutés par le paradoxe de l'horloge circadienne. Bien que tous les eucaryotes possèdent des mécanismes moléculaires régulant les rythmes circadiens, la composition de ces mécanismes varie d'un organisme à l'autre. Cependant, plusieurs d'entre eux sont convaincus que certains des gènes associés à l'horloge circadienne sont conservés dans l'évolution (Dunlap et al., 2004) et qu'il ne reste qu'à les trouver.

1.3.1.2 Compensation de l'horloge circadienne pour la température

Une des caractéristiques qui définit l'horloge circadienne chez tous les êtres vivants est la capacité de maintenir un rythme robuste avec une période avoisinant les 24 h sous différents régimes de température. Ce phénomène important, appelé compensation pour la température, est peu connu chez les plantes. Cependant, ce dernier est largement étudié chez la drosophile et chez *Neurospora*, où les gènes PERIOD (PER), TIMELESS (TIM) et FREQUENCY (FRQ) sont des joueurs de premier plan (Bargiello et Young, 1984; Zehring et al., 1984; McClung et al., 1989; Sehgal et al., 1994). Deux hypothèses majeures découlent de l'étude du phénomène de compensation pour la température chez ces deux organismes. La première stipule que la compensation est un caractère intrinsèque aux gènes de l'horloge circadienne. Ceux-ci, sans l'aide de facteurs externes, pourraient s'autoréguler afin de compenser pour l'écart de température (Price, 1997; Hamblen et al., 1998; Rothenfluh et al., 2000). L'autre hypothèse est que la compensation pour la température de l'horloge serait l'affaire de facteurs autres que les gènes de l'horloge (Liu et al. 1997; Majercak et al., 1999). Cette hypothèse est d'ailleurs soutenue chez les plantes par l'analyse de différents QTLs chez 27 écotypes d'*Arabidopsis* recueillis à différentes longitudes, latitudes et altitudes. Quatorze loci contribuant au phénomène de compensation pour la température ont été identifiés en comparant la capacité de ces plantes à adapter leur rythme circadien à différents régimes de température. Bien que certains QTLs aient été associés à des gènes de l'horloge (*GI*, *APRR3* et possiblement *TOC1*), la grande majorité correspondent à des gènes qui ne sont pas actuellement associés à l'horloge (Edwards et al., 2005). Chez *Arabidopsis*, le rôle de *GI* dans le phénomène de compensation pour la température a récemment été étudié (Gould et al., 2006). Il ressort de cette étude que lors de changements de température, la période durant laquelle sont exprimés les différents gènes de l'horloge (notamment *CCA1*, *LHY* et *TOC1*) demeure inchangée ou presque. Cependant, l'amplitude de la réponse (niveau d'expression) est grandement perturbée, ce qui suggère que l'amplitude serait un

caractère important de l'expression des constituants de l'horloge. De plus, il a été démontré que l'oscillateur central de plantes n'exprimant pas *GI* est grandement affecté à basse et à haute température, démontrant que GI est un gène clé dans le phénomène de compensation pour la température.

1.3.2 Le froid perturbe l'activité photosynthétique.

On reconnaît au froid la propriété de pouvoir ralentir les réactions enzymatiques. Il n'est donc pas étonnant qu'une période d'exposition au froid induise des perturbations de l'activité photosynthétique chez la plante. Les auteurs Mark Stitt et Vaughan Hurry (2002) utilisent l'exemple suivant pour illustrer ces perturbations. Durant la photosynthèse, le CO_2 est combiné avec le ribulose-1,5-biphosphate pour former le glycérate-3-phosphate, qui est par la suite réduit en triose-phosphate en utilisant du NADPH et l'ATP générés par le transport d'électrons de la photosynthèse. Le triose-phosphate est alors exporté dans le cytoplasme et converti notamment en saccharose. Pour que le taux optimal de photosynthèse soit maintenu, il est nécessaire de maintenir un équilibre parfait entre le taux de fixation du carbone et la synthèse de saccharose. Cependant, des études menées à la fin des années 80 ont démontré que suite à une exposition au froid, l'orge et l'épinard montrent rapidement des déficiences au niveau de la synthèse du saccharose (Leegood et Furbank, 1986; Stitt et al., 1988; Sharkey et al., 1986). Par conséquent, la capacité photosynthétique de ces plantes se trouve perturbée.

Il est maintenant connu que le froid entraîne des changements à l'intérieur des chloroplastes (Huner et al., 1998). Ces changements sont notamment provoqués par un débalancement entre la quantité d'énergie lumineuse absorbée par la photochimie et la quantité d'énergie utilisée par le métabolisme de la plante. Ce débalancement provoque une modulation de la pression d'excitation du photosystème II (Huner et al., 1998). La pression d'excitation est le reflet de l'état de réduction de ce photosystème. La modulation de ce paramètre par une exposition à un excès de

lumière ou aux basses températures initie une cascade de signalisation. Cette cascade de signalisation entraîne alors la modulation de l'expression de plusieurs gènes notamment ceuxassociés à la photosynthèse et à la protection contre les stress. Chez les plantes de régions tempérées comme le blé, l'exposition au froid cause une série d'ajustements au niveau de la taille des antennes, de la photochimie et de la dissipation non photochimique de l'énergie (émission de chaleur). Cette série d'ajustements protège la plante contre la photoinhibition et mène souvent à l'acquisition d'une plus grande capacité photosynthétique (Huner et al., 1998). Cependant, lorsque la situation ne peut être rétablie par ces ajustements, on assiste à une photoinhibition sévère causée par une explosion de stress oxydatif affectant les lipides membranaires, les deux photosystèmes et les enzymes du cycle de Calvin. Cette inhibition causée par le stress oxydatif peut alors devenir irréversible, se propager à l'ensemble de la cellule et entraîner la mort de la plante (Huner et al., 1998).

1.3.3 Le froid provoque le stress oxydatif

Le stress oxydatif est fortement induit par le froid ainsi que par une variété de facteurs environnementaux tels que les hautes températures, les UV, la lumière forte, les pathogènes, les herbicides et le manque d'oxygène. Le stress oxydatif, tout comme le gel, est considéré comme l'une des principales causes de pertes en agriculture (Allen et al., 1997; Mittler, 2002; Apel et Hirt, 2004; Bartels and Sunkar, 2005; Foyer et Noctor, 2005).

1.3.3.1 Production des espèces d'oxygène réactives

Par elle-même, une molécule d'oxygène est relativement peu réactive notamment grâce à sa configuration électronique (Elstner, 1987). Cependant, lorsque cette molécule d'oxygène est activée, il s'ensuit la formation de différentes espèces

d'oxygène réactives (ROS). L'excitation de (O_2) entraîne la formation de ROS comme l'oxygène singulet ($^1O_2$). Les radicaux libres superoxyde (O_2^-) et hydroxyle ($OH^·$) et les molécules comme le peroxyde d'hydrogène (H_2O_2) et l'ozone (O_3) sont quant-à eux le résultat du transfert de un, deux ou trois électrons à l'oxygène atmosphérique (Mittler 2002).

Le peroxyde d'hydrogène et l'anion superoxyde sont tous deux normalement produits par de nombreuses réactions enzymatiques catalysées par des lipooxygénases, des peroxydases, des NADPH oxydases et des xanthines oxydases. Ces ROS sont aussi formés par les chaînes de transport d'électrons du chloroplaste (Asada, 1999) de la mitochondrie et de la membrane plasmique (transport d'électrons médié par le cytochrome-*b*) (Elstner, 1987). Cependant, la chaîne de transport d'électrons des mitochondries demeure le site de production des ROS intracellulaires le plus caractérisé chez les eucaryotes. Cette chaîne produit principalement l'anion superoxyde (O_2^-) et le peroxyde d'hydrogène (H_2O_2) en laissant fuir des électrons lors du transport de ces derniers (Gille et Nohl, 2001; Chakraborti et al., 1999; Möller, 2001). Bien que la respiration mitochondriale ait un rôle à jouer lors de la surproduction de ROS chez les plantes stressées, la photorespiration et le fonctionnement de l'appareil photosynthétique en demeurent les principales sources (Mittler, 2002). La production accrue de ROS en conditions de stress résulte en de nombreux dommages cellulaires chez la plante.

1.3.3.2 Dommages causés par le stress oxydatif

Les composants cellulaires les plus susceptibles d'être endommagés par les radicaux libres sont les lipides (peroxydation des acides gras insaturés au niveau des membranes), les protéines (par dénaturation), les carbohydrates et les acides nucléiques. L'étendue des dommages liés au stress oxydatif dépend de l'état préalable des tissus, des propriétés de la membrane plasmique, de la présence d'antioxydants

endogènes et de la capacité de la plante à induire des mécanismes de protection appropriés (Mittler, 2002).

1.3.3.3 Phénomène de tolérance croisée : implication des ROS

La tolérance des plantes à différents stress peut être améliorée en augmentant les niveaux des différents antioxydants et des enzymes de détoxification (Mittler, 2002). Par exemple, des plantes surexprimant la MnSOD ont démontré une tolérance supérieure au gel, à la déshydratation, à l'application de paraquat (méthyle viologène) ainsi qu'un meilleur taux de survie suite à la période hivernale (McKersie et al., 1993, 1996, 1999; Bowler et al., 1991; Slooten et al., 1995). La surexpression de la FeSOD procure aussi une tolérance accrue au paraquat (Van Camp et al., 1996; Van Breusegem et al., 1999) et un meilleur taux de survie suite à la période hivernale (McKersie et al., 2000). La surexpression de la forme chloroplastique de la SOD, la Cu/Zn-SOD du pois chez *Arabidopsis* résulte aussi en une tolérance accrue aux stress liés à l'exposition à la lumière forte et aux basses températures (Sen Gupta et al., 1993a, 1993b).

Ces études sont un exemple du phénomène de tolérance-croisée aux différents stress environnementaux. De plus, bien que les ROS soient perçus comme des sous-produits toxiques issus du métabolisme, de plus en plus d'indices démontrent que ces molécules, lors de conditions de stress, peuvent aussi agir en tant que signaux cellulaires visant à provoquer une réponse spécifique nécessaire à l'acquisition d'une tolérance chez la plante (Mittler et al., 2004).

2.0 Acclimatation au froid et tolérance au gel

L'acclimatation au froid est le processus par lequel les plantes tolérantes acquièrent une tolérance au gel, ce qui leur permet de survivre aux conditions rigoureuses de l'hiver. Deux phases sont étudiées lors de l'acclimatation (Guy, 1990).

Premièrement, la phase d'ajustement métabolique des fonctions cellulaires de base. Celle-ci permet l'ajustement rapide de la plante aux contraintes biophysiques des basses températures. Deuxièmement, l'induction des mécanismes menant à la tolérance au gel complète l'acclimatation.

On associe également des changements physiologiques, biochimiques et moléculaires à l'acclimatation au froid et au développement de la tolérance au gel chez les espèces tolérantes. Parmi ces changements, on note l'accumulation d'osmoprotectants, l'augmentation du transport vésiculaire et la modification de la composition de la membrane plasmique, l'induction de nouveaux gènes (parfois spécifiques aux basses températures) et la production de nouvelles protéines. Toutes ces modifications sont considérées comme faisant partie du processus d'acclimatation au froid qui mène à la tolérance au gel (Guy, 1990). L'essor de la génomique au cours des dernières années a permis d'étudier et de quantifier l'importance de l'induction de nouveaux gènes lors de l'acclimatation au froid chez plusieurs espèces. Ces travaux ont permis de constater que les modifications de l'expression génétique sont sous-jacentes à l'ensemble des changements observés lors de l'acquisition de la tolérance au gel.

2.1 Accumulation d'osmoprotectants

Plusieurs études ont montré qu' à basse température, l'expression de certains gènes et l'activité de certaines enzymes augmentent dans le but de synthétiser des métabolites nécessaires à la tolérance des plantes au gel (Breton et al., 2000). Les osmoprotectants se retrouvent dans tous les organismes, des archéobactéries aux plantes supérieures et animaux (Yancey et al. 1982; McNeil et al. 1999). Dans des conditions physiologiques, ce sont des composés hautement solubles qui ne comportent aucune charge nette. Les osmoprotectants sont aussi non toxiques et ce, à haute concentration. On attribue plusieurs rôles à ce type de composés. Lorsque certaines conditions (salinité, température, humidité) ne sont pas favorables, ils

abaissent le potentiel osmotique dans le cytoplasme, empêchent la précipitation de protéines et stabilisent la membrane plasmique en produisant des interactions avec les groupements polaires des phospholipides (Bohnert et Shen 1999; McNeil et al. 1999). En remplaçant les molécules d'eau, ils permettent une diminution des interactions protéines-solvant de façon à ce que l'association des sous unités et la conformation native des protéines soient maintenues dans un milieu à faible potentiel hydrique. Les osmoprotectants sont donc des acteurs importants dans l'adaptation cellulaire aux diverses conditions environnementales hostiles à la plante (Yancey, 1994). On retrouve principalement deux types d'osmoprotectants: ceux dérivés des acides aminés et ceux dérivés des sucres.

2.1.1 Les sucres comme osmoprotectants

Durant l'acclimatation au froid, l'accumulation de composés reliés aux sucres (saccharose, sorbitol et fructans) a souvent été observée (Olien et Clark 1993; Vágújfalvi et al. 1999). Il n'est donc pas étonnant que chez la luzerne et *Arabidopsis*, l'augmentation de l'activité de certaines enzymes responsables de la formation de ces composés est corrélée avec le degré de tolérance au gel de certains cultivars ou écotypes (Castonguay et Guckert 1996, Taji et al., 2002). Les sucres préviennent la déshydratation des protéines et des membranes en les maintenant dans des conditions analogues à celles observées physiologiquement (Crowe et al. 1992).

Selon Cadieux et al. (1988), l'augmentation des sucres participe à l'augmentation de l'osmolarité intracellulaire abaissant le point de congélation et protégeant ainsi la cellule contre le gel intracellulaire. Il a aussi été proposé que les sucres exercent un effet direct sur la protection de la membrane. Les sucres réducteurs, de même que le tréhalose et les fructans préviendraient la fusion des membranes induites par la déshydratation en agissant comme un tampon entre elles (Demel et al., 1998; Vereyken et al., 2001).

Il a été démontré par notre laboratoire que chez le blé, durant l'acclimatation au froid, le contenu en sucres solubles et réducteurs est plus élevé chez les cultivars de blé tolérants au gel que chez les cultivars de blé sensibles. L'accumulation de sucres, qui sont des composés énergétiques, implique que le blé tolérant possède une balance énergétique plus élevée que son homologue sensible (Perras et Sarhan, 1984).

2.1.2 Osmoprotectants dérivés des acides aminés

Plusieurs osmoprotectants dérivés des acides aminés ont été identifiés chez les plantes soumises à des stress environnementaux. On dénombre parmi ceux-ci, la proline (Hare et al. 1999), le 3-dimethylsulfoniopropionate (McNeil et al. 1999), les polyamines (Kumar et Minocha 1998), la trigonelline (Tramontano et Jouve 1997) et la glycine bétaïne (Nuccio et al. 1999).

2.1.2.1 La proline

La proline s'accumule au cours de plusieurs stress environnementaux et est aussi corrélée avec l'induction florale et l'arrêt de croissance (Hare et al., 1999). La proline peut être synthétisée à partir du glutamate par une enzyme bifonctionnelle possédant les activités catalytiques γ-glutamyl kinase et glutamique-γ-semialdehyde déshydrogénase (Holmberg et Bülow 1998). Le gène qui code pour cette enzyme, le *p5cs*, a été utilisé pour la production de plantes transgéniques dans le but d'augmenter la production de proline (Kishor et al. 1995). Des plants de tabac transgéniques au contenu élevé en proline se sont montrés plus osmotolérants que les plants témoins durant un stress salin et une période de sécheresse. De plus, ils ont démontré une biomasse plus développée que les plants témoins. Les études portant sur l'accumulation de cet acide aminé suggèrent qu'il serait de plus une molécule de signalisation (Maggio et al., 2002 : Nanjo et al., 1999).

2.1.2.2 La glycine bétaïne

Des études ont démontré que l'expression et l'accumulation de plusieurs enzymes impliquées dans la synthèse de la glycine bétaïne sont induites par les stress environnementaux (Charron et al., 2002; McCue et Hanson., 1992; Nakamura et al., 1997; Weretilnyk et Hanson, 1990). De plus, on attribue à la glycine bétaïne plusieurs des résultats les plus spectaculaires en ce qui a trait à l'augmentation de la tolérance aux basses températures (Breton et al. 2000). Parmi ces résultats, l'application exogène de bétaïne a augmenté de 5°C la tolérance au gel des plants de blé acclimatés ou non (Allard et al. 1998). Des tests de cryoprotection ont aussi montré que la glycine bétaïne peut agir comme osmoprotectant en protégeant les liposomes, les bactéries et différents tissus durant la congélation (Higgins et al. 1987; Zhao et al. 1992; Ko et al. 1994; Lloyd et al. 1994).

2.2 Changement de la composition en lipides membranaires et exocytose

Les membranes plasmiques constituent le plus important site de changements structuraux au cours de l'acclimatation au froid. Lyons (1973) et Raison (1973) sont les premiers à avoir proposé qu'une transition de phase thermotropique des lipides membranaires puisse jouer un rôle dans la sensibilité des plantes au froid. Taiz et Zeigler (1991) expliquent qu'une augmentation de la teneur en lipides insaturés augmenterait la fluidité des membranes en diminuant la température de semicristallisation.

En comparant la composition en phophatidylglycérols (PG) de 74 espèces de plantes, un lien fut établi pour la première fois entre la composition lipidique de la membrane et le degré de tolérance au gel chez les végétaux (Roughan 1985). Il existe une corrélation positive entre le taux d'insaturation (saturé vs trans-mono-insaturé) et la sensibilité au froid. Une étude récente a permis de déterminer avec précision que la

quantité de lipides insaturés des thylakoïdes ainsi que des autres membranes augmentait au cours de l'acclimatation au froid et du développement de la tolérance au gel (Welti et al., 2002).

Des changements au niveau de la composition en lipides durant l'acclimatation du seigle ont aussi été démontrés (Uemura et Yoshida 1984; Lynch et Steponkus 1987). On note une augmentation des stérols libres, du taux d'insaturation des phospholipides et une diminution du contenu en cérébrosides. Des changements similaires ont été observés chez *Arabidopsis* (Uemura et al., 1995). Il a aussi été démontré que la fusion de protoplastes de seigle avec des liposomes de différents types de phosphatidylcholine, permettait de simuler l'acclimatation des protoplastes (Steponkus et al.,1988). Le résultat des fusions a démontré qu'une augmentation en lipides mono- et di-insaturés permettait d'imiter le comportement des protoplastes acclimatés.

Au cours de l'acclimatation au froid, le nombre de vésicules membranaires près de la membrane plasmique augmente. Ce phénomène d'exocytose a été constaté chez certains arbustes (Niki et Sakai 1983), des mousses et l'épicotyle de blé et d'orge (Singh et Laroche 1988). Ces vésicules permettraient d'altérer la composition lipidique de la membrane plasmique en fusionnant avec elle lors de l'acclimatation au froid (Singh et Laroche 1988).

2.3 Modifications génétiques reliées à l'acquisition de la tolérance au gel

L'hypothèse voulant que l'activation de nouveaux gènes et la synthèse de nouvelles protéines soient nécessaires à l'acquisition de la tolérance au gel a été proposée il y a maintenant presque 40 ans (Siminovitch et al. 1968; Weiser, 1970). Subséquemment, les travaux de Sarhan et D'Aoust (1975) ont fourni les premiers indices indiquant que l'acclimatation au froid était probablement associée à l'activation de nouveaux gènes. Ils ont observé un changement de la composition en G+C des ARNs lors de l'acclimatation d'un blé tolérant. Ce changement ne survenait

pas chez le blé sensible. Cette observation révélait donc un changement dans la population des ARNm qui pouvait être attribué à l'induction de nouveaux gènes.

Dix ans plus tard, des expériences de traduction *in vitro* ont démontré une modification de l'expression génétique chez l'épinard au cours de l'acclimatation à 5°C (Guy et al. 1985). Au même moment, d'autres données confirmèrent indirectement l'induction de nouveaux gènes lors de l'acclimatation au froid. Il a été montré que les ARN polymérases I et II sont plus actives lors de l'acclimatation au froid chez le blé (Sarhan et Chevrier, 1985). Plus tard, des études de séparation des protéines ainsi que des produits de traduction d'ARN *in vitro* par électrophorèse en deux dimensions ont clairement démontré l'étendue des changements globaux se produisant au cours de l'acclimatation au froid (Danyluk et Sarhan, 1990; Danyluk et al., 1991; Sarhan et Perras, 1987). Suite à ces observations, il a été estimé que l'expression d'environ 10 % des gènes de blé était régulée par le froid. Par la suite, il a été démontré que l'expression de certains gènes induits par le froid était corrélée avec la capacité de la plante à tolérer le gel. (Breton et al., 2003; Danyluk et al., 1994; Houde et al., 1992; Zhang et al., 1993; Frenette Charron et al. 2002; Charron et al., 2003, 2005).

Depuis que Schaffer et Fischer (1988) ont cloné les premiers ADNc induits par les basses températures, plusieurs autres ADNc induits par le froid ont été identifiés et caractérisés notamment grâce à des analyses utilisant des biopuces à ADN (Thomashow, 1999; Fowler et Thomashow, 2002; Seki et al., 2001; Seki et al., 2002). Plus récemment, chez le blé, une collection de plus de 70 000 ESTs a été générée et réunie à près de 200 000 ESTs déjà présents dans les bases de données publiques (Houde et al., 2006). L'analyse électronique de ces données d'expression a permis d'illustrer l'étendue des différents changements métaboliques provoqués par l'exposition du blé au froid et à d'autres stress abiotiques. Cette ressource constitue aussi un outil de premier plan lors d'analyse génétique chez le blé, une espèce dont le génome hexaploïde n'est pas séquencé.

Lors de travaux plus modestes, préliminaires aux travaux présentés par Houde et al., (2006), des ADNc codant pour une protéine de la famille des lipocalines ont été clonés chez le blé d'hiver et chez *Arabidopsis*. Les prochaines sections démontrent l'importance et la nécessité d'étudier plus en détail cette famille de protéines chez les plantes.

3.0 Les lipocalines

Durant les 15 dernières années, la bioinformatique et la cristallographie de protéines ont permis d'identifier une nouvelle famille de protéines, les lipocalines. Les protéines constituant cette famille ont la particularité de lier des ligands et possèdent des similarités au niveau de leur mode d'action et de leurs structures primaire, secondaire et tertiaire. (Pervaiz et Brew, 1985; Flower, 1996). Le terme "famille alpha-2μ globuline", faisant référence à un des premiers membres de cette famille, a d'abord été adopté pour désigner cette famille. Puis, en 1987, Pervaiz et Brew ont été les premiers à suggérer le nom lipocaline (dérivé des mots grecs "lipos" signifiant gras et "calyx" signifiant coupe). Cependant, Sawyer et Richardson (1991) ont soulevé la controverse en préférant utiliser le terme lipocalycine. Le terme lipocaline fait maintenant partie du vocabulaire scientifique et, sans égard pour les arguments de sémantique, sera utilisé dans le présent document.

3.1 Séquence et structure

La famille lipocaline a été définie principalement sur la base de l'homologie de séquence. A priori, les membres de cette famille présentent des séquences étonnamment divergentes. Un alignement des séquences révèle que l'homologie entre certaines lipocalines tombe fréquemment sous le seuil de l'acceptabilité, soit 20%. Cependant, toutes les lipocalines partagent, sous forme de courtes séquences conservées, des motifs qui leur confèrent des propriétés spécifiques à cette famille. La

plupart des lipocalines, de type "kernel", possèdent trois motifs distincts tandis que d'autres membres plus divergents, de type "outlier", n'en possèdent qu'un ou deux. Parmi ces trois motifs, les lipocalines possèdent indubitablement le premier. Cette caractéristique constitue le critère d'identification par excellence des membres de cette famille.

La structure tridimensionnelle des lipocalines est maintenant très bien caractérisée (Flower et al., 1993; Flower 1995; Flower 1996). Lors du repliement, on assiste à la formation d'un baril-β hautement symétrique composé de huit feuillets-β antiparallèles reliés entre eux par des ponts hydrogènes. La propriété de lier un ligand est conférée à ce baril-β par la cavité interne hydrophobe formée par le repliement des feuillets-β et par une boucle externe surplombant la partie ouverte de cette cavité. Par conséquent, c'est la diversité de cavités qui donne lieu à la liaison d'une panoplie de ligands. Chacun des huit feuillets-β, généralement identifiés de A à H, est relié au suivant par une des sept boucles, identifiées de B1 à B7, donnant lieu à la plus simple des topologies pour un baril-β. À l'exception de la première, les boucles sont de petites épingles à cheveux-β ou "β-hairpins". Contrairement aux autres, la boucle B1 est vaste et de type Ω. Elle forme le couvercle flexible qui ferme partiellement la cavité interne du baril-β. Après le feuillet-β H, on retrouve deux structures, soit une hélice alpha et un petit feuillet-β supplémentaire (I). Bien que ces deux structures soient toujours présentes, elles sont de dimensions variables et ne conservent pas la même position dans l'espace par rapport au baril-β. De plus, ces structures n'entrent pas dans la composition de la cavité hydrophobe.

Comme mentionné plus haut, les lipocalines partagent entre elles jusqu'à trois motifs. Ces motifs sont nommés: régions structurales conservées ou "structurally conserved regions (SCR)" (Flower 1995; Flower 1996). SCR1 est le motif ubiquitaire chez les lipocalines. Il est composé de l'hélice $\alpha 3_{10}$ (petite hélice plus rare et moins stable qu'une hélice α) et du feuillet-β A responsable de la fermeture rigide de l'autre extrémité du baril-β. On soupçonne le motif SCR1 de jouer un rôle prépondérant dans

la liaison du ligand. Les deux autres motifs auraient un rôle à jouer dans la stabilisation de l'interaction protéine-ligand. Cependant, ces derniers ne sont pas présents dans toutes les lipocalines. Le motif SCR2 comprend les feuillets-β F et G et la boucle unissant ces derniers, tandis que SCR3 comprend le feuillet-β H et une partie de la boucle adjacente.

3.2 Propriétés des lipocalines

La structure hautement conservée qui caractérise les lipocalines confère à cette famille de protéines une série de propriétés distinctives. On attribue aux lipocalines la capacité de lier des ligands et des récepteurs cellulaires en plus de former des complexes moléculaires. Ces propriétés sont conservées dans l'évolution et définissent maintenant le terme lipocalines.

3.2.1 Liaison de ligands

Les lipocalines sont généralement connues pour leur capacité à lier de petites molécules hydrophobes. Ces molécules peuvent être aussi bien endogènes qu'exogènes. Parmi les molécules déjà connues, certaines possèdent des fonctions biologiques critiques telles que les rétinoïdes (rétinol, acide rétinoïque), l'acide arachidonique, le fer, la bilirubine et la biliverdine, la prostaglandine, et une multitude d'hormones stéroïdiennes. La vaste sélection de molécules pouvant être liées par les lipocalines permet d'envisager un rôle général dans le transport, que ce soit pour retirer des molécules indésirables ou apporter des composés nécessaires au maintien ou à la réparation cellulaire (Åkerström et al 2000).

La capacité de lier des ligands confère un rôle au niveau de la détoxification lors de conditions de stress oxydatif à la lipocaline α-Microglobuline. Il a été démontré que cette lipocaline a la capacité de lier et de détoxifier le groupement

hème ainsi que différents produits d'oxydation générés par les ROS (Allhorn et al., 2003; Larsson et al., 2004; Allhorn et al., 2005).

3.2.2 Liaison aux récepteurs

Un nombre croissant de preuves expérimentales démontrent que les lipocalines ont la possibilité de se lier à des récepteurs spécifiques situés à la surface des cellules et d'être internalisées, notamment par endocytose. Parmi ces lipocalines, on retrouve (par ordre chronologique de la découverte de la liaison à un récepteur) la retinol-binding protein (RBP), l'α-1-Microglobuline (α-1M), la purpurine, la β-lactoglobuline (Blg), l'epididymal retinoic acid binding protein (ERABP), insectianine, l'α-1-acid glycoprotein (AGP), l'odorant-binding protein (OBP), la glycodéline, la lipocaline-1 (Lcn-1) et l'Apolipoprotéine D (ApoD) (Burke et al., 2006). Comme exemple plus spécifique, le complexe rétinol-RBP est internalisé par un mécanisme d'endocytose de récepteurs dans les cellules parenchymateuses (Senoo et al., 1990) et implique probablement un transport utilisant les cavéoles (Malaba et al. 1995). Le récepteur de la RBP a depuis été caractérisé et porte maintenant le nom de mégaline (Saito et al., 1994). Ce récepteur a comme particularité de lier un grand nombre de lipocalines y compris la RBP. Une fois internalisé dans les cellules, le complexe rétinol-RBP se dissocie, rendant ainsi le rétinol disponible pour la croissance, la différenciation cellulaire, la reproduction et la vision normale (Burke et al., 2006). Comme autre exemple, il a été démontré récemment, que l'ApoD interagit avec le récepteur de la leptine. La découverte de cette interaction a permis d'envisager un rôle physiologique pour cette lipocaline. L'ApoD contribuerait à la régulation de l'augmentation du taux de gras corporel probablement par la liaison d'un ligand, encore indéterminé, suite à la stimulation du récepteur de la leptine Ob-Rb. Ce ligand servirait de molécule signal, conférant ainsi à l'ApoD un rôle dans la signalisation cellulaire (Liu et al., 2001).

La principale hypothèse concernant la liaison des lipocalines à des récepteurs implique les trois motifs caractérisant cette famille. Lorsque la protéine est proprement repliée, les trois motifs sont à proximité l'un de l'autre formant ainsi une zone de contact à l'extrémité fermée du baril-β. Cette zone constituerait un site de liaison commun à des récepteurs de surface cellulaire (Flower et al., 1993; North, 1989). Cependant, une étude de mutagenèse a démontré que, pour l'interaction RBP-mégaline, certains résidus situés dans la boucle Ω sont aussi essentiels à la liaison protéine-récepteur (Sivaprasadarao et al., 1993). Il est donc clair que les déterminants structurels responsables de la liaison lipocaline-récepteur résident dans le repliement et la structure conservée des différents membres de cette famille.

3.2.3 Complexation macromoléculaire

La capacité que possèdent certaines lipocalines à former des complexes avec des macromolécules solubles est possiblement leur fonction de reconnaissance moléculaire la moins documentée (Flower, 1995). Dans le plasma sanguin, la RBP, transportant ou non un ligand, est généralement complexée à la protéine transthyrétine (Goodman, 1984). Seulement 4% des RBPs du plasma sont à l'état libre. La transthyrétine, en plus de lier le rétinol, a la propriété de lier des hormones thyroïdiennes dans le plasma sanguin de la plupart des vertébrés (Malpeli et al., 1996). Le complexe RBP-transthyrétine est par la suite impliqué dans l'acheminement du rétinol à certains types de cellules (Yamamoto et al., 1997). Un proche parent de la RBP, la protéine purpurine, est un composant des adhérons. Les adhérons sont de grands complexes macromoléculaires extracellulaires qui provoquent l'adhésion cellule-cellule ou cellule-substrat via une interaction avec un récepteur de surface spécifique (Schubert et LaCorbiere, 1985).

Chez l'humain, grâce à un cinquième pont disulfure, près de 80% de l'ApoD présente dans le plasma existe sous forme de complexe. Les principaux complexes se

forment avec l'apolipoprotéine AII dans les HDL et avec l'apolipoprotéine B-100 dans les LDL et VLDL (Blanco-Vaca et al., 1992). On soupçonne que ces interactions permettent de stabiliser la présence de l'ApoD dans les différents complexes de lipoprotéines pour assurer un transport plus efficace de cette lipocaline.

La boucle Ω située à l'extrémité ouverte du baril-β serait responsable des interactions protéine-protéine. Par conséquent, les variations de longueur, de forme et de composition en acides aminés seraient les principaux facteurs nécessaires à la formation de complexes macromoléculaires variés et hautement spécifiques (Monaco et Zanotti, 1992).

3.3 Les lipocalines chez les plantes

Préalablement à mes travaux, deux protéines, la violaxanthine dé-époxidase (VDE) et la zéaxanthine époxidase (ZEP), ont été identifiées comme ayant certaines caractéristiques structurales des lipocalines (Bugos et al., 1998). Bien que le nom de lipocaline leur ait été rapidement attribué, ces deux protéines suscitent la controverse quant à leur appartenance à la famille des lipocalines (Salier, 2000; Flower, communication personnelle). Comparativement aux lipocalines connues, la VDE et la ZEP, possèdent deux fois plus d'acides aminés, la position et le nombre d'introns ne sont pas conservés, et la structure secondaire à huit feuillets-β est mal définie. Malgré tout, il a été avancé que ces protéines étaient le deuxième exemple tous règnes confondus de lipocalines avec une activité enzymatique, le premier exemple étant la prostaglandine-D synthase (Bugos et al., 1998).

3.3.1 Violaxanthine dé-époxidase et zéaxanthine époxidase

La VDE et la ZEP catalysent respectivement le retrait et l'addition d'un groupement époxyde sur des caroténoïdes. Plus précisément, la violaxanthine dé-

époxidase est localisée dans la lumière des thylakoïdes et catalyse les dé-époxidations de la violaxanthine en anthéraxanthine et de l'anthéraxanthine en zéaxanthine. Ces deux réactions successives nécessitent la présence d'acide ascorbique et d'un pH acide, formé par la présence de pompes à protons activées par la lumière. Pour sa part, la zéaxanthine époxidase est située dans les membranes thylakoidales du côté du stroma. À l'obscurité ou par très faible luminosité, elle catalyse l'époxidation de la zéaxanthine en anthéraxanthine puis en violaxanthine. Les réactions catalysées par ces deux protéines utilisent la ferredoxine et le FAD comme co-substrats. Cette série d'interconversions entre la violaxanthine et la zéaxanthine se nomme cycle des xanthophylles ou cycle de la violaxanthine. Ce cycle protège l'appareil photosynthétique contre les effets nocifs de la photooxydation survenant en condition de lumière excessive. Les différents caroténoïdes produits par ce cycle dissipent l'excès d'énergie en générant de la chaleur (Bugos et al., 1998).

Bien que la logique veuille qu'une plante ayant un contenu supérieur en zéaxanthine puisse être plus tolérante au stress photooxydatif, la réalité est toute autre. Des résultats préliminaires ont indiqué que des plantes de tabac surexprimant la violanxanthine dé-époxidase et montrant un taux de conversion accru de la violaxanthine en zéaxanthine ne démontrent pas un niveau supérieur de photoprotection (Hieber et al., 2002). Dans le même ordre d'idées, le mutant *npq1* chez *Arabidopsis thaliana*, qui est incapable de former de la zéaxanthine, montre un développement normal sous le chaud soleil de la Californie. Ceci laisse présager que d'autres mécanismes doivent exister pour protéger la plante de la photoinhibition causée par une forte intensité lumineuse. Cependant, des feuilles détachées de ce mutant ont montré une sensibilité accrue suite à un traitement avec une lumière forte de courte durée sous conditions contrôlées, et les lipides des membranes chloroplastiques ont montré une peroxydation accrue. Ces dommages photooxydatifs (taches blanches et zones nécrotiques au niveau des feuilles) sont aussi amplifiés chez le mutant par rapport au type sauvage suite à des traitements au froid et à la sécheresse (Niyogi et al., 1998).

Malgré le fait que leur appartenance à la famille des lipocalines soit critiquée, les enzymes du cycle des xanthophylles sont néanmoins des protéines jouant un rôle de premier plan lors de conditions de stress abiotiques. Le seul fait de pouvoir établir une certaine ressemblance entre ces protéines et la famille des lipocalines laisse envisager que les lipocalines de plantes pourraient jouer des rôles de premier plan dans la protection de la cellule végétale.

4.0 Bibliographie

Alabadí, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Más, P. et Kay, S.A. (2001) Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. Science 293: 880-883.

Alabadí, D., Yanovsky, M.J., Más, P., Harmer, S.L. et Kay, S.A. (2002) Critical role for CCA1 and LHY in maintaining circadian rhythmicity in *Arabidopsis*. Curr. Biol. 12: 757-761.

Allhorn, M., Lundqvist, K., Schmidtchen, A. et Åkerström, B. (2003) Heme-scavenging role of alpha1-microglobulin in chronic ulcers. J Invest Dermatol. 121: 640-646.

Allhorn, M., Klapyta, A. et Åkerström B. (2005) Redox properties of the lipocalin alpha1-microglobulin: reduction of cytochrome c, hemoglobin, and free iron. Free Radic Biol Med. 38: 557-567

Allard, F., Houde, M., Kröl, M., Ivanov, A., Huner, N.P.A. et Sarhan, F. (1998) Betaine improves freezing tolerance in wheat. Plant Cell Physiol. 39: 1194-1202.

Allen, R.D., Webb, R.P. et Schake, S.A. (1997) Use of transgenic plants to study antioxidant defenses. Free Radic. Biol. Med. 23: 473–479.

Åkerström, B., Flower, D.R. et Salier, J.P. (2000) Lipocalins: unity in diversity. Biochim. Biophys. Acta 1482: 1-8.

Apel, K. et Hirt, H. (2004) Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55: 373–399.

Asada, K. (1999) The water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. Ann. Rev. Plant Physiol. Plant Mol. Biol. 50: 601–639.

Bargiello, T.A. et Young, M.W. (1984) Molecular genetics of a biological clock in *Drosophil*a. Proc. Natl. Acad. Sci. USA. 81: 2142-2146.

Bartels, D. et Sunkar, R. (2005) Drought and salt tolerance in plants. CRC Crit. Rev. Plant Sci. 24: 23–58.

Blanco-Vaca, F., Via, D.P., Yang, C.Y., Massey, J.B. et Pownall, H.J. (1992) Characterization of disulfide-linked heterodimers containing apolipoprotein D in human plasma lipoproteins. J. Lipid Res. 12: 1785-1796.

Bohnert, H. et Shen, B. (1999) Transformation and compatible solutes. Sci. Hortic. 78: 237-260.

Bowler, C., Slooten, L., Vandenbranden, S., De Rycke, R., Botterman, J., Sybesma, C., Van Montagu, M. et Inzé, D. (1991) Manganese superoxide dismutase can reduce cellular damage mediated by oxygen radicals in transgenic plants. EMBO J. 10: 1723–1732.

Breton G, Danyluk J, Ouellet F, Sarhan F. (2000). Biotechnological applications of plant freezing associated proteins. Biotechnol. Annu. Rev. 6: 57-99.

Breton, G., Danyluk, J., Charron, J.B.F. et Sarhan, F. (2003) Expression profiling and bioinformatic analyses of a novel stress-regulated multispanning transmembrane protein family from cereals and *Arabidopsis*. Plant Physiol. 132: 64-74.

Bugos, R.C., Hieber, A.D. et Yamamoto, H.Y. (1998) Xanthophyll cycle enzymes are members of the lipocalin family, the first identified from plants. J. Biol. Chem. 273: 15321-15324.

Burke, B.J., Redondo, C., Redl, B. et Findlay, J.B.C. (2006) Lipocalin Receptors: Into the Spotlight. *Dans* Lipocalins. *Éditeurs*: B. Åkerström, N. Borregaard, D.R. Flower, et J.P. Salier. Landes Bioscience. Georgetown, TX. pp.167-176.

Burke, M.J., Gusta, L.V., Quamme, H.A., Weiser, C.J. et Li, P.H. (1976) Freezing and injury in plants. Ann. Rev. Plant Physiol. 27: 507-528.

Cadieux, C., Sarhan, F. et Perras, M. (1988) Osmotic adjustement and photosynthetic electron transport response to cold hardening in winter and spring wheat. Plant Physiol. Biochem. 26: 313-322.

Castonguay, Y. et Guckert, A. (1996) Adaptation of forage legumes to cold climates. Grassland and Land use systems. EGF Meetings contributions 532: 911-917.

Chakraborti, T., Das, S., Mondal, M., Roychoudhury, S. et Chakraborti, S. (1999) Oxidant, mitochondria and calcium: an overview. Cell. Signal. 11: 77–85.

Charron, J.B., Breton, G., Danyluk, J., Muzac, I., Ibrahim, R.K. et Sarhan, F. (2002) Molecular and biochemical characterization of a cold-regulated phospho-ethanolamine N-methyltransferase from wheat. Plant Physiol. 129: 363-373.

Charron, J.B., Ouellet, F., Pelletier, M., Danyluk, J., Chauve, C. et Sarhan, F. (2005) Identification, expression, and evolutionary analyses of plant lipocalins. Plant Physiol. 139: 2017-2028.

Crowe, J.H., Hoekstra, F.A. et Crowe, L.M. (1992) Anhydrobiosis. Ann. Rev. Plant Physiol. 54: 579-599.

Danyluk, J. et Sarhan, F. (1990) Differential mRNA transcription during the induction of freezing tolerance in spring and winter wheat. Plant Cell Physiol. 31: 609-619.

Danyluk, J., Rassart, E. et Sarhan, F. (1991) Gene expression during cold and heat shock in wheat. Biochem. Cell Biol. 69: 383-391.

Danyluk, J., Houde, M., Rassart, E. et Sarhan, F. (1994) Differential expression of a gene encoding an acidic dehydrin in chilling sensitive and freezing tolerant gramineae species. FEBS Lett. 344: 20-24.

Demel, R.A., Dorrepaal, E., Ebskamp, M.J., Smeekens, J.C. et de Kruijff, B. (1998) Fructans interact strongly with model membranes. Biochim. Biophys. Acta. 1375: 36-42.

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

Doyle, M.R., Davis, S.J., Bastow, R.M., McWatters, H.G., Kozma-Bognar, L., Nagy, F., Millar, A.J. et Amasino, R.M. (2002) The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. Nature 419: 74-77.

Dunlap, J.C., Loros, J. et DeCoursey, P.J. (2004) Chronobiology: Biological Timekeeping. Sinauer, Sunderland, UK. 382 p.

Edwards, K.D., Lynn, J.R., Gyula, P., Nagy, F. et Millar AJ. (2005) Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. Genetics 170: 387-400.

Elstner, E.F. (1987) Metabolism of activated oxygen species. *Dans* Biochemistry of plants. *Éditeur*: D.D. Davies. Academic Press, London, UK. pp.253–315.

Farré, E.M., Harmer, S.L., Harmon, F.G., Yanovsky, M.J. et Kay, S.A. (2005) Overlapping and distinct roles of PRR7 and PRR9 in the *Arabidopsis* circadian clock Curr. Biol. 15: 47-54.

Flower, D.R., North, A.C. et Attwood, T.K. (1993) Structure and sequence relationships in the lipocalins and related proteins. Protein Sci. 2: 753-761.

Flower, D.R. (1995) Multiple molecular recognition properties of the lipocalin protein family. J. Mol. Recognit. 8: 185-195.

Flower, D.R. (1996) The lipocalin protein family: structure and function. Biochem. J. 318: 1-14.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G. et Putterill, J. (1999) GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. EMBO J. 18: 4679-4688.

Fowler, S. et Thomashow, M.F. (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690.

Foyer, C.H. et Noctor, G. (2005) Redox homeostasis and antioxidant signaling: A metabolic interface between stress perception and physiological responses. Plant Cell 17: 1866–1875.

Frenette-Charron, J.B., Breton, G., Badawi, M. et Sarhan, F. (2002) Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. FEBS Lett. 517: 129-132.

Gille, L. et Nohl, H. 2001. The ubiquinol/bc_1 redox couple regulates mitochondrial oxygen radical formation. Arch. Biochem. Biophys. 388: 34-38.

Goodman, D.S. (1984) Overview of current knowledge of metabolism of vitamin A and carotenoids. J. Natl. Cancer Inst. 73: 1375-1379.

Gould, P.D., Locke, J.C., Larue, C., Southern, M.M., Davis, S.J., Hanano. S,, Moyle, R., Milich, R., Putterill, J., Millar, A.J. et Hall, A. (2006) The molecular basis of temperature compensation in the *Arabidopsis* circadian clock. Plant Cell 18: 1177-1187.

Guy, C.L., Niemi, K.J. et Brambl, R. (1985) Altered gene expression during cold acclimation of spinach. Proc. Natl. Acad. Sci. USA. 82: 3673-3677.

Guy, C.L. (1990) Cold acclimation and freezing stress tolerance: role of protein metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 187-223.

Hamblen, M.J., White, N.E., Emery, P.T., Kaiser, K. et Hall, J.C. (1998) Molecular and behavioral analysis of four period mutants in *Drosophila melanogaster* encompassing extreme short, novel long, and unorthodox arrhythmic types. Genetics 149: 165-178.

Hare, P.D., Cress, W.A. et van Staden, J. (1999). Proline synthesis and degradation: a model system for elucidating stress-related signal transduction. J. Exp. Bot. 50: 413-434.

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

Harmer, S.L. et Kay, S.A. (2005) Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*. Plant Cell 17: 1926-1940.

Hazen, S.P., Schultz, T.F., Pruneda-Paz, J.L., Borevitz, J.O., Ecker, J.R. et Kay, S.A. (2005) LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc. Natl. Acad. Sci. USA. 102: 10387-10392.

Hieber, A.D., Bugos, R.C et Yamamoto, H.Y. (2000) Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. Biochim. Biophys. Acta 1482: 84-91.

Hieber, A.D., Bugos, R.C., Verhoeven, A.S. et Yamamoto, H.Y. (2002) Overexpression of violaxanthin de-epoxidase: properties of C-terminal deletions on activity and pH-dependent lipid binding. Planta 214: 476-483.

Higgins, J., Hodges, N.A., Olliff, C.J. et Phillips, A.J. (1987) A comparative investigation of glycinebetaine and dimethylsulphoxide as liposome cryoprotectants. J. Pharm. Pharmacol. 39: 577-582.

Holmberg, N. et Bülow, L. (1998) Improving stress tolerance in plants by gene transfer. Trends Plant Sci. 3: 61-66.

Houde, M., Dhindsa, R.S. et Sarhan, F. (1992) A molecular marker to select for freezing tolerance in Gramineae. Mol. Gen. Genet. 234: 43-48.

Houde, M., Belcaid, M., Ouellet, F., Danyluk, J., Monroy, A.F., Dryanova, A., Gulick, P., Bergeron, A., Laroche, A., Links, M.G., MacCarthy, L., Crosby, W.L. et Sarhan, F. (2006) Wheat EST resources for functional genomics of abiotic stress. BMC Genomics 7: 149.

Huner, N.P.A., Oquist, G. et Sarhan, F. (1998) Energy balance and acclimation to light and cold. Trends Plant Sci. 3: 224-230.

Kaczorowski, K.A. et Quail, P.H. (2003) *Arabidopsis* PSEUDO-RESPONSE REGULATOR7 is a signaling intermediate in phytochrome-regulated seedling deetiolation and phasing of the circadian clock. Plant Cell 15: 2654-2665.

Kishor, P.B.K, Hong, Z., Miao, G.H., Hu, C.A. et Verma, D.P.S. (1995) Overexpression of D1-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. Plant Physiol. 108: 1387-1394.

Ko, R., Smith, L.T. et Smith, G.M. (1994) Glycine betaine confers enhanced osmotolerance and cryotolerance on Listeria monocytogenes. J.Bacteriol. 176: 426-431.

Kumar, A. et Minocha, S.C. (1998) Transgenic Manipulation of Polyamine Metabolism. Amsterdam: Harwood.Ladyman JAR, Hitz WD, Hanson AD. 1980. Translocation and metabolism of glycine betaine by barley plants in relation to water stress. Planta 150: 191-196.

Larsson, J., Allhorn, M. et Kerstrom, B. (2004) The lipocalin alpha(1)-microglobulin binds heme in different species. Arch Biochem Biophys. 432: 196-204.

Leegood, R.C. et Furbank, R.T. (1986) Stimulation of photosynthesis by 2% oxygen at low-temperatures is restored by phosphate. Planta 168: 84-93.

Limin, A.E., Danyluk, J., Chauvin, L.P., Fowler, D.B. et Sarhan, F. (1997) Chromosome mapping of low-temperature induced Wcs120 family genes and regulation of cold-tolerance expression in wheat. Mol. Gen. Genet. 253: 720-727.

Liu, Y., Garceau, N.Y., Loros, J.J. et Dunlap, J.C. (1997) Thermally regulated translational control of FRQ mediates aspects of temperature responses in the neurospora circadian clock. Cell 89: 477-486.

Liu, Z., Chang G.Q. et Leibowitz, S.F. (2001) Apolipoprotein D interacts with the long-form leptin receptor: a hypothalamic function in the control of energy homeostasis. FASEB J. 15: 1329-1331.

Locke, J.C.W., Southern, M.M., Kozma-Bognar, L., Hibberd, V., Brown, P.E., Turner, M.S. et Millar, A.J. (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. Mol. Syst. Biol. 1: 2005.0013.

Lloyd, A.W., Ollif, C.J. et Rutt, K.J. (1994) A study of modified betaines as cryoprotective additives. J. Pharm. Pharmacol. 46: 704-707.

Lynch, D.V. et Steponkus, P.L. (1987) Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (Secale cereale L. cv Puma). Plant Physiol. 83: 761-767.

Lyons, J.M. (1973) Chilling injury in plants. Ann. Rev. Plant Physiol. 24: 445-466.

Maggio, A., Miyazaki, S., Veronese, P., Fujita, T., Ibeas, J.I., Damsz, B., Narasimhan, M.L., Hasegawa, P.M., Joly, R.J. et Bressan, R.A. (2002) Does proline accumulation play an active role in stress-induced growth reduction? Plant J. 31: 699-712.

Majercak, J., Sidote, D., Hardin, P.E. et Edery, I.. (1999) How a circadian clock adapts to seasonal decreases in temperature and day length. Neuron 24: 219-230.

Malaba, L., Smeland, S., Senoo, H., Norum, K.R., Berg, T., Blomhoff, R. et Kindberg, G.M. (1995) Retinol-binding protein and asialo-orosomucoid are taken up by different pathways in liver cells. J. Biol. Chem. 270: 15686-15692.

Malpeli, G., Folli, C. et Berni, R. (1996) Retinoid binding to retinol-binding protein and the interference with the interaction with transthyretin. Biochim. Biophys. Acta 1294: 48-54.

Makino, S., Matsushika, A., Kojima, M., Oda, Y. et Mizuno, T. (2001) Light response of the circadian waves of the APRR1/TOC1 quintet: when does the quintet start singing rhythmically in *Arabidopsis*? Plant Cell Physiol. 42: 334-339.

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

Matsushika, A., Imamura, A., Yamashino, T. et Mizuno, T. (2002) Aberrant expression of the light-inducible and circadian-regulated APRR9 gene belonging to the circadian-associated APRR1/TOC1 quintet results in the phenotype of early flowering in *Arabidopsis thaliana*. Plant Cell Physiol. 43: 833-843.

McClung, C.R., Fox, B.A. et Dunlap, J.C. (1989) The Neurospora clock gene frequency shares a sequence element with the Drosophila clock gene period. Nature 339: 558-562.

McClung, C.R. (2001) Circadian rhythms in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52:139-162.

McClung, C.R. (2006) Plant circadian rhythms. Plant Cell 18: 792-803.

McCue, K.F. et Hanson, A.D. (1992) Salt-inducible betaine aldehyde dehydrogenase from sugar beet: cDNA cloning and expression. Plant Mol. Biol. 18: 1-11.

McKersie, B.D., Chen, Y., de Beus, M., Bowley, S.R., Bowler, C., Inze, D., D'Halluin, K. et Botterman, J. (1993) Superoxide dismutase enhances tolerance of freezing stress in transgenic alfalfa (*Medicago sativa* L.). Plant Physiol. 103: 1155–1163.

McKersie, BD., Bowley, SR., Harjanto, E. et Le Prince, O. (1996) Water-deficit tolerance and field performance of transgenic alfalfa overexpressing superoxide dismutase. Plant Physiol. 111: 1177–1181.

McKersie, B.D., Bowley, S.R., et Jones, K.S. (1999) Winter survival of transgenic alfalfa overexpressing superoxide dismutase. Plant Physiol. 119: 839–848.

McKersie, B.D., Murnaghan, J., Jones, K.S. et Bowley, S.R. (2000) Iron-superoxide dismutase expression in transgenic alfalfa increases winter survival without a detectable increase in photosynthetic oxidative stress tolerance. Plant Physiol. 122: 1427–1438.

McNeil, S., Nuccio, M. et Hanson, A. (1999) Betaines and related osmoprotectants. Targets for metabolic engineering of stress resistance. Plant Physiol. 120: 945-949.

Michael, T.P., Salome, P.A., Yu, H.J., Spencer, T.R., Sharp, E.L., McPeek, M.A., Alonso, J.M., Ecker, J.R. et McClung, C.R. (2003) Enhanced fitness conferred by naturally occurring variation in the circadian clock. Science 302: 1049-1053.

Millar, A.J. (2004) Input signals to the plant circadian clock. J. Exp. Bot. 55: 277-283.

Mittler, R. (2002) Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci. 7: 405–410.

Mittler, R., Vanderauwere, S., Gollery, M., et van Breusegem, F. (2004) Reactive oxygen gene network of plants. Trends Plant Sci. 9: 490–498.

Mizuno, T. et Nakamichi, N. (2005) Pseudo-Response Regulators (PRRs) or True Oscillator Components (TOCs). Plant Cell Physiol. 46: 677-685.

Möller, I.M. (2001) Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52: 561–91.

Monaco, H.L. et Zanotti, G. (1992) Three-dimensional structure and active site of three hydrophobic molecule-binding proteins with significant amino acid sequence similarity. Biopolymers 32: 457-465.

Murakami, M., Yamashino, T. et Mizuno, T. (2004) Characterization of circadian-associated APRR3 pseudo-response regulator belonging to the APRR1/TOC1 quintet in *Arabidopsis thaliana*. Plant Cell Physiol. 45: 645-650.

Nakamichi, N., Kita, M., Ito, S., Yamashino, T. et Mizuno, T. (2005) PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of *Arabidopsis thaliana*. Plant Cell Physiol. 46: 686-698.

Nakamura, T., Yokota, S., Muramoto, Y., Tsutsui, K., Oguri, Y., Fukui, K. et Takabe, T. (1997) Expression of a betaine aldehyde dehydrogenase gene in rice, a glycine betaine nonaccumulator, and possible localization of its protein in peroxisomes. Plant J. 11: 1115-1120.

Nanjo, T., Kobayashi, M., Yoshiba, Y., Sanada, Y., Wada, K., Tsukaya, H., Kakubari, Y.,Yamaguchi-Shinozaki, K. et Shinozaki, K. (1999) Biological functions of proline in morphogenesis and osmotolerance revealed in antisense transgenic *Arabidopsis thaliana*. Plant J. 18: 185-193.

Niki, T. et Sakai, A. (1983) Effect of cycloheximide on the freezing tolerance and ultrastructure of cortical parenchyma cells from mulberry twigs. Can. J. Bot. 61: 2205–2221.

Nishida, I. et Murata, N. (1996) Chilling sensitivity in plants and cyanobacteria: The crucial contribution of membrane lipids. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 541-568.

Niyogi, K.K., Grossman, A.R. et Bjorkman, O. (1998) *Arabidopsis* mutants define a central role for the xanthophyll cycle in the regulation of photosynthetic energy conversion. Plant Cell 10: 1121-1134.

North, A.C. (1989) Three-dimensional arrangement of conserved amino acid residues in a superfamily of specific ligand-binding proteins. Int. J. Biol. Macromol. 11: 56-58.

Nuccio, M.L., Rhodes, D., McNeil, S.D. et Hanson, A.D. (1999) Metabolic engineering of plants for osmotic stress resistance. Curr. Opin. Plant Biol. 2: 128-134.

Olien, C.R. et Clark, J.L. (1993) Changes in soluble carbohydrate composition of barley, wheat, and rye during winter. Agr.J. 85: 21-29.

Perras, M. et Sarhan, F. (1984) Energy state of spring and winter wheat during cold hardening. Soluble sugars and adenine nucleotides. Physiol. Plant. 60: 129-132.

Pervaiz, S. et Brew, K. (1985) Homology of beta-lactoglobulin, serum retinol-binding protein, and protein HC. Science 228: 335-337.

Pervaiz, S. et Brew, K. (1987) Homology and structure-function correlations between alpha 1-acid glycoprotein and serum retinol-binding protein and its relatives. FASEB J. 1: 209-214.

Prasad, T.K., Anderson, M.D., Martin, B.A. et Stewart, C.R. (1994) Evidence for chilling-induced oxidative stress in maize seedlings and a regulatory role for hydrogen peroxide. Plant Cell 6: 65-74.

Price, J.L. (1997) Insights into the molecular mechanisms of temperature compensation from the *Drosophila* PERIOD and TIMELESS mutants. Chronobiol. Int. 14: 455-468.

Raison, J.K. (1973) The influence of temperature-induced phase changes on kinetics of respiratory and other membrane-associated enzymes. J. Bioenerg 4: 258-309.

Rothenfluh, A., Abodeely, M., Price, J.L. et Young, M.W. (2000) Isolation and analysis of six timeless alleles that cause short- or long-period circadian rhythms in *Drosophila.* Genetics 156: 665-675.

Roughan, P.G. (1985) Phosphatidylglycerol and chilling sensitivity in plants. Plant Physiol. 77: 740-746.

Saito, A., Pietromonaco, S., Loo, A.K. et Farquhar, M.G. (1994) Complete cloning and sequencing of rat gp330/"megalin," a distinctive member of the low density lipoprotein receptor gene family. Proc. Natl. Acad. Sci. USA. 91: 9725-9729.

Salier, J.P. (2000) Chromosomal location, exon/intron organization and evolution of lipocalin genes. Biochim. Biophys. Acta 1482: 25-34.

Salomé, P.A. et McClung, C.R. (2005) What makes *Arabidopsis* tick: Light and temperature entrainment of the circadian clock. Plant Cell Environ. 28: 21-38.

Sarhan, F. et D'Aoust, M.J. (1975) RNA synthesis in spring and winter wheat during cold acclimation. Physiol. Plant. 35: 62-65.

Sarhan, F. et Chevrier, N. (1985) Regulation of RNA synthesis by DNA-dependent RNA polymerases and RNases during cold acclimation in winter and spring wheat. Plant Physiol. 78: 250-255.

Sarhan, F. et Perras, M. (1987) Accumulation of a high molecular weight protein during cold hardening of wheat (Triticum aestivum L.). Plant Cell Physiol. 28: 1173-1179.

Sato, E., Nakamichi, N., Yamashino, T. et Mizuno, T. (2002) Aberrant expression of the *Arabidopsis* circadian-regulated APRR5 gene belonging to the APRR1/TOC1 quintet results in early flowering and hypersensitiveness to light in early photomorphogenesis. Plant Cell Physiol. 43: 1374-1385.

Sawyer, L. et Richardson, J.S. (1991) Using appropriate nomenclature. Trends Biochem. Sci. 16: 11.

Schaffer, M.A. et Fischer, R.L. (1988) Analysis of mRNA that accumulate in response to low temperature identifies a thiol protease gene in tomato. Plant Physiol. 87: 431-436.

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

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

Schubert, D. et LaCorbiere, M. (1985) Isolation of an adhesion-mediating protein from chick neural retina adherons. J. Cell Biol. 101: 1071-1077.

Sehgal, A., Price, J.L., Man, B. et Young, M.W. (1994) Loss of circadian behavioral rhythms and per RNA oscillations in the *Drosophila* mutant timeless. Science 263: 1603-1606.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. et Shinozaki, K. (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13: 61-72.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y. et Shinozaki, K. (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31: 279-292.

Sen Gupta, A., Heinen, J.L., Holaday, AS., Burke, J.J. et Allen, R.D. (1993a). Increased resistance to oxidative stress in transgenic plants that overexpress chloroplastic Cu/Zn superoxide dismutase. Proc. Natl. Acad. Sci. USA. 90: 1629–1633.

Sen Gupta, A., Webb, R.P., Holaday, A.S. et Allen, R.D. (1993b). Over-expression of superoxide dismutase protects plants from oxidative stress: Induction of ascorbate peroxidase in superoxide dismutase-overexpressing plants. Plant Physiol. 103: 1067–1073.

Senoo, H., Stang, E., Nilsson, A., Kindberg, G.M., Berg, T., Roos, N., Norum, K.R. et Blomhoff, R. (1990) Internalization of retinol-binding protein in parenchymal and stellate cells of rat liver. J. Lipid Res. 31: 1229-1239.

Sharkey, T.D., Stitt, M., Heineke, D., Gerhardt, R., Raschke, K. et Heldt, H.W. (1986) Limitation of photosynthesis by carbon metabolism. II. O2-insensitive CO2 uptake results from limitations of triose phosphate utilization. Plant Physiol. 81: 1123-1129.

Siminovitch, D., Rheaume, B., Pomeroy, K. et Lepage, M. (1968) Phospholipid, protein, and nucleic acid increases in protoplasm and membrane structures associated with development of extreme freezing resistance in black locust tree cells. Cryobiology 5: 202-205.

Singh, J. et Laroche, A. (1988) Freezing tolerance in plants: a biochemical overview. Biochemistry and Cell Biol. 66: 650-657.

Sivaprasadarao, A., Boudjelal, M. et Findlay, J.B. (1993) Lipocalin structure and function. Biochem. Soc. Trans. 21: 619-622.

Slooten, L., Capiau, K., Van Camp, W., Van Montagu, M., Subesma, C. et Inzé, D. (1995) Factors affecting the enhancements of oxidative stress tolerance in transgenic tobacco overexpressing manganese superoxide dismutase in the chloroplasts. Plant Physiol. 107: 737–750.

Steponkus, P.L. (1984) Role of the plasma membrane in freezing injury and cold acclimation. Annu. Rev. Plant Physiol. 35: 543-584.

Steponkus, P.L., Uemura, M., Balsamo, R.A., Arvinte, T. et Lynch, D.V. (1988) Transformation of the cryobehavior of rye protoplasts by modification of the plasma membrane lipid composition. Proc Natl Acad Sci USA. 85: 9026-9030.

Stitt, M., Grosse, H. et Woo, K.C. (1988) Interactions between sucrose synthesis and CO2 fixation. II. Alterations of fructose 2,6-bisphosphate during photosynthetic oscillations. J. Plant Physiol. 133:138-143.

Stitt, M. et Hurry, V. (2002) A plant for all seasons: alterations in photosynthetic carbon metabolism during cold acclimation in *Arabidopsis*. Curr. Opin. Plant Biol. 5: 199-206.

Tramontano, W.A. et Jouve, D. (1997) Trigonelline accumulation in salt-stressed legumes and the role of other osmoregulators as cell cycle control agents. Phytochemistry 44: 1037-1040.

Taiz, L. et Zeigler, E. (1991) Plant physiology. Benjamin/Cummings Pub Co. Redwood City, CA. 559 p.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K. et Shinozaki, K. (2002) Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. Plant J. 29: 417-426.

Thomashow, M.F. (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Ann. Rev. Plant Physiol. Plant Mol. Biol. 50: 571-599.

Uemura, M. et Yoshida, S. (1984) Involvement of plasma membrane alterations in cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). Plant Physiol. 75: 818-826.

Uemura, M., Joseph, R.A. et Steponkus, P.L. (1995) Cold acclimation of *Arabidopsis thaliana* (Effect on plasma membrane lipid composition and freeze-induced lesions). Plant Physiol. 109: 15-30.

Vágújfalvi, A., Kerepesi, I., Galiba, G., Tischner, T. et Sutka, J. (1999) Frost hardiness depending on carbohydrate changes during cold acclimation in wheat. Plant Sci. 144: 85-92.

Van Breusegem, F., Slooten, L., Stassart, J.M., Moens, T., Botterman, J., Van Montagu, M. et Inze, D. (1999) Overproduction of *Arabidopsis thaliana* FeSOD confers oxidative stress tolerance to transgenic maize. Plant Cell Physiol. 40: 515–523.

Van Camp, W., Capiau, K., Van Montagu, M., Inze, D. et Stoolen, L. (1996) Enhancement of oxidative stress tolerance in transgenic tobacco plants overproducing Fe-superoxide dismutase in chloroplasts. Plant Physiol. 112: 1703–1714.

Vereyken, I.J., Chupin, V., Demel, R.A., Smeekens, S.C. et De Kruijff, B. (2001) Fructans insert between the headgroups of phospholipids. Biochim. Biophys. Acta 1510: 307-320.

Weiser, C.J. (1970) Cold resistance and injury in woody plants. Science 169: 1269-1278.

Welti, R., Li, W., Li, M., Sang, Y., Biesiada, H., Zhou, H.E., Rajashekar, C.B., Williams, T.D. et Wang, X. (2002) Profiling membrane lipids in plant stress responses. Role of phospholipase D alpha in freezing-induced lipid changes in *Arabidopsis*. J. Biol. Chem. 277: 31994-32002.

Weretilnyk, E.A. et Hanson, A.D. (1990) Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. Proc. Natl. Acad. Sci. USA. 87: 2745-2749.

Yamamoto, Y., Yoshizawa, T., Kamio, S., Aoki, O., Kawamata, Y., Masushige, S. et Kato, S. (1997) Interactions of transthyretin (TTR) and retinol-binding protein (RBP) in the uptake of retinol by primary rat hepatocytes. Exp. Cell Res. 234: 373-378.

Yancey, P.H., Clark, M.E., Hand, S.C., Bowlus, R.D. et Somero, G.N. (1982) Living with water stress: evolution of osmolyte systems. Science 217: 1214-1222.

Yancey, P. (1994). Compatible and counteracting solutes. *Dans* Cellular and molecular physiology of cell volume regulation. *Éditeur* K. Strange. CRC Press. Boca Raton, FL. pp 81-109.

Zehring, W.A., Wheeler, D.A., Reddy, P., Konopka, R.J., Kyriacou, C.P., Rosbash, M. et Hall, J.C. (1984) P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic *Drosophila melanogaster*. Cell 39: 369-376.

Zhang, L., Dunn, M.A., Pearce, R.S. et Hughes, M.A. (1993) Analysis of organ specificity of low-temperature-responsive gene family in rye (*Secale cereale L.*). J. Exp. Bot. 44: 1787-1793.

Zhao, Y., Aspinall, D. et Paleg, L.G. (1992) Protection of membrane integrity in *Medicago sativa* L. by glycinebetaine against the effects of freezing. J. Plant Physiol. 140: 541-543.

PROBLÉMATIQUE

Les céréales représentent une des cultures les plus importantes sur la planète. Les coûts occasionnés par les gels précoces à l'automne et tardifs au printemps assombrissent le potentiel économique relié à ces cultures dans les pays nordiques comme le Canada. L'amélioration de la tolérance au gel par génie génétique ou par sélection assistée par des marqueurs moléculaires pourrait augmenter et diversifier ces productions, notamment en prolongeant la saison de croissance.

Outre leur importance économique, les céréales possèdent une variabilité génétique en ce qui a trait à la tolérance au gel. Cette variabilité nous permet de distinguer entre les changements reliés à la croissance lors des basses températures et ceux impliqués dans l'acquisition de la tolérance au gel. Le blé, un membre important des espèces céréalières, possède cette variabilité. Cette caractéristique fait donc de cette espèce un excellent modèle d'étude.

Des travaux préalables à mes études ont permis d'identifier chez le blé un clone codant pour une protéine faisant partie de la famille des lipocalines. L'expression de ce gène est régulée par les basses températures et associée avec le niveau de tolérance des différents cultivars. Les lipocalines se retrouvent chez les animaux vertébrés et invertébrés, les plantes et les bactéries. Quoique principalement connues comme protéines de transport, il est maintenant clair que les membres de cette famille accomplissent une variété de fonctions. Certaines lipocalines telles que l'Apolipoprotéine D sont largement étudiées en ce qui a trait à leur implication dans différentes maladies telles que l'Alzheimer, les syndromes Parkinsoniens, et différents types de cancer. L'identification, chez les plantes, de protéines démontrant une homologie de séquence avec les lipocalines a soulevé des questions quant à leur rôle dans les plantes. L'éventail des processus biochimiques et physiologiques dans lesquels les lipocalines sont impliquées et la distance phylogénétique entre les organismes inférieurs, les animaux et les plantes suggèrent que les membres de cette

famille de protéines jouent des rôles essentiels qui ont dû être maintenus au courant de l'évolution.

Considérant la vaste distribution des lipocalines (animaux, plantes, insectes et bactéries), et l'importance de leurs fonctions, il est apparu important d'identifier et caractériser les lipocalines végétales. Pour ce faire, une approche intégrée de compilation de données de séquence, profil d'expression, analyses phylogénétiques, et prédictions bioinformatiques afin de caractériser en détail l'implication de ces protéines dans l'acquisition de la tolérance au gel. L'ensemble des informations recueillies jusqu'à présent mène à croire que les lipocalines de plantes possèdent un potentiel important dans notre stratégie globale d'amélioration des plantes céréalières dans le but de leur conférer une plus grande tolérance au gel et autres stress abiotiques.

CHAPITRE II

Molecular and Structural Analyses of a Novel Temperature Stress Induced Lipocalin from Wheat and *Arabidopsis*

Jean-Benoit Frenette Charron, Ghislain Breton, Mohamed Badawi, Fathey Sarhan

FEBS Letters (2002) 517: 129-132

J'ai été impliqué dans toutes les facettes des travaux associés à cet article. J'ai élaboré le design expérimental et effectué les analyses bioinformatiques. J'ai aussi fait les analyses d'expression des ARN messagers chez le blé. J'ai finalement rédigé le manuscrit et conçu les figures. Ghislain Breton m'a conseillé lors de l'élaboration du design expérimental et des analyses de bioinformatique. Il a de plus réalisé les analyses d'expression des ARNm chez *Arabidopsis thaliana* et a édité le manuscrit. Mohamed Badawi, en formation à cette époque, a assisté lors des analyses d'expression des ARN messagers.

Résumé

Deux ADNc correspondant à une nouvelle lipocaline ont été identifiés chez le blé et *Arabidopsis*. Les deux ADNc, nommés *Tatil* et *Attil* pour « *Triticum aestivum* temperature-induced lipocalins » et « *Arabidopsis thaliana* temperature-induced lipocalins », codent pour des polypeptides de 190 et 186 acides aminés respectivement. L'analyse des séquences a démontré la présence de trois régions conservées qui caractérisent les lipocalines. De plus, cette analyse a établi une ressemblance entre cette nouvelle lipocaline et trois protéines reliées au sens évolutif: l'Apolipoprotéine D humaine, la lipocaline bactérienne Blc et la lipocaline d'insecte Lazarillo. Une comparaison des structures tertiaires probables de TaTIL et de l'Apolipoprotéine D humaine suggère que ces dernières diffèrent au niveau de leur site d'attachement à la membrane et de leur site de liaison au substrat. De plus, des analyses de type northern ont démontré que les ARN messagers de *Tatil* et *Attil* s'accumulent au cours de l'acclimatation au froid et suite à un choc thermique. Les fonctions possibles de cette nouvelle catégorie de lipocalines végétales au cours des stress de température sont discutées.

Mots clés : acclimatation au froid; ; apolipoprotéine D; *Arabidopsis thaliana*; lipocaline; stress à la chaleur; *Triticum aestivum* L.

Abstract

Two cDNAs corresponding to a novel lipocalin were identified from wheat and *Arabidopsis*. The two cDNAs designated *Tatil* for *Triticum aestivum* L. temperature induced lipocalin and *Attil* for *Arabidopsis thaliana* temperature induced lipocalin, encode polypeptides of 190 and 186 amino acids respectively. Structure analyses indicated the presence of the three structurally conserved regions that characterize lipocalins. Sequence analyses revealed that this novel class of plant lipocalin share homology with three evolutionarily related lipocalins: the mammalian apolipoprotein D, the bacterial lipocalin and the insect Lazarillo. The comparison of the putative tertiary-structures of both the human apolipoprotein D and the wheat *Ta*TIL suggest that the two proteins differ in membrane attachment and ligand interaction. Northern analyses demonstrated that *Tatil* and *Attil* transcripts are upregulated during cold acclimation and heat-shock treatment. The putative functions of this novel class of plant lipocalins during temperature stresses are discussed.

Abbreviations

ABA, abscisic acid; *At*, *Arabidopsis thaliana*; ApoD, apolipoprotein D; Blc, bacterial lipocalin; GPI, glycosylphosphatidylinositol; SCR, structurally conserved regions; SSC, saline sodium citrate; *Ta*, *Triticum aestivum* L.; TIL, temperature induced lipocalin.

Introduction

Lipocalins are a large and diverse group of small, mostly extracellular proteins that are found in vertebrates and invertebrate animals, plants and bacteria. They are characterized by a conserved ligand-binding pocket, which gives them the ability to bind small, principally hydrophobic molecules. The list of potential ligands is constantly growing and includes diverse molecules such as steroids, pheromones, and odorant molecules. Lipocalins were early established as transport proteins, but it is becoming increasingly clear that some of them may be implicated in many other important functions such as modulation of cell growth and metabolism, binding of cell-surface receptors, nerve growth and regeneration, regulation of the immune response, smell reception, cryptic coloration, membrane biogenesis and repair, and induction of apoptosis [1]. In plants, only one study reported the presence of two lipocalin-like proteins. Those lipocalins were found to be key enzymes of the xanthophyll cycle responsible for the protection against photo-oxidative damage [2].

Considering the wide distribution of lipocalins in animals, insects and bacteria, and their expression under conditions of environmental stress [1], it is important to survey the plant genome for stress regulated lipocalin-like proteins. Towards this goal, we searched our ESTs database, generated from cold acclimated wheat, and that of *Arabidopsis* ESTs collection from GenBank^{tm}, to determine if some of the clones isolated so far present sequence homology with lipocalins. Two homologous cDNAs, from wheat (*Tatil),* and *Arabidopsis thaliana* (*Attil)*, were found to encode proteins that contain the three structurally conserved regions (SCRs) that characterize lipocalins. Those proteins share significant homology with three established members of the lipocalin family, the mammalian apolipoprotein D, the bacterial lipocalin Blc, and the insect Lazarillo. The structure, regulation and putative function of this novel class of plant lipocalin during environmental stresses are discussed.

Materials and methods

2.1. Plant material and growth conditions

In this study we used two spring wheat genotypes (*Triticum aestivum* L. cv Glenlea and cv Concorde), 4 winter wheat genotypes (*T. aestivum* L. cv Monopole, cv Absolvent, cv Fredrick, and cv Norstar,), winter rye (*Secale cereale* L. cv Musketeer), oat (*Avena sativa* L. cv Laurent), barley (*Hordeum vulgare* L. cv Winchester) and *Arabidopsis thaliana* ecotype Columbia.

Cereal seeds were germinated in moist sterilized vermiculite for 5 days in the dark and 2 days under artificial light (225 μE) at 25°C / 20°C (day/night). Control plants were maintained under the same conditions while cold acclimation and other stress treatments were performed as previously described [3]. *Arabidopsis* plants were grown in pots in a 1:1 mixture of Promix (Premier) and vermiculite in a growth cabinet for 40 days under 8 hour artificial light (70 μE) at 22°C (70% R.H.). For cold treatment, plants were placed under the same light conditions at 4°C. For heat-shock treatment the plants were exposed to 45°C for 1 hour. This condition elicits a typical heat-shock response.

2.2. Cloning and molecular analysis

The wheat lipocalin cDNA was isolated from a Lambda Zap II library (Stratagene) constructed from poly $(A)^+$ RNA that was isolated from one day cold-acclimated winter wheat (*T. aestivum* L. cv Norstar) [4]. The *Arabidopsis* lipocalin was identified by homology search [5] using the wheat sequence against the *Arabidopsis* ESTs database. The identified clone (ID: 120O12) from the PRL2 library [6] was ordered from the *Arabidopsis* Biological Resource Center. Complete DNA sequences of wheat and *Arabidopsis* clones were determined from both strands. Cereals RNA extractions were performed as described previously [7]. Total RNA

from *Arabidopsis* was isolated using the Tri-Reagent (Molecular Research Center) according to the manufacturer protocol. Total RNA (7.5 μg) samples were mixed with ethidium bromide before electrophoresis on formaldehyde-agarose gels [8]. This allowed visual evaluation of RNA quantity and loads on gels. After electrophoresis, RNA was transferred to nitrocellulose membranes (Osmonics) in 20X SSC. The filters were baked for 2 h at 80°C prior to hybridization with corresponding ^{32}P-labelled p*Tatil* and p*Attil* inserts. Filters were washed at 65°C with several buffer changes of decreasing SSC concentration (5-0.1 X) and autoradiographed on Kodak MS films with MS intensifying screens (Kodak) at -80°C. Relative levels of *Tatil* and *Attil* mRNA transcripts were determined by densitometry scanning of the Northern blots using the ImageQuant 4.2 software (Molecular Dynamics).

Analysis and sequence comparisons were carried out with programs available on the ExPASy Molecular Biology Server.

Results and discussion

A novel plant lipocalin, *Tatil*, was identified from our collection of cold induced EST using the BLAST X software [5]. The full-length clone was then isolated from a wheat cDNA library and named *Tatil* for *Triticum aestivum* L. temperature induced lipocalin. The longest ORF of the cDNA is 570 bp and encodes a protein of 190 amino acids (aa) with an ATG codon at nucleotide 115 and a stop codon at nucleotide 686. The calculated molecular mass is 22 kDa and its theoretical pI is 5.5.

Search in the GenBank™ ESTs database revealed high homology (74 % identity, 83% similarity) with a putative protein from *Arabidopsis thaliana* that we have named *At*TIL for *Arabidopsis thaliana* temperature induced lipocalin. The complete sequencing of this *Arabidopsis* clone revealed that the cDNA encodes a 186 aa protein. Sequence analysis revealed that the N-terminal portion of both wheat and *Arabidopsis* proteins possess the three lipocalin SCRs. The SCR 1 region is located within aa 15-31 (GLDVARYMGRWYEIASF) in *Ta*TIL and within aa 12-28 (GLNVERYMGRWYEIASF) in *At*TIL and posses the two conserved amino acids G and W [1,9] (Fig.1). The SCR 2 of *Ta*TIL is found at the C-terminal portion of the protein within aa 105-119 (YWVLYVDDDYQYALV) while in *At*TIL it is found within aa 101-115 (YWVLYIDPDYQHALI) (Fig.1). Generally SCR 2 contains a TDY triplet at the positions underlined [1,9]. In *Ta*TIL and *At*TIL only the central D is present. SCR 3 is also found at the C-terminal portion of both proteins within aa 129-144 (ILCRKTHIEEEVNQL) in *Ta*TIL and within aa 125-140 in *At*TIL (ILSRTAQMEEETYKQL) (Fig.1). The conserved R residue that characterizes this fingerprint is present in both sequences [1,9].

Further sequence analysis of *Ta*TIL indicated the presence of only one cysteine at aa 130 and a putative N-glycosylation site at aa 60. On the other hand, the entire primary sequence of *At*TIL does not contain any cysteine, although the putative N-glycosylation site is found at aa 56. Putative C-terminal cleavage sites are

predicted by several target peptide prediction programs (DGPI, PSORT [10], and SignalP [11]) to be at aa 172 in *Ta*TIL and at aa 168 in *At*TIL. Considering this putative cleavage site, the calculated molecular mass of the mature protein of wheat and *Arabidopsis* is 20 kDa with a pI of 5.2.

The homology search revealed that *Ta*TIL (accession no. AY077702) and its ortholog from *Arabidopsis* (accession no. AY062789) share significant homology with three evolutionarily related lipocalins: the human apolipoprotein D precursor (ApoD) (accession no. P05090), the *E.coli* outer membrane lipoprotein Blc precursor (accession no. P39281), and the american grasshopper Lazarillo precursor (accession no. P49291) (Fig. 1). They respectively share 29 %, 31 %, and 23 % identity and 46 %, 54 % and 40 % similarity with *Ta*TIL. Among all lipocalins, Blc, ApoD, and Lazarillo are the only ones known to be anchored to biological membranes [12]. Thus it is possible that *Ta*TIL and *At*TIL are also membranes associated proteins.

The sequence analysis also revealed that *Ta*TIL and *At*TIL, like the *E.coli* Blc, are distinguished from most lipocalins by the absence of intramolecular disulphide bonds. However, they are potentially N-glycosylated like human ApoD and Lazarillo. When the three SCRs of these five proteins are aligned, the start methionines from *Ta*TIL and *At*TIL are positioned precisely at the cleavage sites of the N-terminal signal peptides of the three other proteins (Fig. 1). This alignment suggests that *Ta*TIL and *At*TIL do not possess an N-terminal signal peptide like *E.coli* Blc, Human ApoD and Lazarillo. The N-terminal portion of *Ta*TIL is composed of hydrophilic residues followed by few hydrophobic residues. In *At*TIL, the hydrophobic section is even less accentuated. This profile does not fit the standard hydrophobic nature of the N-terminal signal peptide identified in ApoD, Blc and Lazarillo.

Like Lazarillo, *Ta*TIL and *At*TIL are longer than human ApoD and Blc at their C-terminal end and possess a similar putative cleavage site (Fig. 1). The hydrophobic C-terminal tail after the cleavage site enables Lazarillo to receive a glycosylphosphatidylinositol (GPI) anchor [13]. This may suggest that *Ta*TIL and *At*TIL also receive a GPI anchor. The GPI anchor is a post-translational addition of a

lipid occurring in the endoplasmic reticulum lumen, which links proteins to the external face of the plasma membrane. This type of modifications has been reported in plants [14]. The fact that the N-glycosylation site is conserved between wheat and *Arabidopsis* TIL orthologs support the possibility that those proteins are processed in the endoplasmic reticulum lumen.

Another type of attachment to the membrane can also be suggested for *Ta*TIL and *At*TIL. It has been proposed that human ApoD is associated with the external face of the membrane by a hydrophobic loop [12,15,16]. *Ta*TIL and *At*TIL also possess an hydrophobic stretch of seven amino acids that is inserted into a loop between two ß-strands (Fig. 1). This insertion is not in the same loop as the human ApoD. It is between ß-strands 5 and 6 instead of 7 and 8 (Fig. 2, B1 and C2). However it is possible that this stretch favors the attachment to the plasma membrane. The model presented in figure 2, reveals a difference in the loop scaffold covering the large cup-shaped cavity that characterize the lipocalins (Fig. 2, B2). The loop scaffold in *Ta*TIL and *At*TIL is two amino acids longer than human ApoD and has a proline moiety at position 32 and 29 respectively (Fig. 1). Those modifications may suggest that the plant TIL has different binding specificity.

Northern blot analysis revealed that the *Tatil* transcripts accumulate to high levels upon exposure to low temperature (10-fold) and heat-shock treatments (10-fold) and to a lesser extent by water stress (3.5-fold). ABA, high salt and wounding treatments have no measurable effect (Fig. 3A). The *Tatil* transcripts accumulate gradually to a maximum level after 36 days of cold acclimation. Upon deacclimation, the level of transcripts returned to those seen in the control non-acclimated plants. The accumulation of *Tatil* transcripts in wheat was found to be tissue-specific, as they were detected only in cold acclimated leaves (Fig. 3A). Transcripts accumulation of *Attil* revealed that the dicot ortholog is also induced by low temperature (6 fold) and heat-shock treatments (9-fold) (Fig 3A). *Arabidopsis* genome sequence analysis (accession no. AB024029) revealed that the promoter of *Attil* does not contain any

known heat shock responsive element although two low temperature responsive elements are found [17,18].

RNA-blot hybridization studies also demonstrated that cold acclimation induced the accumulation of the *Tatil* transcripts in both less tolerant and hardy wheat (Fig. 3B). However, this increase was greater in the hardy winter cultivars. Low levels of expression are also found in oat and barley, two less cold tolerant species (Fig. 3B). This difference of accumulation indicates that the *Tatil* expression is correlated with the plants capacity to develop freezing tolerance.

Temperature stresses are known to induce membrane injuries. The membrane-anchored lipocalins (Blc, ApoD, Lazarillo and possibly *Ta*TIL and *At*TIL) all appear to be expressed in response to conditions that cause membrane stresses [12], which suggest a biological role in membrane biogenesis and repair under severe stress conditions.

*Ta*TIL and *At*TIL like human ApoD may possess a wide variety of potential ligands of different structures and functions. Mammalian ApoD is reported to bind arachidonic acid, bilirubine, steroid hormones (progesterone and pregnenolone) and cholesterol [19]. It is interesting to mention that plants also synthesize a wide variety of steroid hormones called brassinosteroids. Treatment with 24-epibrassinolide, a brassinosteroid, increases the tolerance to heat and cold stresses in plants [20]. The enhanced resistance to temperature stress was attributed to membrane stability and osmoregulation. These results suggest that part of the temperature-responses in plants may involve brassinosteroids as signaling molecules to elicit the expression of steroid binding proteins such as lipocalin. It is also known that sterol insertion in the plasma membrane increases its fluidity at low temperature and maintains the phospholipids order at high temperature [21]. *Ta*TIL may be involved in the transport of those sterol molecules to the membrane in response to stress conditions.

The specific expression of *Ta*TIL in leaf suggests that the protein may play a role in the chloroplast function during temperature stresses. However, there is no evidence of the presence of a chloroplast signal peptide to confirm this assumption.

Recent work on the OEP7 protein from spinach demonstrated that the interaction of this protein with the outer membrane of the chloroplast is independent of a classical cleavable targeting signal or membrane-channel protein [22]. It is possible that the *Ta*TIL is anchored to the outer chloroplast membrane in a similar manner under temperature stresses.

References

[1] Flower, D.R. (1996) Biochem. J. 318, 1-14.

[2] Bugos, R.C., Hieber, A.D. and Yamamoto, H.Y. (1998) J. Biol. Chem. 273, 15321-15324.

[3] Danyluk, J., Carpentier, E. and Sarhan, F. (1996) FEBS Lett. 389, 324-327.

[4] Houde, M., Danyluk, J., Laliberte, J.F., Rassart, E., Dhindsa, R.S. and Sarhan, F. (1992) Plant Physiol. 99, 1381-1387.

[5] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403-410.

[6] Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., et al. (1994) Plant Physiol. 106, 1241-1255.

[7] Danyluk, J. and Sarhan, F. (1990) Plant Cell Physiol. 31, 609-619.

[8] Rosen, K.M. and Villa-Komaroff, L. (1990) Focus 12, 23-24.

[9] Flower, D.R., North, A.C. and Attwood, T.K. (1993) Protein Sci. 2, 753-761.

[10] Nakai, K. and Horton, P. (1999) Trends Biochem. Sci. 24, 34-36.

[11] Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Int. J. Neural. Syst. 8, 581-599.

[12] Bishop, R.E. (2000) Biochim. Biophys. Acta 1482, 73-83.

[13] Ganfornina, M.D., Sanchez, D. and Bastiani, M.J. (1995) Development 121, 123-134.

[14] Morita, N., Nakazato, H., Okuyama, H., Kim, Y. and Thompson, G.A. Jr. (1996) Biochim. Biophys. Acta 1290, 53-62.

[15] Peitsch, M.C. and Boguski, M.S. (1990) New Biol. 2, 197-206.

[16] Bishop, R.E., Penfold, S.S., Frost, L.S., Holtje, J.V. and Weiner, J.H. (1995) J. Biol. Chem. 270, 23097-23103.

[17] Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Nucleic Acids Res. 27, 297-300.

[18] Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. and Tabata, S. J. (2000) DNA Res. 7, 31-63.

[19] Rassart, E., Bedirian, A., Do Carmo, S., Guinard, O., Sirois, J., Terrisse, L. and Milne, R. (2000) Biochim. Biophys. Acta 1482, 185-198.

[20] Clouse, S.D. and Sasse, J.M. (1998) Annu. Rev. Plant Physiol. Plant Mol. 49, 427-451.

[21] Demel, R.A. and De Kruijff, B. (1976) Biochim. Biophys. Acta 457, 109-132.

[22] Schleiff, E., Tien, R., Salomon, M. and Soll, J. (2001) Mol. Biol. Cell 12, 4090-4102.

[23] Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J. (1998) Bioinformatics 14, 892-893.

[24] Peitsch, M.C. (1996) Biochem. Soc. Trans. 24, 274-279.

Figure 1. Alignment of the deduced amino acid sequences of wheat *Ta*TIL and *Arabidopsis At*TIL with related lipocalins.

Identical residues are in black and similar residues are in gray. The three-structurally conserved regions that provide a signature for the lipocalins are indicated above. Putative cleavage sites are in red with yellow letters. The secondary structure is predicted from already published models [15,16] and analyzed with the Jpred2 program [23]. Red arrows, green rectangles and blue rectangle, represent ß-strands, α-helix, and 3_{10} α-helix respectively.

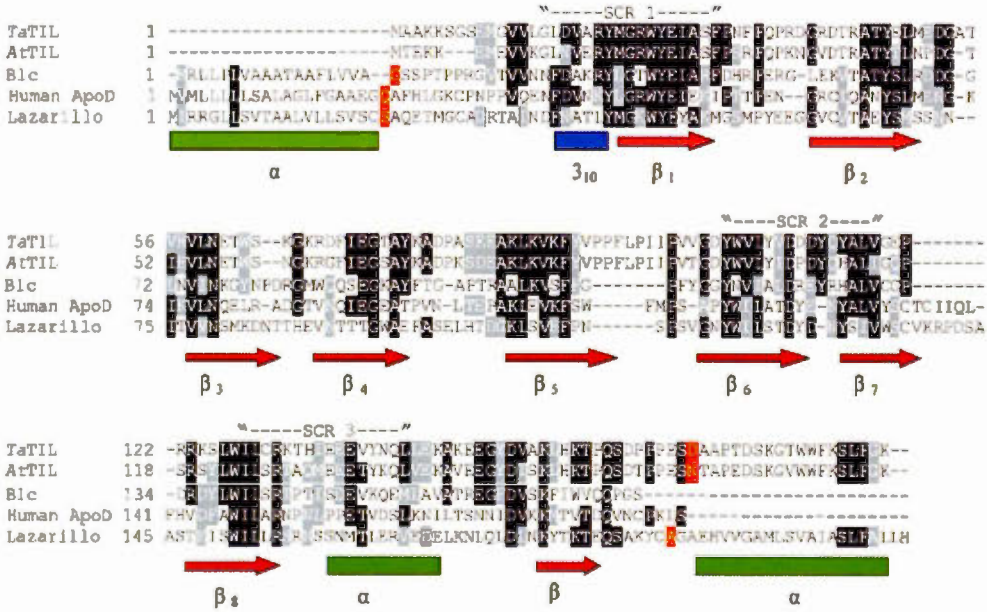

"-----SCR 1-----"
TaTIL 1
AtTIL 1
Blc 1
Human ApoD 1
Lazarillo 1
α
3₁₀
β 1
β 2
"-----SCR 2-----"
TaTIL 56
AtTIL 52
Blc 72
Human ApoD 74
Lazarillo 75
β 3
β 4
β 5
β 6
β 7
"-----SCR 3-----"
TaTIL 122
AtTIL 118
Blc 134
Human ApoD 141
Lazarillo 145
β 8
α
β
α

Figure 2. Tertiary-structure models of human ApoD and wheat TaTIL.

Tertiary structure analyses were carried out using the SWISS-MODEL program [24]. The lower blast limit was set at 0.00001 and the human ApoD model (PDB ID: 2APD) [15] was used as template. The initial result was then resubmitted through the optimizing mode of the program. The final result was then visualized using the Swiss-Pdb Viewer and the model was adapted according to sequence comparison. Differences between the wheat and the human models were superposed and colored. Sections of *Ta*TIL that differ from human ApoD are presented in red. Sections of human ApoD that differ from *Ta*TIL are in blue. Grey sections are common to both models.

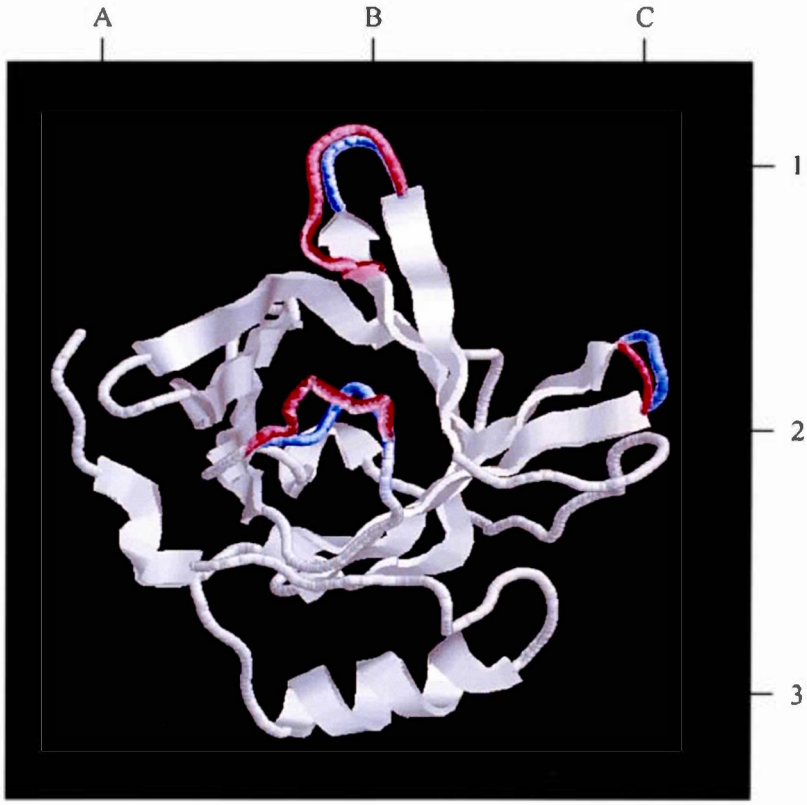
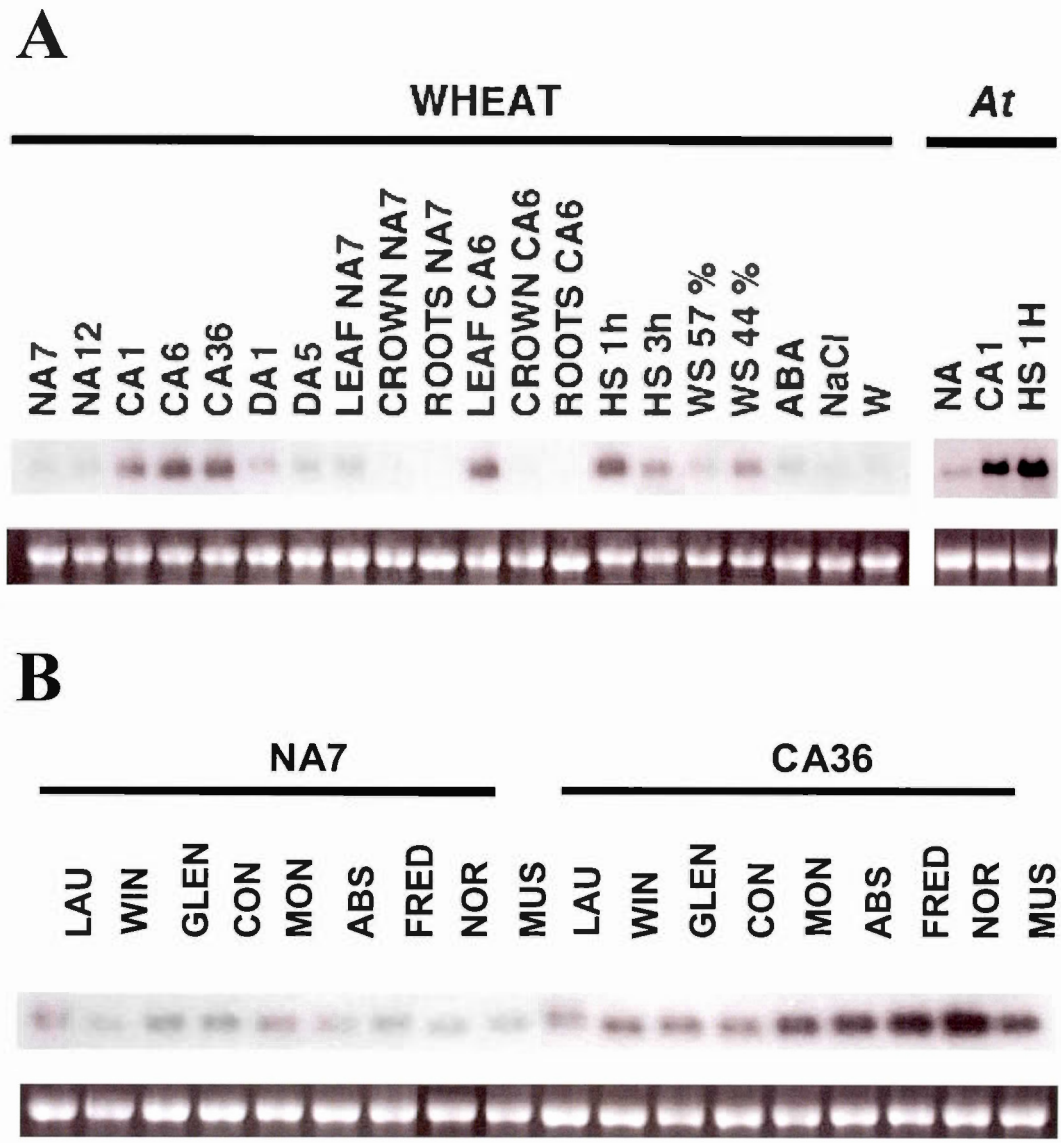

A
B
C
1
2
3

Figure 3. Upregulation of Tatil and Attil during cold acclimation and heat shock.

(A) Accumulation and tissue specificity of *Tatil* and *Attil* mRNAs under different stress conditions in winter wheat Fredrick and *Arabidopsis*. The 28S ribosomal band stained with ethidium bromide is included to show RNA loads (7.5 μg). NA7, NA12, non-acclimated plants grown for 7 and 12 days; CA1, CA6, CA36, cold acclimated plants for 1, 6 and 36 days; DA1 and DA5, cold-acclimated plants (36 days) were deacclimated for 1 and 5 days; NaCl, plants treated with 300mM NaCl for 18 h; ABA, plants treated with 0.1 mM ABA (Sigma) for 18 h; WS, water stressed plants with a relative water content of 57 % and 44 %; HS, plants exposed to 40°C for 1h and 3 h (heat shock); W, wounding stress for 3 h. Leaf, crown and roots of non-acclimated plants (NA7) and cold-acclimated for 6 days (CA6). *At*, *Arabidopsis* plants; NA, non acclimated plants; CA1, cold acclimated for 1 day; HS 1h, exposed to 45°C for 1h (heat shock).

(B) Accumulation of *Tatil* mRNAs during cold acclimation in spring and winter wheat and other cereal species. Total RNA (7.5 μg) from shoots of: two spring wheat genotypes (*Triticum aestivum* L. cv Glenlea (Glen), LT_{50} (lethal temperature that kills 50% of the seedlings) -8°C; and cv Concorde (Con), LT_{50} -8°C), 4 winter wheat genotypes (*T. aestivum* L. cv Monopole (Mon), LT_{50} -15°C; cv Absolvent (Abs), LT_{50} -16°C; cv Fredrick (Fred), LT_{50} -16°C; and cv Norstar (Nor), LT_{50} -19°C), winter rye (*Secale cereal* L. cv Musketeer (Mus), LT_{50} -21°C), oat (*Avena sativa* L. cv Laurent (Lau), LT_{50} -6°C), barley (*Hordeum vulgare* L. cv Winchester (Win), LT_{50} -7°C). NA, non-acclimated plants grown for 6 days; CA, cold acclimated plants for 36 days.


A
WHEAT
At
NA7
NA12
CA1
CA6
CA36
DA1
DA5
LEAF NA7
CROWN NA7
ROOTS NA7
LEAF CA6
CROWN CA6
ROOTS CA6
HS 1h
HS 3h
WS 57 %
WS 44 %
ABA
NaCl
W
NA
CA1
HS 1H
B
NA7
CA36
LAU
WIN
GLEN
CON
MON
ABS
FRED
NOR
MUS
LAU
WIN
GLEN
CON
MON
ABS
FRED
NOR
MUS

CHAPITRE III

Identification, expression and evolutionary analyses of plant lipocalins

Jean-Benoit Frenette Charron, François Ouellet, Mélanie Pelletier, Jean Danyluk, Cédric Chauve, and Fathey Sarhan

Plant Physiology (2005) 139: 2017-2028

Pour les travaux associés à cet article, j'ai élaboré le design expérimental et effectué: les analyses bioinformatiques, les analyses phylogénétiques et les analyses de localisation cellulaire. La partie technique reliée au PCR quantitatif a été réalisée par Marie Champoux. J'ai, pour ma part, analysé et interprété les données de PCR quantitatif. J'ai aussi rédigé le manuscrit et conçu les figures. François Ouellet, a participé à l'élaboration du design expérimental, à la rédaction et plus particulièrement à l'édition finale du manuscript. Mélanie Pelletier a effectué une partie de l'analyse phylogénétique (maximum likelihood). Cédric Chauve a supervisé Mélanie Pelletier. Jean Danyluk a été une personne ressource lors des analyses d'expression des ARN messagers en plus de fournir les différents échantillons d'ARN totaux utilisés dans cet article.

Résumé

Les lipocalines sont un groupe de protéines caractérisées chez les bactéries et chez les animaux vertébrés et invertébrés. Cependant, très peu de choses sont connues sur les lipocalines de plantes. Nous avons récemment publié nos résultats portant sur le clonage et la caractérisation des premières vraies lipocalines de plantes. Dans la présente communication, nous rapportons l'identification et la caractérisation des lipocalines et protéines "lipocalin-like" de plantes en utilisant une approche intégrée de compilation de données, profilage d'expression, localisation cellulaire, analyses phylogénétiques, et prédictions bioinformatiques. Cette approche a permis de classifier les lipocalines de plantes en deux groupes : « temperature-induced lipocalins » (TILs) et « chloroplastic lipocalins » (CHLs). Deux autres groupes de protéines, violaxanthine dé-époxidase (VDE) et zéaxanthine époxidase (ZEP), possèdent une légère similitude avec les lipocalines TILs et ont été classifiées sous le nom de protéines "lipocalin-like". Les protéines CHLs, VDEs, et ZEPs possèdent des séquences transit qui destinent ces protéines aux chloroplastes. Quant à elles, les protéines TILs sont localisées à la membrane plasmique malgré le fait qu'elles ne possèdent aucun peptide signal défini. Les analyses de PCR quantitatif en temps réel a révélé que l'expression des lipocalines et des protéines "lipocalin-like" de blé (*Triticum aestivum* L.) est régulée par les stress abiotiques et est corrélée avec la capacité de la plante à développer une tolérance au gel. Cette corrélation est supportée par le fait que les lipocalines sont présentes chez l'algue rouge *Porphyra yezoensis*, tolérante à la dessiccation, et chez la levure cryotolérante *Debaryomyces hansenii*. Ceci suggère une association possible entre les lipocalines et les organismes tolérants aux stress. Considérant les propriétés des lipocalines, leur spécificité tissulaire, leur implication dans la réponse aux stress de température et leur association avec le chloroplaste et la membrane plasmique des tissues verts, nous formulons l'hypothèse que les lipocalines de plantes possèdent une fonction de protection de l'appareil photosynthétique contre les stress de température. Les analyses phylogénétiques suggèrent que les lipocalines TILs chez les plantes supérieures proviennent d'un gène bactérien présent chez un eucaryote unicellulaire primitif. Cependant, CHLs, VDEs et ZEPs ont probablement évolué d'un gène ancestral provenant d'une cyanobactérie après la formation de l'endosymbiote duquel le chloroplaste origine.

Mots clés : apolipoprotéine D; *Arabidopsis thaliana*; cycle des xanthophylles; lipocaline; membrane plasmique; tolérance au gel; *Triticum aestivum* L.

Abstract

Lipocalins are a group of proteins that have been characterized in bacteria, invertebrate and vertebrate animals. However, very little is known about plant lipocalins. We have previously reported the cloning of the first true plant lipocalins. Here we report the identification and characterization of plant lipocalins and lipocalin-like proteins using an integrated approach of data mining, expression studies, cellular localization and phylogenetic analyses. Plant lipocalins can be classified into two groups, temperature-induced lipocalins (TILs) and chloroplastic lipocalins (CHLs). In addition, violaxanthin de-epoxidases (VDEs) and zeaxanthin epoxidases (ZEPs) can be classified as lipocalin-like proteins. CHLs, VDEs and ZEPs possess transit peptides that target them to the chloroplast. On the other hand, TILs do not show any targeting peptide but localization studies revealed that the proteins are found at the plasma membrane. Expression analyses by Quantitative Real-Time PCR showed that the expression of the wheat (*Triticum aestivum* L.) lipocalins and lipocalin-like proteins is associated with abiotic stress response and is correlated with the plant's capacity to develop freezing tolerance. In support of this correlation, data mining revealed that lipocalins are present in the desiccation tolerant red algae *Porphyra yezoensis* and the cryotolerant marine yeast *Debaryomyces hansenii*, suggesting a possible association with stress tolerant organisms. Considering the plant lipocalins properties, tissue specificity, response to temperature stress, and their association with chloroplasts and plasma membranes of green leaves, we hypothesize a protective function of the photosynthetic system against temperature stress. Phylogenetic analyses suggest that TIL lipocalin members in higher plants were probably inherited from a bacterial gene present in a primitive unicellular eukaryote. On the other hand, CHLs, VDEs and ZEPs may have evolved from a cyanobacterial ancestral gene after the formation of the cyanobacterial endosymbiont from which the chloroplast originated.

Introduction

Lipocalins are an ancient and functionally diverse family of mostly extracellular proteins found in bacteria, protoctists, plants, arthropods, and chordates (Suzuki et al., 2004). They have been implicated in many important functions such as modulation of cell growth and metabolism, binding of cell-surface receptors, nerve growth and regeneration, regulation of the immune response, smell reception, cryptic coloration, membrane biogenesis and repair, induction of apoptosis, animal behavior and environmental stress response (Akerstrom et al., 2000; Bishop, 2000; Frenette Charron et al., 2002).

The lipocalin fold is a highly symmetrical all-ß structure dominated by a single eight-stranded antiparallel ß-sheet closed back on itself to form a continuously hydrogen-bonded ß-barrel. This ß-barrel encloses a ligand-binding site composed of both an internal cavity and an external loop scaffold (Flower et al., 2000). The structural diversity of cavity and scaffold gave rise to a variety of different binding specificities, each capable of accommodating ligands of different size, shape and chemical character (Flower et al., 2000). Lipocalins generally bind small hydrophobic ligands such as retinoids, fatty acids, steroids, odorants and pheromones, and interact with cell surface receptors (Flower, 2000; Flower et al., 2000).

Phylogenetic analyses of lipocalins are possible due to the highly conserved three-dimensional structure (Ganfornina et al., 2000). Three structurally conserved regions (SCRs) related to features of the ß-barrel are conserved: SCR1 (strand A and the 3_{10}-like helix preceding it), SCR2 (portions of strands F and G, and the loop linking them) and SCR3 (portion of strand H, the beginning of the following helix and the loop in-between). It has been suggested that bacterial lipocalins were inherited by unicellular eukaryotes and then passed on to both plants and metazoans (Bishop, 2000). According to this hypothesis, primitive metazoans spread a low number of ancient lipocalins into some of their successors, the arthropods and chordates. These primordial lipocalins were likely similar to the Lazarillo and ApoD

proteins. Alongside the chordate radiation, the ApoD-like ancestral lipocalin suffered duplications. On one hand, it gave rise to the ancestor of retinol-binding proteins (RBPs), and on the other hand, to one or more ancestors of all other paralogous groups of lipocalins that diverged into current chordate lipocalins (Sánchez et al., 2003).

Although the evolution of metazoan lipocalins is well documented (Ganfornina et al., 2000; Gutiérrez et al., 2000; Salier, 2000; Sánchez et al., 2003), very little is known of the evolution of their plant couterparts. The first evidence of the presence of putative plant lipocalins was reported by Bugos et al. (1998). These are violaxanthin de-epoxidases (VDEs) and zeaxanthin epoxidases (ZEPs), key enzymes involved in the biosynthesis of the xanthophyll pigments required for photoprotection of the photosynthetic apparatus. They share the common substrate antheraxanthin and are therefore believed to exhibit similar tertiary structure. However, the peculiar architecture of these two proteins raised doubt as of their true belonging to the lipocalin family (Ganfornina et al., 2000; Salier, 2000).

We have recently reported the identification of the first true plant lipocalins from wheat and *Arabidopsis* (Frenette Charron et al., 2002). The two cDNAs designated *TaTIL* for *Triticum aestivum* L. temperature-induced lipocalin and *AtTIL* for *Arabidopsis thaliana* temperature-induced lipocalin encode polypeptides of 190 and 186 amino acids respectively. Structure analyses indicated the presence of the three typical SCRs that characterize lipocalins. Sequence analyses revealed that these first true plant lipocalins share similarities with three evolutionarily-related lipocalins: the mammalian apolipoprotein D (ApoD), the bacterial lipocalin (Blc) and the insect Lazarillo protein. The comparison of the putative tertiary structures of the human ApoD and the wheat *Ta*TIL-1 suggests that the two proteins differ in membrane attachment and ligand interaction.

To further identify and characterize other plant lipocalins and study their putative functions, we used an integrated approach of data mining of EST databases, bioinformatic predictions, structural features, cellular localization, expression,

phylogenetic and comparative genomics analyses. These analyses revealed that plants possess proteins that can be classified as true lipocalins (TILs and the chloroplastic CHL lipocalins) and lipocalin-like proteins (VDEs and ZEPs). The features and evolutionary origin of these proteins in plants are discussed.

Results

Identification of TaTIL homologs

The recently identified wheat lipocalin *Ta*TIL-1 was used to search GenBank databases. The search revealed that plants possess several homologs of this protein. A combination of EST sequencing and *in silico* reconstruction allowed the generation of 45 complete *Ta*TIL-related protein sequences from plants (Table I and Supplemental Table IV). Based on size, structure, the presence of the 3 SCRs and sequence similarity, these proteins were clustered into two distinct groups, TILs and CHLs. Thirty seven TIL members sharing over 57% identity and 70% overall similarity with *Ta*TIL-1 are found in 25 different species. Wheat possesses two different TIL members, TIL-1 and TIL-2, which share 67% identity and 79% similarity. A short region at the N-terminus differentiates these two members in monocot species but is absent in dicots. Data mining of the rice genome revealed the existence of genes encoding TIL-1 and TIL-2 members (on chromosomes 2 and 8 respectively), whereas the *Arabidopsis thaliana* genome only has TIL-1 (on chromosome 5). In total, twelve plant species contain two TIL members. Sequence analyses revealed that *TIL* genes encode proteins ranging from 179 to 201 amino acids with a calculated molecular mass of 19 to 23 kDa. All TIL homologs show a conserved putative N-glycosylation site. DGPI, PSORT, and SignalP predict a putative C-terminal cleavage site in 8 out of the 37 proteins (Fig. 1, Supplemental Fig. 1). Considering this cleavage site, the calculated molecular mass of the mature TIL proteins would be 2 kDa shorter than the corresponding precursor.

The second group, CHLs, was found in 8 species and shares 25% identity and 35% similarity with *Ta*TIL-1 (Table I). CHL homologs encode proteins ranging from 328 to 340 amino acids with a calculated molecular mass of 36 to 39 kDa. TargetP and ChloroP predict N-terminal chloroplastic transit peptides with high scores (over 0.800) (Fig. 1, Supplemental Fig. 2). However, those transit peptides do not show any

conservation in cleavage site position nor in length (17 to 68 amino acids). Pairwise sequence alignments (Supplemental Fig. 2) predict chloroplast transit peptide cleavage sites near the beginning of SCR1 in both monocot and dicot sequences. CHL homologs also possess 8 conserved cysteine residues probably involved in the three-dimensional structure of the protein by forming disulfide bridges (Fig. 1B, Supplemental Fig. 2).

Lipocalin-like proteins: Violaxanthin De-Epoxidases and Zeaxanthin Epoxidases

It has been suggested that xanthophyll cycle enzymes are lipocalin members (Bugos et al., 1998; Hieber et al., 2000). However, no ESTs corresponding to VDEs or ZEPs were found using *Ta*TIL-1 as query. VDE and ZEP protein sequences from *Arabidopsis* were thus used to search GenBank databases. This search identified 8 and 12 complete sequences corresponding to VDEs and ZEPs, respectively (Fig. 1, Supplemental Figs. 3 and 4). VDE and ZEP sequences share less that 15% similarity with *Ta*TIL-1. VDE homologs encode proteins ranging from 446 to 478 amino acids with a calculated molecular mass of 50 to 55 kDa (Table I). Each of the 8 VDE homologs possesses an N-terminal transit peptide that targets the protein to the chloroplast (Fig. 1B, Supplemental Fig. 3). Considering the cleavage of the transit peptide, the calculated molecular mass of the mature VDE proteins would be in the range of 39 to 40 kDa. VDE homologs show a conserved putative N-glycosylation site and 14 conserved cysteine residues. Of those 14 cysteine residues, 11 form a cysteine-rich region in the N-terminal portion. The C-terminal portion contains 47% charged residues, most of which being glutamic acid residues forming a glutamic acid-rich region. All VDE proteins possess the first lipocalin signature SCR1 next to the cysteine-rich region. In 6 of the 8 VDE sequences, SCR1 fits the consensus while 3 VDE sequences show discrepancies according to the Prosite database. All VDEs exhibit the two invariant amino acids G and W that are key features of SCR1 (Flower et al., Sansom 2000). SCR3 is also found in the glutamic acid-rich C-terminal region,

and the conserved R residue that characterizes this fingerprint is conserved. SCR2 is not present in the VDE sequences.

ZEP homologs encode proteins ranging from 626 to 763 amino acids with a calculated molecular mass of 68 to 80 kDa (Table I). As for the VDE proteins, they possess an N-terminal transit peptide that targets the protein to the chloroplast (Fig 1, Supplemental Fig. 4). After the cleavage of the transit peptide, the calculated molecular mass of the mature ZEP proteins ranges from 60 to 72 kDa. ZEP homologs possess a conserved putative N-glycosylation and 2 conserved cysteine residues. In addition, ZEP proteins contain an ADP-binding domain in their N-terminal portion and an FAD-binding domain in their C-terminal portion (Marin et al., 1996). The two invariant amino acids G and W that are key features of SCR1 are also present. ZEPs differ from TIL and CHL lipocalins in that they do not possess SCR2 and SCR3. ZEP proteins show 28% identity and 44% similarity with mono-oxygenases and oxydases that contain ADP-binding and FAD-binding domains found in bacteria and cyanobacteria (Supplemental Fig. 5).

Localization of the TIL-1 lipocalin

No targeting peptide was found in TIL-1. However, Blc and Lazarillo are known to be anchored to the plasma membrane (PM; Bishop, 2000). We therefore performed transient expression analysis of GFP fusion proteins in onion epidermal cells to establish the subcellular location of the *At*TIL-1 protein. To assess for a possible effect of the GFP moiety on subcellular localization, three constructs were generated (Fig. 2A). The results show that the GFP::TIL fusion accumulates specifically at the PM (Fig. 2B). The two other constructs showed the same localization pattern (data not shown). In contrast, the fluorescence is visible throughout the cell when the GFP protein is in its native state (negative control) (Fig. 2B). These data show that TIL proteins accumulate at the PM.

The *At*TIL localization result obtained by transient expression in onion cells was confirmed by biochemical fractionation in wheat. The immunoblot results in Fig. 2C show that cold acclimation (CA) induces a high accumulation of *Ta*TIL-1 in an enriched PM fraction of cold-acclimated wheat but not in nuclei. The protein is also detected in a total soluble extract but at a lower level.

Expression studies

Induction by abiotic stresses

Expression analyses of the wheat lipocalin genes was carried out using quantitative real-time PCR. The data show that a low temperature (LT) treatment induces the accumulation of the *TaTIL-1*, *TaTIL-2* and *TaZEP* transcripts in both less tolerant (Manitou) and hardy wheat (Norstar) (Fig. 3A). This increase is greater in the hardy winter cultivars. *TaCHL* and *TaVDE* transcripts also accumulate during CA but only in the tolerant wheat. To determine whether plant lipocalin genes are regulated by other stresses, plants were subjected to different treatments. Results in Figure 3B show that a heat shock induces *TaTIL-1* expression while it represses *TaCHL* and *TaVDE*. There is no significant change in response to water or salt stress. The *TaTIL-1* and *TaCHL* transcripts accumulate differentially in various wheat cultivars showing different levels of freezing tolerance (FT), indicating that their expression is associated with the plant's capacity to develop FT (Fig. 4).

TaTIL-1, *TaTIL-2*, *TaCHL*, *TaVDE* and *TaZEP* transcripts all accumulate in response to CA in green leaves (Fig. 5). The maximal accumulation is seen after 6 days of CA for *TaCHL* and 36 days for *TaTIL-1* and *TaZEP*. When the plants are deacclimated at 20°C for 1 and 5 days, all transcripts decline to the non-acclimated control levels. *TaTIL-1* and *TaTIL-2* transcripts accumulate to higher level in crown compared to leaves after 36 days of CA. The results also show that *TaTIL-2* is the only wheat lipocalin expressed in roots.

Regulation during the diurnal cycle

Recent evidence has emerged on the regulation of stress-regulated gene expression by the circadian clock (Fowler et al., 2005). In addition, Thompson et al. (2000) reported that the expression of *ZEP* genes in tomato is under the control of circadian regulation. Based on the classification of ZEP proteins as lipocalin-like proteins, we thus performed expression analyses to determine if oscillation in transcript accumulation of these genes occur during a 16 h light / 8 h dark regime at 20°C (Fig. 6A). *TaZEP* transcripts accumulate to lower levels during dark periods while they accumulate up to 15-fold in the presence of light, reaching a maximum after 4 h near diurnal time 12:00. This oscillation was observed over three cycles. *TaTIL-1* also demonstrated a diurnal oscillation over three cycles. However, the oscillation is less pronounced and reaches a maximal accumulation level at diurnal time 00:00 and a minimal accumulation level at diurnal time 16:00. *TaTIL-2, TaCHL* and *TaVDE* transcripts accumulation is not under the control of a diurnal regulation. To evaluate the effect of LT on the diurnal oscillation, plants were exposed to 4°C under the same light/dark regime (Fig. 6B). Upon exposure to LT, the diurnal oscillation of *TaZEP* transcript accumulation is deregulated.

Evolution of lipocalins

To investigate the evolutionary origin of plant lipocalins, we searched for homologs in ancient plants and algae. Data mining of non-redundant sequence databases, EST databases and other genome projects showed that sequences encoding homologs of TILs and CHLs are found in ancient plants like mosses, coniferals, gnetales and cycads (Table I). No entries encoding TIL or CHL homologs were found for the green algae *Chlamydomonas reinhardtii* or the red algae *Cyanidioschyzon merolae*. However, three plant lipocalin-related ESTs were identified in the red algae *Porphyra yezoensis*. A survey of 14 cyanobacterial genome project databases

revealed that the cyanobacterium *Gloeobacter violaceus* PCC 7421 is the only cyanobacterial strain that possesses a lipocalin gene. A search of 31 fungi genomes revealed lipocalin homologs in two different fungi, the yeast *Debaryomyces hansenii* CBS767 and the foliar plant pathogen *Magnaporthe grisea* strain 70-15.

The relationships between plant lipocalins, ancient lipocalins and other family members were determined by building a phylogenetic tree (Fig. 7). Our goal was not to redesign the evolution scheme of the lipocalin family but to trace the origin of plant lipocalins. We therefore used the dataset from Ganfornina et al. (2000) and appended the plants, algae, cyanobacteria and fungi sequences (this study) and the newly identified epididymal lipocalin sequence (Suzuki et al., 2004). To reduce the complexity, we removed closely-related sequences from the original alignment of Ganfornina et al. (2000). However, each of the 14 clades was represented. We thus aligned 84 lipocalin sequences and reconstructed phylogenetic trees using the Neighbor-Joining method (Fig. 7A) and the ML-based method (Fig. 7B). For the latter, we first computed a global tree (Supplemental Fig. 9) and then refined the part of the tree containing the new sequences using the corresponding subset of the initial alignment and the same phylogenetic reconstruction methodology.

In comparison with the original alignment by Ganfornina et al. (2000), our alignment contains long gaps due to the presence of ZEP and VDE sequences and the lower conservation of the SCR2 and SCR3 signatures. Both trees obtained from this alignment are supported by strong bootstrap values and agree well with the 14 major lipocalin clades already identified (Supplemental Figs. 8 and 9; Ganfornina et al., 2000; Sánchez et al., 2003). The branching pattern suggests that the plant TILs, the yeast *Dh*LIP, the cyanobacterium *Gv*Blc and the red algae *Py*LIP diverged early from bacterial lipocalins (ML bootstrap values of 819 and 700). This is in agreement with a previous phylogenetic study that included a plant lipocalin (Sánchez et al., 2003). The fungus *Mg*LIP and plant CHL lipocalins are incorporated into clade II along with the insect Lazarillo and the mammalian ApoD. The two lipocalin-like groups, VDEs and ZEPs, are in clade XII with a1GP lipocalins. The a1GP protein has been described as

an outlier lipocalin due to the lower conservation of motifs SCR2 and SCR3 (Ganfornina et al., 2000). As these motifs are not well conserved in VDEs and ZEPs, it is not surprising to see these three proteins in the same clade. However, the branching pattern inside this clade is not supported by high bootstrap values in both trees. It is worth noting that exclusion of the VDE and ZEP sequences from the phylogenetic analysis results in the positioning of clade XII near clade XIII in the trees, as reported by Ganfornina et al. (2000). Apart from this, the branching patterns of the trees are not affected.

Another small difference between our analyses and those of others (Ganfornina et al., 2000; Gutiérrez et al., 2000; Sánchez et al., 2003; Suzuki et al., 2004) is that in the ML tree, the two lipocalins Hsap.Lcn5 and Ggal.QS-21 are relocated to the miscellaneous clade that already contains Mmus.Lcn11, Hsap.Lcn9 and Lviv.ESP. In the NJ tree, only Ggal.QS-21 is relocated to this clade.

Discussion

We recently reported the identification of the first true lipocalins from plants, *TaTIL-1* and *AtTIL*, which possess the three SCRs that characterize lipocalins (Frenette Charron et al., 2002). Data mining of various databases using the *Ta*TIL-1 sequence as query resulted in the identification of all available full length plant lipocalins. Protein sequence alignments revealed that these proteins can be classified into four groups based on structural features conserved among typical lipocalins. Two of these groups, TILs and CHLs, are bonafide lipocalins. Monocotyledonous species possess genes encoding two different members of the TIL group, *TIL-1* and *TIL-2*, which are regulated by abiotic stresses. On the other hand, there is no conclusive evidence of the existence of these two forms in dicotyledonous plants. Members of the *CHLs* group are expressed specifically in photosynthetic tissues of higher plants in response to LT exposure. The presence of a transit peptide at their N-terminus suggests that they may play a role in the chloroplast during CA.

*Ta*TIL-1 and *At*TIL proteins share similarity with three evolutionarily related lipocalins, ApoD, Blc and Lazarillo (Frenette Charron et al., 2002). Since the latter two proteins are known to be anchored to membranes, we hypothesized that TILs were also membrane-associated. Our localization studies showed that TIL-1 is indeed localized at the PM level. This result is supported by proteomic analyses of PM proteins from *Arabidopsis* (Kawamura and Uemura, 2003). TILs do not bear a signal peptide, therefore bioinformatic analyses were used to determine which type of attachment is responsible for the PM localisation. These analyses suggested the presence of a C-terminal cleavage site and a favourable environment for the addition of a GPI anchor (proper hydrophobic tail length and hydrophilic region length) in 8 of the 37 reconstructed TIL proteins. Addition of a GPI anchor would result in the cleavage of the C-terminal end of TILs. To determine if this is the case, a C-terminal TIL::GFP fusion was tested by transient expression. The addition of a GPI anchor would result in the cleavage of the TIL::GFP fusion in two separate proteins, TIL and

GFP, and GFP would be able to move freely in the cytoplasm. Our results demonstrate that TILs are associated with the PM but not via a GPI anchor since the GFP fluorescence is always observed at the PM level. The fact that the N-terminal, internal and C-terminal GFP fusions are localized at the PM suggests that TILs could be targeted to this site via the hydrophobic loop between β-strands 5 and 6, as we suggested previously (Frenette Charron et al., 2002).

Despite the presence of SCRs in members of the other two groups, VDEs and ZEPs, many questions have been raised on their true belonging to the lipocalin family. The size and the exon–intron architecture of these xanthophyll cycle enzymes show no significant similarity to the genomic organization of typical lipocalin genes (Gutiérrez et al., 2000; Salier, 2000). VDEs are predicted to be lipocalin-like proteins with a central barrel structure flanked by a cysteine-rich N-terminal domain and a glutamate-rich C-terminal domain (Fig.1; Hieber et al., 2002). ZEPs possess ADP and FAD-binding domains and only fit the description of lipocalins based on a low SCR1 similarity (Fig.1). On the other hand, the 44% sequence similarity with mono-oxygenases would instead classify ZEP proteins in the latter family. According to our phylogenetic analyses, VDEs and ZEPs are positioned in clade XII together with a1GP. Since a1GP is found only in marsupials and placental mammals, it is unlikely that this grouping reflects a genuine evolutionary relationship. Given the features of VDEs and ZEPs, the strict definition of lipocalins and their positioning in the phylogenetic trees, it is difficult to consider them part of the lipocalin family. Rather, they could be classified as lipocalin-like proteins. The apparent fusion of a true plant lipocalin with other proteins during evolution was proposed to explain the atypical structures of these enzymes (Ganfornina et al., 2000). The appearance of proteins with novel functions would have been an evolutionary advantage in that it would have provided plants with enhanced protection against photooxidative damages.

Important clues regarding the evolution of plant lipocalins in our study come from the finding of a lipocalin homolog in the cyanobacterium *Gloeobacter violaceus*

PCC 7421. Cyanobacteria are unicellular organisms that carry a complete set of genes for oxygenic photosynthesis, the most fundamental life process on earth. The chloroplasts in higher plants are believed to have evolved from cyanobacterial ancestors who developed an endosymbiontic relationship with a eukaryotic host cell (Delwiche et al., 1995). To this day, sequence information is available for 14 complete and 2 partial genomes of cyanobacteria. Among these, only *G. violaceous* possesses a lipocalin gene. Unlike most recent cyanobacteria, this strain lacks thylakoids, and phycobilisomes are attached to the PM. Recent molecular phylogenetic analyses show that the *G. violaceus* is a member of early branching of the cyanobacterial lineage (Delwiche et al., 1995) and could thus be the oldest known cyanobacterium. This suggests that *G. violaceous* or a close relative might have been the initial donor that gave rise to the chloroplast structures of higher plants. These observations reveal that certain lipocalins were associated with photosynthetic membranes early in the evolution.

No TIL homologs were found in the primitive photosynthetic green algae *Chlamydomonas reinhardtii* nor in the red algae *Cyanidioschyzon merolae*, two species for which extensive genomic information is available. However, a homolog was found in the red algae *Porphyra yeonisis*. Red algae (*Rhodophyta*) are photoautotrophic eukaryotes characterized by a lack of flagella and the presence of phycobiliproteins within the plastid (Bold and Wynne, 1985; South and Whittick, 1987). *Porphyra* species are blade-forming red seaweeds and are among the simplest of red algae. Some are extremely tolerant to dessiccation and are found in the highest, driest reaches of the littoral zone in cold temperate and boreal regions. The presence of a lipocalin in this species may be related to its desiccation tolerance. The close positioning, in the phylogenetic trees, of the cyanobacterial *Gv*Blc with *Py*Lip from a photosynthetic red algae supports the hypothesis of lateral transfer.

Another novel finding is the identification of lipocalins in two fungi species, *D. hansenii* and *M. grisea*. *D.* (*Torulaspora*) *hansenii* is a cryotolerant marine yeast that tolerates salinity levels up to 24%, whereas common yeast growth is inhibited at

10% salinity. *D. hansenii* is the most common species found in all types of cheeses (Fleet, 1990). It is also common in other dairy products (Seiler and Busse, 1990) because of its ability to grow in the presence of high salt at LT and to metabolize lactic and citric acids. *M. grisea*, the causal agent of rice blast disease, is one of the most devastating threats to food security worldwide (Zeigler et al., 1994). *M. grisea* shows excellent adaptation abilities to a wide spectrum of stresses (Ikeda et al., 2001). The presence of lipocalins in these fungi species may explain their abiotic stress tolerance. It is possible that chloroplast-associated lipocalins and lipocalin-like proteins could have arisen from a lateral transfer following infection by a fungus such as *M. grisea*. It has been suggested that such transfers could explain the presence of *M. grisea* DNA in plant genomes (Kim et al., 2000). On the other hand, gene duplication and/or fusion can also be proposed to explain the presence of the different proteins in higher plants genomes.

The plant lipocalins and lipocalin-like proteins properties, their tissue specificity and their transcript accumulation in response to temperature stress suggest a possible protection role against stress damage. Their association with the chloroplast (CHLs, VDEs and ZEPs) and the plasma membrane (TILs) in the green leaves supports the idea that these proteins may act as a scavenger of potentially harmful molecules known to be induced by temperature stress and excess light. The lipocalin-like proteins VDEs and ZEPs catalyze the interconversions between the carotenoids violaxanthin, antheraxanthin and zeaxanthin in higher plants under stress condition to form the zeaxanthin that protects the photosynthetic apparatus against the effect of excessive light (Havaux and Kloppstech, 2001). Our previous work demonstrated that the photosynthetic acclimation to LT mimics the photosynthetic acclimation to high light because both conditions result in a comparable reduction state of photosystem II (Ndong et al., 2001). Based on this comparison and the data in the present report, we hypothesize that the other plant lipocalins and the CHLs in particular may protect the photosynthetic system against the deleterious effect of

temperature stress. The work is in progress to determine the exact function of these novel members of plant lipocalins.

Materials and methods

Data mining

*Ta*TIL homologs were identified using the *Ta*TIL-1 protein sequence as query (Acc. No. AAL75812) using TBLASTN against the GenBank EST database. Overlapping ESTs were assembled using the CAP3 Assembly software (http://fenice.tigem.it/bioprg/interfaces/cap3.html) and a consensus cDNA sequence was deduced when three or more identical sequences could be aligned. The degree of sequence identity was determined using ALIGN on the Biology Workbench (http://workbench.sdsc.edu/) and NCBI BLAST 2 sequences (http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). Sequences were aligned and analyzed using ClustalW on the Biology Workbench. Shading of amino acids was performed with BOXSHADE (http://ulrec3.unil.ch/software/boxshade/boxshade.html). PSORT, iPSORT (http://psort.nibb.ac.jp/), TargetP v1.01 (http://www.cbs.dtu.dk), SignalP v2.0 (http://www.cbs.dtu.dk) and ChloroP (http://www.cbs.dtu.dk) were used to detect specific targeting sequences. For functional domain identification, we first used ScanProsite to scan the Prosite database then most of the software available on the ExPASy server (http://ca.expasy.org). DGPI was used for GPI-anchoring site prediction (http://129.194.186.123/GPI-anchor/index_en.html).

Plant material and growth conditions

In this study, we used two spring wheat genotypes (*Triticum aestivum* L. cv Glenlea, LT_{50} (lethal temperature that kills 50% of the seedlings) -8°C; and cv Concorde, LT_{50} -8°C), and four winter wheat genotypes (*T. aestivum* L. cv Monopole, LT_{50} -15°C; cv Absolvent, LT_{50} -16°C; cv Fredrick, LT_{50} -16°C; and cv Norstar, LT_{50} -19°C). Plants were grown in a mixture of 50% black earth and 50% Pro-Mix (Premier) for 7 days

under 16 h day photoperiod with a light intensity of 250 μmol m^{-2} s^{-1} at 20°C. Heat shock (40°C) and cold treatments (4°C) were performed by changing the temperature in the growth chamber while salt stress (0.3 M NaCl) and osmotic stress (30% (w/v) PEG-6000) were performed by saturating the soil with these solutions. A total of 8 seedlings were harvested on dry ice at different time points as stated in the figures and immediately frozen at -70°C.

Cellular localization of TILs

Transient expression of GFP fusions

AtTIL cDNA fragments were PCR-amplified using the primers described in Supplemental Table I then cloned in the pAVA321 vector (von Arnim et al., 1998) to generate 3 constructs. The chimeric genes encode GFP::TIL and TIL::GFP fusion proteins, and another protein in which the GFP protein is inserted within the TIL sequence (TI::GFP::IL) (Fig. 2A). Plasmid DNA was coated onto M17 tungsten particles (Bio-Rad) and delivered into onion epidermal cells by particle bombardment (Shieh et al., 1993). Images were captured on a MRC1024 confocal system with a Nikon Eclipse TE300 inverted microscope and analyzed using the LaserSharp software (Bio-Rad).

Subcellular fractionation

Organellar protein fractions were prepared from leaves of control and 7-day cold-acclimated winter wheat cv Norstar. Plasma membranes (PM) were isolated by two-phase partitioning as described by Zhou et al. (1994). Nuclei were isolated as previously described (Vazquez-Tello et al., 1998), then nuclear proteins were extracted using the TRI-Reagent® (Molecular Research Center) following the manufacturer's recommendations. Total soluble proteins were prepared as described

(Vazquez-Tello et al., 1998). Samples were separated on 12% SDS-PAGE gels, and the rabbit anti-*Ta*TIL-1 antibody was used for the immunoblot analysis. Detection was performed with a peroxidase-coupled anti-rabbit IgG secondary antibody and the Western Lightning Chemiluminescence Reagent Plus® (PerkinElmer).

Expression analyses by Quantitative Real-Time PCR

RNA isolation and cDNA synthesis

Total RNA was isolated using the RNeasy Plant Mini Kit (QIAGEN). For the expression analyses in the different wheat cultivars, RNA was separated on a formaldehyde agarose gel, transfered to a positively-charged nylon membrane then hybridized sequentially to *TaTIL* and *TaCHL* ^{32}P-labeled probes. All other expression analyses were performed using quantitative real-time PCR. Purified RNA (2.8 µg) was reverse transcribed in a 20 µL reaction volume using the SuperScript™ II First-Strand Synthesis System for RT-PCR (Invitrogen). Parallel reactions were run for each RNA sample in the absence of SuperScript II (no RT control) to assess for genomic DNA contamination. The reactions were terminated by heat inactivation at 70°C for 15 min. Subsequently, the cDNA products were treated with 2 units of RNase H for 20 min at 37°C, then diluted in water to 20 ng μl^{-1} and stored at –20°C.

Design of gene-specific primers

The genome of hexaploid wheat contains 3 genomes inherited from 3 diploid ancestors. Primers were specifically designed to monitor the expression of the 3 copies of each gene in the same RT reaction. In addition, primers for *TaCHL*, *TaVDE* and *TaZEP* were designed onto exon junctions to avoid genomic DNA amplification. The gene architecture of *TaTIL-1* and *TaTIL-2* did not allow for the design of LUX™ primers on the exon-exon junction. Fluorescent LUX™ primers as well as non-

fluorescent primers (Supplemental Table I) were designed using a combination of the LUX™ Designer-Desktop Version (Invitrogen) and the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). BLASTN searches were performed to confirm the gene specificity of the primers. Primers were synthesized by Invitrogen.

PCR amplification

Quantitative Real-Time PCR assays were performed in quadruplicate on an ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) using the 18S ribosomal RNA as internal standard. From the diluted cDNA, 1 µL (20 ng) was used as template in a 50 µL PCR reaction containing 1x Platinum Quantitative PCR SuperMix-UDG, 0.15 µM of non-fluorescent primer, 0.3 µM of LUX™ fluorescent primer, and ROX reference dye. The PCR thermal cycling parameters were 50°C for 2 min, 95°C for 2 min followed by 50 cycles of 95°C for 20 s and 60°C for 1 min. Each experiment was replicated at least three times.

Data analysis

All calculations and statistical analyses were performed using the SDS RQ Manager 1.1 software using the $2^{-\Delta\Delta Ct}$ method with a RQ Min/Max confidence set at 95 % (Livak and Schmittgen 2001). The error bars display the calculated maximum (RQMax) and minimum (RQMin) expression levels that represent standard error of the mean expression level (RQ value). Collectively, the upper and lower limits define the region of expression within which the true expression level value is likely to occur (SDS RQ Manager 1.1 software user manual; Applied Biosystems Inc.). Amplification efficiency (98% to 100%) for the six primer sets was determined by amplification of cDNA dilution series using 80, 20, 10, 5, 2.5 and 1.25 µg per reaction (data not

shown). Specificity of the RT-PCR products was assessed by gel electrophoresis. A single product with the expected length was detected for each reaction.

Phylogenetic analyses

Proteins used in this analysis and the FASTA files are presented in Supplemental Tables II and III. The ClustalX v-1.83 software (Thompson et al., 1997) was used to generate the sequence alignment using the following parameters: gap opening penalty of 15.0, gap extension penalty of 0.30, and the substitution Gonnet scoring matrix. The alignment was adjusted manually in order to respect the position of the three SCRs, the glycosylation sites and the cysteine residues, then used to generate phylogenetic trees based on two methods. The Neighbor-Joining (NJ) tree was generated using TREECON for Windows (version 1.3b) with a 100 bootstrap replicates and with the distance calculation set to the Poisson correction (van de Peer and Wachter, 1994). PHYML (Guindon and Gascuel, 2003) was used to perform maximum likelihood (ML) analyses with the evolution model JTT (Jones et al., 1992) and all other parameters set to their default value. One thousand bootstrap replicates were performed with Seqboot and a consensus tree was computed with Consense following a strict-majority rule. The latter two programs are part of the PHYLIP v-3.6 package (Felsenstein, 1993). The ML trees were rooted with the VchoLpro taxon and displayed with TreeView v-1.6.6 (Page, 1996).

Upon request, all novel materials described in this publication will be made available in a timely manner for non-commercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requestor.

Acknowledgments

We thank Drs. Sánchez and Ganfornina (Departamento de Bioquímica y Biología Molecular y Fisiología-IBGM, Universidad de Valladolid-CSIC, Valladolid, Spain) for providing their lipocalin sequence alignment. We also thank M. Champoux, C. Plouffe, G. Brault, N.A. Kane and D. Flipo for technical assistance.

Literature cited

Akerstrom BD, Flower R, Salier JP (2000) Lipocalins: unity in diversity. Biochim. Biophys. Acta 1482: 1-8

Bishop RE (2000) The bacterial lipocalins. Biochim. Biophys. Acta 1482: 73-83

Bishop RE, Penfold SS, Frost LS, Holtje JV, Weiner JH (1995) Stationary phase expression of a novel *Escherichia coli* outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins. J. Biol. Chem. 270: 23097-23103

Bold HC, Wynne MJ (1985) Introduction to the algae. Structure and Reproduction. Prentice Hall, Inc., Englewood Cliffs

Bugos RC, Hieber AD, Yamamoto HY (1998) Xanthophyll cycle enzymes are members of the lipocalin family, the first identified from plants. J. Biol. Chem. 273: 15321-15324

Delwiche CF, Kuhsel M, Palmer JD (1995) Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. Mol. Phylogenet. Evol. 4: 110-128

Felsenstein J (1993) PHYLIP, Phylogeny Inference Package, Version 3.6. Distributed by the Author. Department of Genetics, University of Washington, Seattle.

Fleet G (1990) Yeasts in dairy products. J. Appl. Bacteriol. 68: 199-211

Flower DR (2000) Beyond the superfamily: the lipocalin receptors. Biochim. Biophys. Acta 1482: 327-336

Flower DR, North AC, Sansom CE (2000) The lipocalin protein family: structural and sequence overview. Biochim. Biophys. Acta 1482: 9-24

Fowler SG, Cook D, Thomashow MF (2005) Low temperature induction of *Arabidopsis* CBF1, 2, and 3 is gated by the circadian clock. Plant Physiol. 137: 961-968

Frenette Charron JB, Breton G, Badawi M, Sarhan F (2002) Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. FEBS Lett. 517: 129-132

Ganfornina MD, Gutiérrez G, Bastiani M, Sánchez D (2000) A phylogenetic analysis of the lipocalin protein family. Mol. Biol. Evol. 17: 114-126

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52: 696-704

Gutiérrez G, Ganfornina MD, Sánchez D (2000) Evolution of the lipocalin family as inferred from a protein sequence phylogeny. Biochim. Biophys. Acta 1482: 35-45

Havaux M, Kloppstech K (2001) The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in *Arabidopsis npq* and *tt* mutants. Planta 213: 953-966

Hieber AD, Bugos RC, Verhoeven AS, Yamamoto HY (2002) Overexpression of violaxanthin de-epoxidase: properties of C-terminal deletions on activity and pH-dependent lipid binding. Planta 214: 476-483

Hieber AD, Bugos RC, Yamamoto HY (2000) Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. Biochim. Biophys. Acta 1482: 84-91

Ikeda K, Nakayashiki H, Takagi M, Tosa Y, Mayama S (2001) Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. Mol. Gen. Genet. 266: 318-325

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. CABIOS 8: 275-282

Kawamura Y, Uemura M (2003) Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation. Plant J. 36: 141-154

Kim NS, Park NI, Kim SH, Kim ST, Han SS, Kang KY (2000) Isolation of TC/AG repeat microsatellite sequences for fingerprinting rice blast fungus and their possible horizontal transfer to plant species. Mol. Cells 10: 127-134

Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. *Methods* 25:402–408.

Marin E, Nussaume L, Quesada A, Gonneau M, Sotta B, Hugueney P, Frey A, Marion-Poll A (1996) Molecular identification of zeaxanthin epoxidase of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. EMBO J. 15: 2331-2342

Ndong C, Danyluk J, Huner NPA and Sarhan F (2001) Survey of gene expression in winter rye during changes in growth temperature, irradiance or excitation pressure. Plant Mol. Biol. 45: 691-703

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12: 357-358

Peitsch MC, Boguski MS (1990) Is apolipoprotein D a mammalian bilin-binding protein? New Biol. 2: 197-206

Salier JP (2000) Chromosomal location, exon/intron organization and evolution of lipocalin genes. Biochim. Biophys. Acta 1482: 25-34

Sánchez D, Ganfornina MD, Gutiérrez G, Marín A (2003) Exon–intron structure and evolution of the lipocalin gene family. Mol. Biol. Evol. 20: 775–783

Seiler H, Busse M (1990) The yeasts of cheese brines. Int. J. Food Microbiol. 11: 289-303

Shieh MW, Wessler SR, Raikhel NV (1993) Nuclear targeting of the maize R protein requires two nuclear localization sequences. Plant Physiol. 101: 353-361

South GR, Whittick A (1987) Introduction to Phycology. Blackwell Scientific Publications, London

Suzuki K, Lareyre JJ, Sánchez D, Gutiérrez G, Araki Y, Matusik RJ, Orgebin-Crist MC (2004) Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2. Gene 339: 49-59

Thompson AJ, Jackson AC, Parker RA, Morpeth DR, Burbidge A, Taylor IB (2000) Abscisic acid biosynthesis in tomato: regulation of zeaxanthin epoxidase and 9-cis-epoxycarotenoid dioxygenase mRNAs by light/dark cycles, water stress and abscisic acid. Plant Mol. Biol. 42: 833-845

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882

van de Peer Y, de Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput. Appl. Biosci. 10: 569-570

Vazquez-Tello A, Ouellet F, Sarhan F (1998) Low temperature-stimulated phosphorylation regulates the binding of nuclear factors to the promoter of *wcs120*, a wheat cold-specific gene. Mol. Gen. Genet. 257: 157-166

von Arnim AG, Deng XW, Stacey MG 1998. Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene 221: 35-43

Zeigler RS, Tohme J, Nelson J, Levy M, Correa F (1994) Linking blast population analysis to resistance breeding: A proposed strategy for durable resistance. In RS Ziegler, SA Leong, PS Teng, eds. Rice blast disease. CAB International, Wallingford, pp 267-292

Zhou BL, Arakawa K, Fujikawa S, Yoshida S (1994) Cold-induced alterations in plasma membrane proteins that are specifically related to the development of freezing tolerance in cold-hardy winter wheat. Plant Cell Physiol. 35: 175-182

Table 1. *Nomenclature and characteristics of plant lipocalins and lipocalin-like proteins*

*Ta*TIL homologs were identified using the *Ta*TIL-1 protein sequence as query using TBLASTN against the GenBank EST database. Overlapping ESTs were assembled using the CAP3 assembly software and a consensus cDNA sequence was deduced when three or more identical sequences could be aligned. The degree of sequence identity was determined using ALIGN on the Biology Workbench and the National Center for Biotechnology Information BLAST 2 sequences. Chl, Chloroplastic; ND, not determined; PM, plasma membrane associated.

Name	Class	Species	Accession No.	Precursor Protein	Mature Protein	Subcellular Location	Identity/ Similarity[a]
				kD (Amino Acid)	*kD (Amino Acid)*		
TILs							
*Pp*TIL	Mosses	*Physcomitrella patens*	DQ222991	22 (189)	22 (189)	PM	57%/72%
*Tr*TIL	Mosses	*Tortula ruralis*	DQ223011	21 (186)	21 (186)	PM	59%/70%
*Pt*TIL	Conifers	*Pinus taeda*	DQ222992	21 (182)	21 (182)	PM	62%/77%
*Hv*TIL-1	Monocots	*Hordeum vulgare*	DQ222974	22 (190)	22 (190)	PM	96%/97%
*Hv*TIL-2	Monocots	*H. vulgare*	DQ222978	21 (182)	21 (182)	ND	66%/78%
*Os*TIL-1	Monocots	*Oryza sativa*	XP_466697	22 (195)	22 (195)	PM	86%/91%
*Os*TIL-2	Monocots	*O. sativa*	XP_482610	21 (179)	21 (179)	ND	69%/80%
*So*TIL	Monocots	*Saccharum officinarum*	DQ222989	22 (193)	22 (193)	PM	88%/95%
*Sb*TIL-1	Monocots	*Sorghum bicolor*	DQ222976	23 (201)	23 (201)	PM	86%/92%
*Sb*TIL-2	Monocots	*S. bicolor*	DQ222980	21 (187)	21 (187)	ND	64%/77%
*Ta*TIL-1	Monocots	*Triticum aestivum*	AAL75812	22 (190)	22 (190)	PM	100%
*Ta*TIL-2	Monocots	*T. aestivum*	DQ222977	21 (182)	21 (182)	ND	67%/79%
*Zm*TIL-1	Monocots	*Zea mays*	DQ222975	23 (198)	23 (198)	PM	87%/95%
*Zm*TIL-2	Monocots	*Z. mays*	DQ222979	21 (181)	21 (181)	ND	66%/79%
*At*TIL	Dicots	Arabidopsis	BAB10998	21 (186)	21 (186)	PM	75%/83%
*Bn*TIL	Dicots	*Brassica napus*	DQ222996	21 (187)	21 (187)	PM	70%/81%
*Cs*TIL	Dicots	*Citrus sinensis*	DQ223001	22 (186)	22 (186)	PM	73%/84%
*Gm*TIL	Dicots	*Glycine max*	DQ222990	21 (184)	21 (184)	PM	67%/81%
*Gm*TIL′	Dicots	*G. max*	DQ222982	21 (184)	21 (184)	ND	67%/81%
*Ga*TIL	Dicots	*Gossypium arboreum*	DQ223000	21 (185)	21 (185)	PM	73%/85%
*Ga*TIL′	Dicots	*G. arboreum*	DQ222986	21 (179)	21 (179)	ND	70%/81%
*Ls*TIL	Dicots	*Lactuca sativa*	BQ852119	21 (185)	21 (185)	PM	74%/86%
*Le*TIL	Dicots	*Lycopersicon esculentum*	DQ222988	21 (185)	21 (185)	PM	74%/86%
*Le*TIL′	Dicots	*L. esculentum*	DQ222981	21 (185)	21 (185)	ND	74%/86%
*Mt*TIL	Dicots	*Medicago truncatula*	DQ222994	21 (184)	21 (184)	PM	70%/82%
*Mt*TIL′	Dicots	*M. truncatula*	DQ222983	19 (168)	19 (168)	ND	74%/86%
*Mc*TIL	Dicots	*Mesembryanthemum crystallinum*	DQ222999	22 (187)	22 (187)	PM	72%/83%
*Mc*TIL′	Dicots	*M. crystallinum*	DQ222985	22 (187)	22 (187)	ND	75%/84%
*Pb*TIL	Dicots	*Populus balsamifera*	DQ223002	21 (185)	21 (185)	PM	73%/85%
*Pb*TIL′	Dicots	*P. balsamifera*	DQ222987	21 (185)	21 (185)	ND	74%/85%
*Pot*TIL	Dicots	*Populus tremula*	DQ223003	21 (185)	21 (185)	PM	73%/84%
Pot × *Ptr*TIL	Dicots	*P. tremula* × *Populus tremuloides*	DQ223004	21 (185)	21 (185)	PM	75%/86%
*Pa*TIL	Dicots	*Prunus armeniaca*	DQ222998	21 (185)	21 (185)	PM	70%/84%
*Prp*TIL	Dicots	*Prunus persica*	DQ222997	21 (185)	21 (185)	PM	70%/84%
*St*TIL	Dicots	*Solanum tuberosum*	DQ222995	21 (186)	21 (186)	PM	73%/85%
*St*TIL′	Dicots	*S. tuberosum*	DQ222984	21 (185)	21 (185)	ND	74%/85%
*Vv*TIL	Dicots	*Vitis vinifera*	DQ222994	22 (185)	22 (185)	PM	72%/85%
CHLs							
*Hv*CHL	Monocots	*H. vulgare*	DQ223006	37 (336)	26 (230)	Chl	25%/35%
*Os*CHL	Monocots	*O. sativa*	XP_473969	37 (342)	26 (231)	Chl	26%/36%
*Sb*CHL	Monocots	*S. bicolor*	DQ223005	38 (340)	26 (231)	Chl	25%/34%
*Ta*CHL	Monocots	*T. aestivum*	DQ223009	37 (339)	26 (230)	Chl	26%/36%
*At*CHL	Dicots	Arabidopsis	CAB41869	39 (353)	26 (230)	Chl	25%/40%
*Gm*CHL	Dicots	*G. max*	DQ223010	36 (328)	26 (231)	Chl	26%/36%
*In*CHL	Dicots	*Ipomoea nil*	DQ223007	37 (334)	26 (231)	Chl	29%/41%
*St*CHL	Dicots	*S. tuberosum*	DQ223008	38 (333)	26 (231)	Chl	27%/39%
VDEs							
*Os*VDE_jap	Monocots	*O. sativa* var. *japonica*	AAL09678	50 (446)	40 (351)	Chl	14%
*Os*VDE_ind	Monocots	*O. sativa* var. *indica*	AAF97601	50 (446)	40 (351)	Chl	14%
*So*VDE	Monocots	*S. officinarum*	CAB59211	54 (472)	40 (351)	Chl	13%
*Ta*VDE	Monocots	*T. aestivum*	AAK38177	51 (454)	40 (351)	Chl	13%
*At*VDE	Dicots	Arabidopsis	NP_172331	52 (462)	40 (352)	Chl	14%

(*Table continues on following page.*)

Table I. *(Continued from previous page.)*

Name	Class	Species	Accession No.	Precursor Protein	Mature Protein	Subcellular Location	Identity/ Similarity[a]
*Cs*VDE	Dicots	*C. sinensis*	AAL67858	54 (473)	39 (344)	Chl	13%
*Nt*VDE	Dicots	*Nicotiana tabacum*	AAC50031	55 (478)	40 (347)	Chl	14%
*Ls*VDE	Dicots	*L. sativa*	AAC49373	54 (473)	40 (351)	Chl	14%
ZEPs							
*Cr*ZEP	Algae	*C. reinhardtii*	AAO34404	81 (763)	72 (676)	Chl	10%
*Csp*ZEP	Algae	*Chlamydomonas* sp. *W80*	AAO48941	78 (727)	68 (627)	Chl	9%
*Os*ZEP	Monocots	*O. sativa*	BAB39765	68 (626)	60 (548)	Chl	12%
*At*ZEP_col	Dicots	Arabidopsis	AAM13144	74 (667)	65 (586)	Chl	11%
*At*ZEP_ler	Dicots	Arabidopsis	AAG17703	74 (667)	65 (586)	Chl	11%
*At*ZEP_?	Dicots	Arabidopsis	AAF82390	74 (667)	65 (586)	Chl	11%
*Ca*ZEP	Dicots	*Capsicum annuum*	Q96375	72 (660)	63 (582)	Chl	12%
*Cu*ZEP	Dicots	*Citrus unshiu*	BAB78733	73 (664)	64 (586)	Chl	11%
*Le*ZEP	Dicots	*L. esculentum*	P93236	73 (669)	64 (583)	Chl	11%
*Np*ZEP	Dicots	*Nicotiana plumbaginifolia*	CAA65048	73 (663)	64 (583)	Chl	11%
*Nt*ZEP	Dicots	*N. tabacum*	S69548	73 (663)	64 (583)	Chl	11%
*Pa*ZEP	Dicots	*P. armeniaca*	O81360	72 (661)	63 (580)	Chl	12%

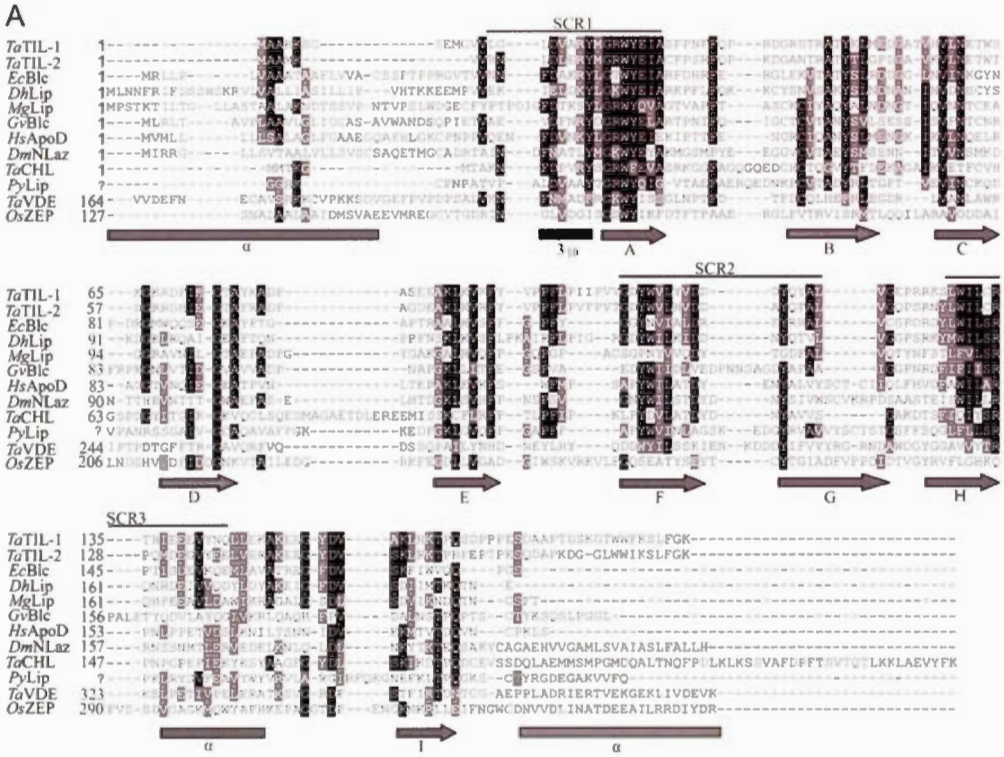
[a]With respect to the wheat *Ta*TIL-1; for VDEs and ZEPs, only the similarity is indicated.

Fig. 1. Structure of plant lipocalins and lipocalin-like proteins.

A. Alignment of the deduced amino acid sequences of wheat lipocalins with a select set of related lipocalins. Identical residues are in black and similar residues are in grey. The three SCRs that provide a signature for lipocalins are indicated. The secondary structure predicted from already published models (Bishop et al., 1995; Peitsh and Boguski, 1996) is presented under the alignment. Arrows, grey rectangles and the black rectangle represent ß-strands, α-helix and the 3_{10} α-helix respectively.

B. Schematic representation of plant lipocalins.

Plant lipocalins can be classified in four groups. White boxes labeled 1, 2 and 3 represent the three structurally conserved regions (SCR1, SCR2, and SCR3) that characterize lipocalins. Black rectangles represent chloroplast transit peptides. Dark grey rectangles with roman numerals represent respectively: I = cysteine-rich region, II = glutamic acid-rich region, III = ADP-binding site and IV = FAD-binding site. Stars and lozenges represent conserved N-glycosylation sites and conserved cysteine residues, respectively.


A
SCR1
SCR2
SCR3
TaTIL-1
TaTIL-2
EcBlc
DhLip
MgLip
GvBlc
HsApoD
DmNLaz
TaCHL
PyLip
TaVDE
OsZEP
α
3₁₀
A
B
C
D
E
F
G
H
α
I
α

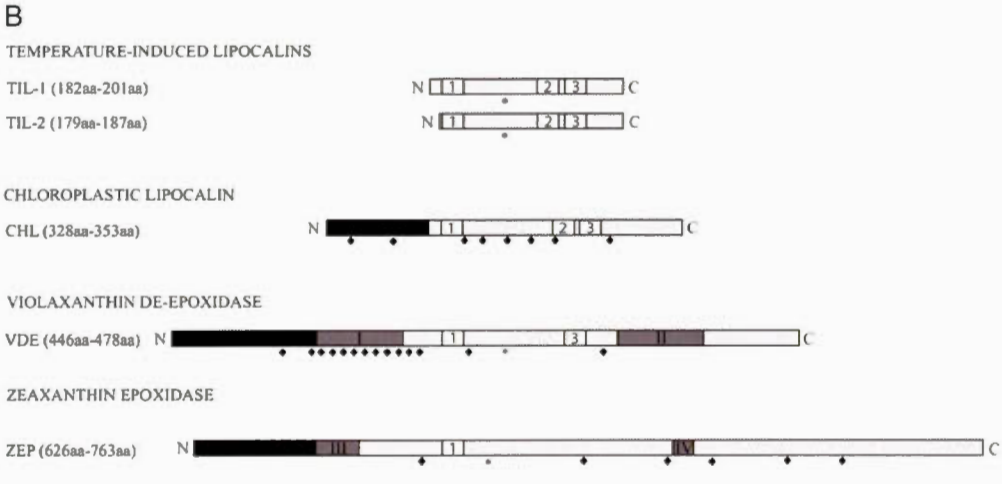

B
TEMPERATURE-INDUCED LIPOCALINS
TIL-1 (182aa-201aa)
TIL-2 (179aa-187aa)
N
C
CHLOROPLASTIC LIPOCALIN
CHL (328aa-353aa)
VIOLAXANTHIN DE-EPOXIDASE
VDE (446aa-478aa)
ZEAXANTHIN EPOXIDASE
ZEP (626aa-763aa)

Fig. 2. Cellular localization of the plant TIL lipocalins.

A. Schematic representation of GFP fusions used in the transient expression experiments. N and C are the amino and carboxy termini of the proteins, respectively; 1, 2 and 3 indicate the 3 structurally conserved regions (SCRs).

B. Transient expression assays of GFP-TIL fusions. Plasmids carrying the fusions were transformed into onion epidermal cells by microprojectile bombardment. Confocal images of GFP fluorescence were captured 20 hours after transformation. Only the GFP::*At*TIL data is shown since the three constructs gave the same fluorescence pattern. The color figure is shown in Supplemental Fig. 6.

C. Biochemical fractionation analysis. Wheat protein extracts were prepared and subjected to SDS-PAGE and western blot analyses. Upper panels, western blot results obtained using the anti-*Ta*TIL antibody (dil. 1/25,000, 10 sec. exposure for the plasma membrane fractions; dil. 1/2,500, 5 min. exposure for the other fractions). Lower panel, Coomassie Brilliant Blue-stained gel showing the quality of the preparations. Typical protein patterns are observed for each fraction. NA, non-acclimated plants grown for 7 days; CA7, plants grown for 7 days at 24°C then cold acclimated at 4°C for 7 days.

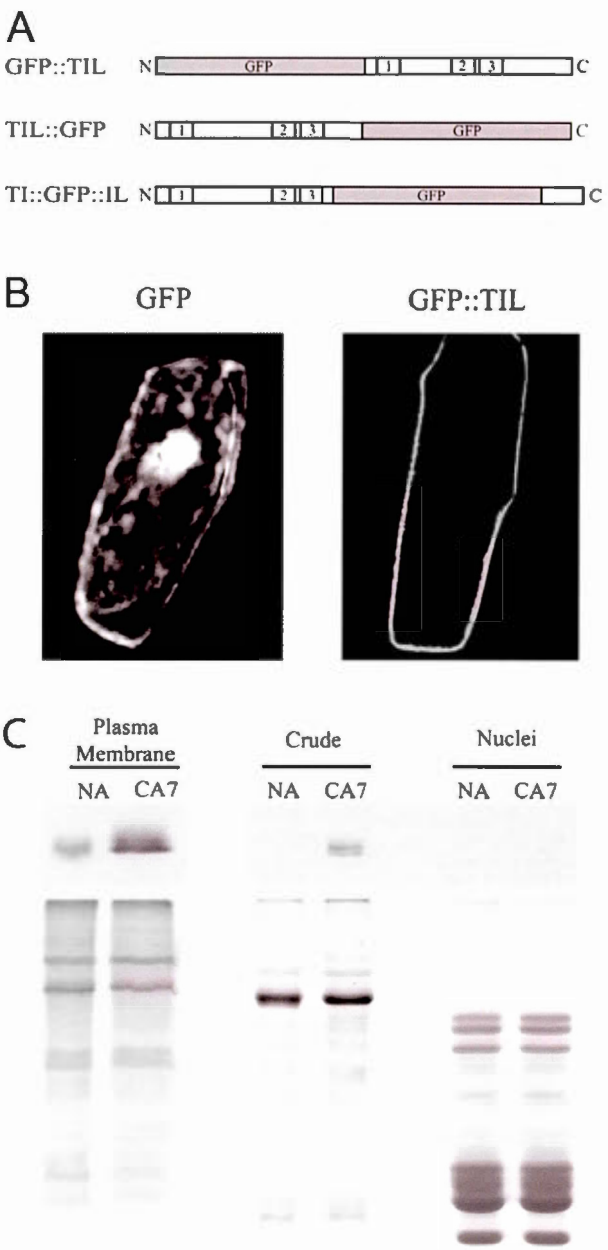

A
GFP::TIL
N
GFP
1
2
3
C
TIL::GFP
N
1
2
3
GFP
C
TI::GFP::IL
N
1
2
3
GFP
C
B
GFP
GFP::TIL
C
Plasma Membrane
NA
CA7
Crude
NA
CA7
Nuclei
NA
CA7

Fig. 3. Expression analysis of wheat lipocalins in response to abiotic stresses.

Plants were treated then total RNA was isolated from leaves, reverse transcribed, and subjected to quantitative real-time PCR. Relative transcript abundance was calculated and normalized with respect to the 18S rRNA transcript level. Data shown represent mean values obtained from four independent amplification reactions (n=4), and the error bars indicate the range of possible RQ values defined by the standard error of the delta Ct's. This experiment was repeated three times with similar results.

A. Expression during cold acclimation of wheat seedlings. Tissues were sampled at 18:00. Spring wheat (*T. aestivum* L. cv Manitou; LT_{50} (temperature that kills 50% of the seedlings) of -8°C) and winter wheat (*T. aestivum* L. cv Norstar; LT_{50} -19°C) were cold acclimated at 4°C for the indicated number of days. NA, non-acclimated plants grown for 7 days; CA7 to CA56, plants cold acclimated for 7 to 56 days.

B. Expression following exposure of wheat seedlings to abiotic stresses. Winter wheat (*T. aestivum* L. cv Norstar; LT_{50} -19°C) plants were treated as follows: WS, dehydrated in a 30% PEG solution (water stress); HS, exposed to 40°C for 1 h (heat shock); NaCl, treated with 300 mM NaCl for 12 and 22 h (salt stress). The sampling time (08:00 and 18:00) indicates treatment periods of 12 and 22 hours, respectively.

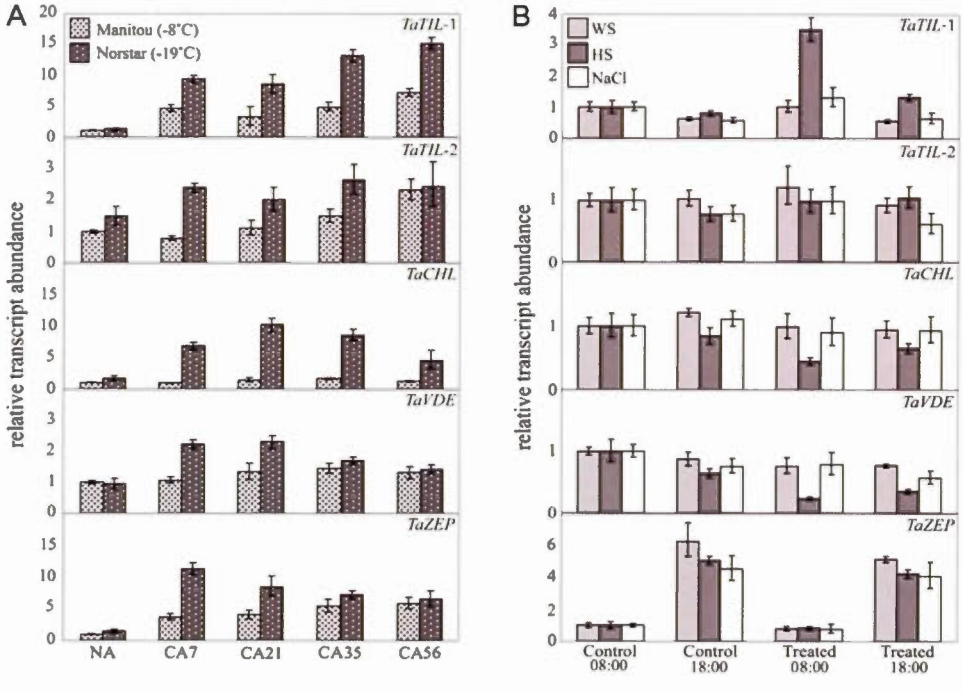

A
Manitou (-8°C)
Norstar (-19°C)
TaTIL-1
TaTIL-2
TaCHL
TaVDE
TaZEP
relative transcript abundance
NA
CA7
CA21
CA35
CA56
B
WS
HS
NaCl
TaTIL-1
TaTIL-2
TaCHL
TaVDE
TaZEP
relative transcript abundance
Control 08:00
Control 18:00
Treated 08:00
Treated 18:00

Fig. 4. Expression analysis of wheat TaTIL and TaCHL lipocalins in various wheat cultivars showing varying levels of freezing tolerance.

Non-acclimated control plants (NA) were maintained at 20°C for 7 days while other plants were cold-acclimated (CA36) at 4°C for 36 days after the 7-day germination period. Tissues were sampled at 18:00. RNA was extracted and analyzed on a denaturing RNA gel blot. Spring wheat genotypes (Triticum aestivum L. cv Glenlea (GLEN; LT50 -8°C) and cv Concorde (CON; LT50 -8°C); winter wheat genotypes (T. aestivum L. cv Monopole (MON; LT50 -15°C), cv Absolvent (ABS; LT50 -16°C), cv Fredrick (FRED; LT50 -16°C), and cv Norstar (NOR; LT50 -19°C)); rRNA, ethidium bromide-stained 28S ribosomal RNA included to show RNA loads (7.5 μg).

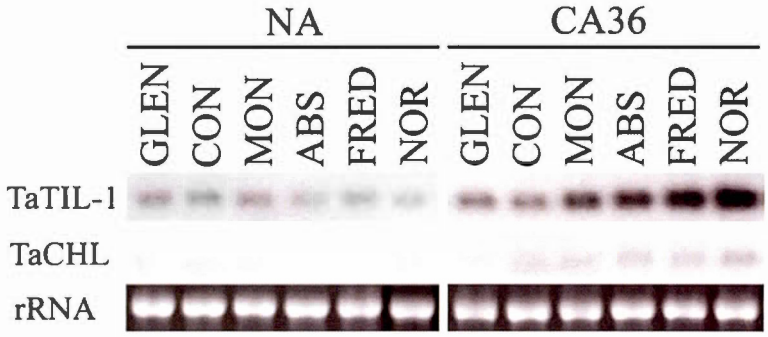

NA
CA36
GLEN
CON
MON
ABS
FRED
NOR
GLEN
CON
MON
ABS
FRED
NOR
TaTIL-1
TaCHL
rRNA

Fig. 5. Expression analysis of wheat lipocalins in different tissues.

Plants were grown for 7 days at 20°C. Non-acclimated control plants (NA) were maintained at 20°C for 1 and 6 days. Cold-acclimated plants (CA) were transfered at 4°C for 1, 6 and 36 days. After CA, some plants were transfered at 20°C for 1 and 5 days for deacclimation (DA). Leaf, crown and root tissues were sampled at 18:00 and RNA was extracted and analyzed as described in Fig. 3.

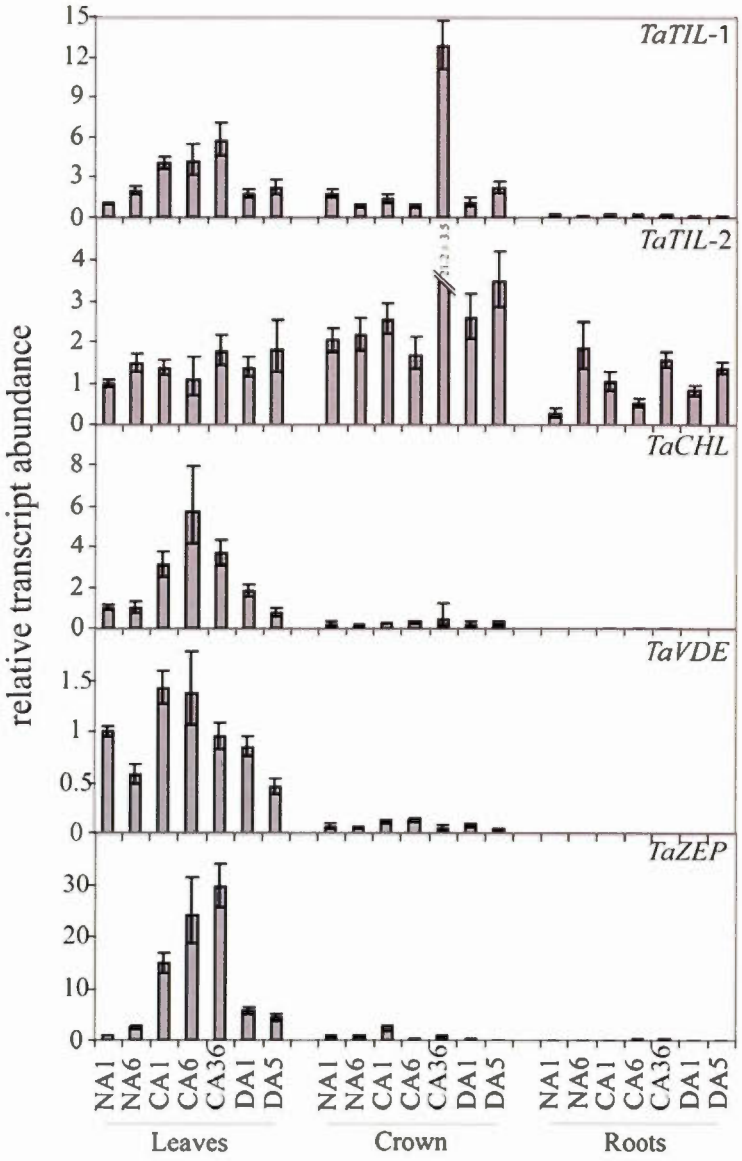

relative transcript abundance
TaTIL-1
TaTIL-2
TaCHL
TaVDE
TaZEP
NA1
NA6
CA1
CA6
CA36
DA1
DA5
Leaves
Crown
Roots

Fig. 6. Expression analysis of wheat lipocalins in response to diurnal cycles.
Plants were germinated for 8 days at 20°C under a 16/8 h day/night photoperiod. Beginning on day 8 at 8:00, plants were grown at 20°C for 12 hours (from 08:00 to 20:00) then kept at 20°C for 52 hours (A, controls) or transferred at 4°C for 52 hours (B, cold-acclimated). Leaf samples were harvested at each time-point, and RNA was extracted and analyzed as described in Fig. 3. Grey areas represent the dark periods.

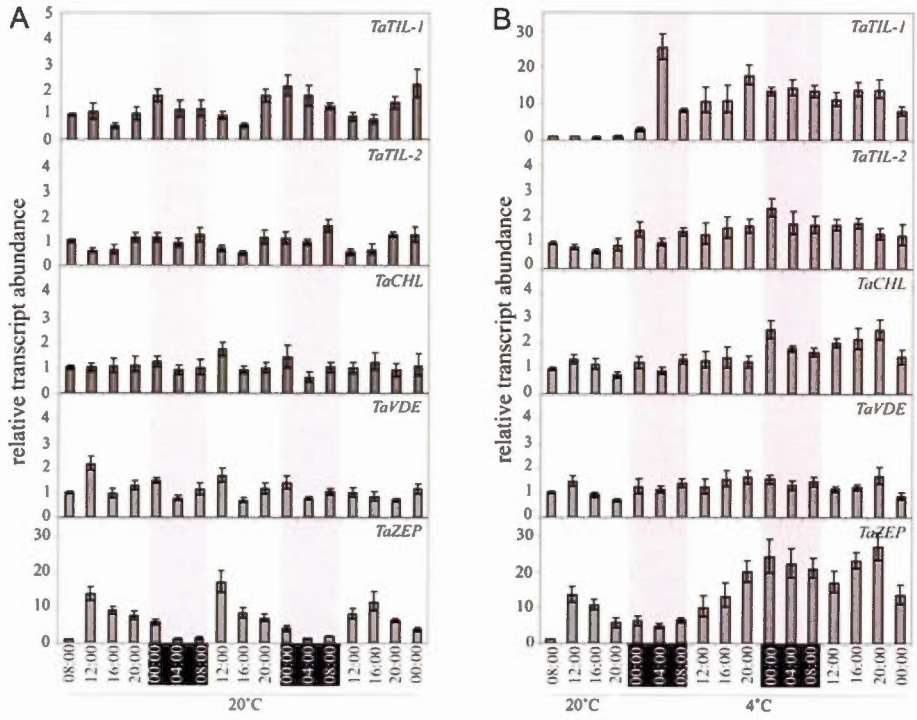
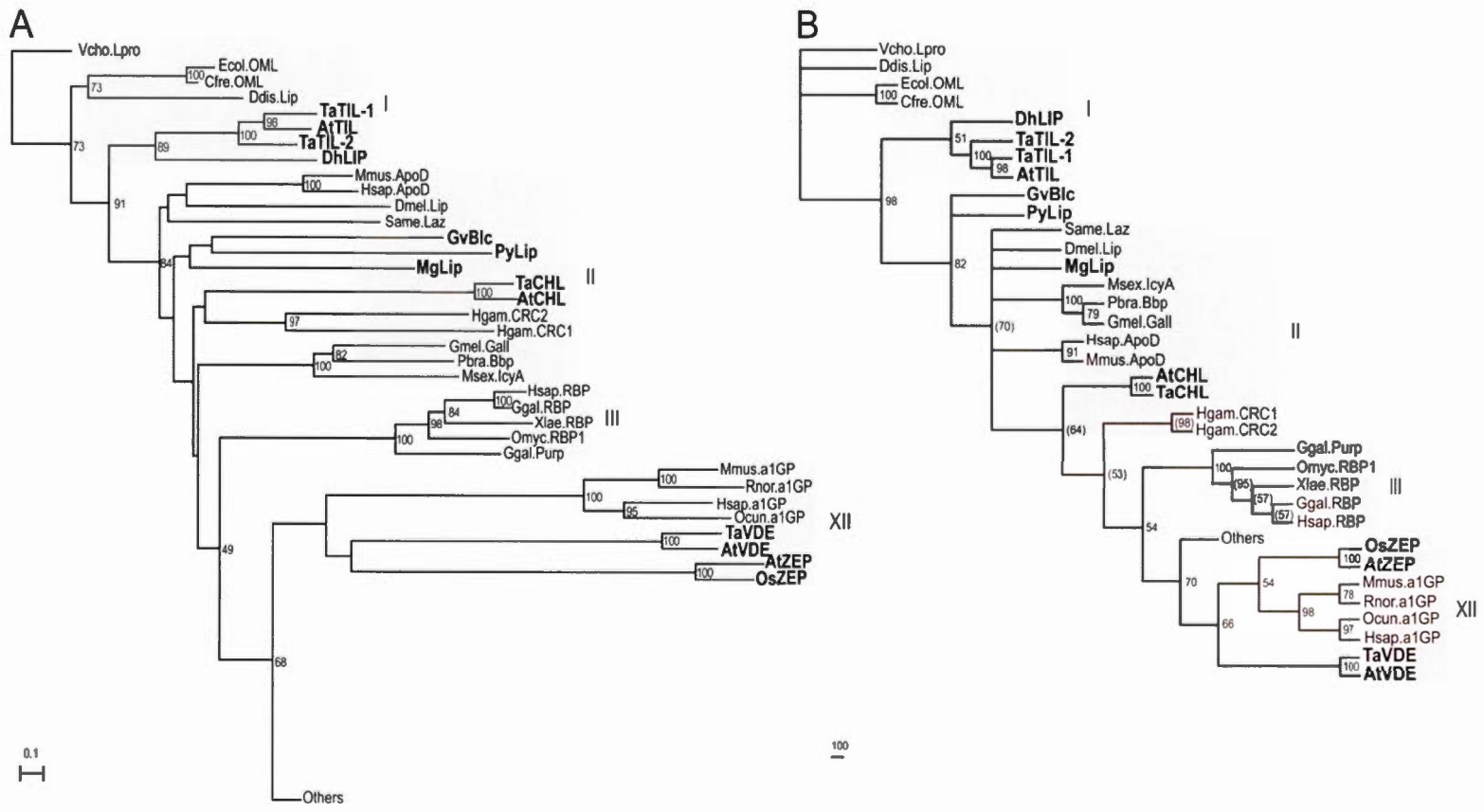

A
B
TaTIL-1
TaTIL-2
TaCHL
TaVDE
TaZEP
relative transcript abundance
20°C
4°C

Fig. 7. Phylogenetic analyses of selected lipocalins.

A. The Neighbor-Joining tree was built from the alignment presented in Supplemental Fig. 7 and rooted with the VchoLpro taxon. Only part of the tree is shown. The global tree is presented in Supplemental Fig. 8. The scale bar represents the branch length (number of amino acid substitutions/site).

B. The maximum likelihood tree was built from the alignment presented in Supplemental Fig. 7 and rooted with the VchoLpro taxon. Bootstrap values in parenthesis indicate nodes that were refined for better resolution. Only part of the tree is shown. The global tree is presented in Supplemental Fig. 9. The scale bar represents the branch length (number of amino acid substitutions/100 residues). The grey boxes, identified by roman numerals, represent the clades nomenclature (Gutiérrez et al., 2000).


A
Vcho.Lpro
Ecol.OML
Cfre.OML
Ddis.Lip
TaTIL-1
AtTIL
TaTIL-2
DhLIP
I
Mmus.ApoD
Hsap.ApoD
Dmel.Lip
Same.Laz
GvBlc
PyLip
MgLip
TaCHL
AtCHL
II
Hgam.CRC2
Hgam.CRC1
Gmel.Gall
Pbra.Bbp
Msex.IcyA
Hsap.RBP
Ggal.RBP
Xlae.RBP
Omyc.RBP1
Ggal.Purp
III
Mmus.a1GP
Rnor.a1GP
Hsap.a1GP
Ocun.a1GP
TaVDE
AtVDE
AtZEP
OsZEP
XII
Others
0.1
B
Vcho.Lpro
Ddis.Lip
Ecol.OML
Cfre.OML
DhLIP
TaTIL-2
TaTIL-1
AtTIL
I
GvBlc
PyLip
Same.Laz
Dmel.Lip
MgLip
Msex.IcyA
Pbra.Bbp
Gmel.Gall
Hsap.ApoD
Mmus.ApoD
II
AtCHL
TaCHL
Hgam.CRC1
Hgam.CRC2
Ggal.Purp
Omyc.RBP1
Xlae.RBP
Ggal.RBP
Hsap.RBP
III
Others
OsZEP
AtZEP
Mmus.a1GP
Rnor.a1GP
Ocun.a1GP
Hsap.a1GP
XII
TaVDE
AtVDE
100

Supplemental Table I. Primer sequences used in cellular localization and Real-Time PCR experiments.

Primer Name	Direction	Primer Sequence	Amplicon Size (bp)
AtTIL-F1	Forward	5'-CAATCTAGAACTATTTGCCGAAGAGAGA-3'	560
AtTIL-R1	Reverse	5'-GCAGATCTATGACAGAGAGAAGAAAGAGATG-3'	
AtTIL-F2	Forward	5'-TACCATGGTGACAGAGAAGAAAGAGATG-3'	560
AtTIL-R2	Reverse	5'-TACCCATGGATTTGCCGAAGAGAGA-3'	
AtTIL-F2	Forward	5'-TACCATGGTGACAGAGAAGAAAGAGATG-3'	460
AtTIL-R3	Reverse	5'-TACCATGGCACTCTGAGGAGTCTTGTG-3'	
AtTIL-F3	Forward	5'-ATAGATCTGACACACCACCTGAGTCC-3'	100
AtTIL-R4	Reverse	5'-CAATCTAGAACTATTTGCCGAAGAGAGA-3'	
TaTIL-1 FAM	Forward	5'-CACGTCAAGGAGGAAGGCTACGACG[FAM]G-3'	131
TaTIL-1	Reverse	5'-CATTTACCAAAGAGCGACTTGAACC-3'	
TaTIL-2	Forward	5'-TACATGGGCCGGTGGTACG-3'	43
TaTIL-2 FAM	Reverse	5'-GATCGGACGGGAAGCACGCGA[FAM]C-3'	
TaCHL	Forward	5'-GCAAAGGACACCAGCTTTATTCAGATATAC-3'	175
TaCHL FAM	Reverse	5'-GACCAGACATCATCTCCGCAAGCTGG[FAM]C-3'	
TaTDE	Forward	5'-GCTCAAGGAATGCAGGATCGAG-3'	63
TaTDE FAM	Reverse	5'-CAACCTGCTGCACATGATGGG[FAM]TG-3'	
TaZEP FAM	Forward	5'-CAGCATGTTGGAATGCCTTTGATGC[FAM]G-3'	59
TaZEP	Reverse	5'-AGCTTTGAGCTGTTGCCACCT-3'	
18S RNA JOE	Forward	5'-GAACATCGGTCTGTGATGCCCTTAGATG[JOE]TC-3'	83
18S RNA	Reverse	5'-GGCCAAGGCTATATACTCGTTGAATAC-3'	

Supplemental Table II. Lipocalins and lipocalin-like proteins used for the alignment (suppl. fig. 7).

Protein	Species	Abbreviation	Accession numbers*	Clade
Outer membrane lipoprotein	*Vibrio cholerae*	Vcho.Lpro	X64097	I
Outer membrane lipoprotein	*Escherichia coli*	Ecol.OML	AE000487	I
Outer membrane lipoprotein	*Citrobacter freundii*	Cfre.OML	U21727	I
Outer membrane lipoprotein	*Gloeobacter violaceus*	*Gv* Blc	BAC88907	I
Putative lipocalin	*Dictyostelium discoideum*	Ddis.Lip	C24642	I
Lipocalin	*Debaryomyces hansenii*	*Dh* Lip	CAG88663	I
Temperature induced lipocalin 2	*Triticum aestivum*	*Ta* TIL-2		I
Temperature induced lipocalin 1	*Triticum aestivum*	*Ta* TIL-1	AAL75812	I
Temperature induced lipocalin	*Arabidopsis thaliana*	*At* TIL	BAB10998	I
Lipocalin	*Porphira yezoensis*	*Py* Lip		II
Outer membrane lipoprotein	*Magnaporthe grisea*	*Mg* Blc	EAA51774	II
Schistocerca americana	*Schistocerca americana*	Same.Laz	U15656	II
Lipocalin	*Drosophila melanogaster*	Dmel.Lip	AC004439	II
Lipocalin	*Arabidopsis thaliana*	*At* CHL	AAK59669	II
Lipocalin	*Triticum aestivum*	*Ta* CHL		II
Crustacyanin 1	*Homarus gammarus*	Hgam.CRC2	P80007	II
Crustacyanin 2	*Homarus gammarus*	Hgam.CRC1	P80029	II
Insecticyanin A	*Manduca sexta*	Msex.IcyA	X64714	II
Galleria mellonella	*Galleria mellonella*	Gmel.Gall	X64715	II
Bilin-binding protein	*Pieris brassicae*	Pbra.Bbp	X76568	II
Apolipoprotein D	*Mus musculus*	Mmus.ApoD	X82648	II
Apolipoprotein D	*Homo sapiens*	Hsap.ApoD	J02611	II
Purpurin	*Gallus gallus*	Ggal.Purp	M17538	III
Retinol-binding protein	*Xenopus laevis*	Xlae.RBP	J02718	III
Retinol-binding protein	*Oncorhynchus mykiss*	Omyc.RBP1	P24774	III
Retinol-binding protein	*Gallus gallus*	Ggal.RBP	X77960	III
Retinol-binding protein	*Homo sapiens*	Hsap.RBP	X00129	III
Beta-lactoglobulin	*Trichosurus vulpecula*	Tvul.BL	U34289	IV
Pregnancy protein 14	*Homo sapiens*	Hsap.PP14	M61886	IV
Beta-lactoglobulin B	*Bos taurus*	Btau.BLB	X14712	IV
Beta-lactoglobulin B	*Sus scrofa*	Sscr.BLB	X54976	IV
Epididymal secretory protein	*Mus musculus*	Mmus.Lcn11	AAQ81972	M
Endometrial P19 protein	*Equus caballus*	Ecab.p19p	X98459	M
Epididymal secretory protein	*Homo sapiens*	Hsap.Lcn9	AAQ81975	M
Epididymal secretory protein	*Lacerta vivipara*	Lviv.LSP	X63151	M
Quiescence-specific protein 21	*Gallus gallus*	Ggal.QS-21	M55644	V M
Choroid plexus lipocalin 1	*Xenopus laevis*	Xlae.cpl1	X84414	V
Choroid plexus lipocalin	*Bufo marinus*	Bmar.lip	Q01584	V
Prostaglandin D synthase	*Homo sapiens*	Hsap.PGDS	M61900	V
Prostaglandin D synthase	*Mus musculus*	Mmus.PGDS	X89222	V
Epididymal secretory protein	*Mus musculus*	Mmus.Lcn12	AAQ63836	V
Epididymal secretory protein	*Homo sapiens*	Hsap.Lcn12	BC041168	V
Neutrophil gelatinase lipocalin	*Homo sapiens*	Hsap.NGAL	X83006	V
Neutrophil gelatinase lipocalin	*Mus musculus*	Mmus.NGAL	P11672	V
Alpha-1 microglobulin	*Xenopus laevis*	Xlae.A1mg	D87752	VI
Alpha-1 microglobulin	*Mus musculus*	Mmus.A1mg	D28812	VI
Alpha-1 microglobulin	*Homo sapiens*	Hsap.A1mg	X04494	VI

Alpha-1 microglobulin	*Salmo salar*	Ssal.A1mg	L26598	VI
Alpha-1 microglobulin	*Pleuronectes platessa*	Ppla.A1mg	X63762	VI
Complement C8γ subunit	*Homo sapiens*	Hsap.C8GC	X06465	VII
Complement C8γ subunit	*Oryctolagus cuniculus*	Ocun.C8GC	L26979	VII
Major urinary protein	*Mus musculus*	Mmus.mMUP	X03525	VIII
Major urinary protein 5	*Mus musculus*	Mmus.MUP5	M16360	VIII
Major urinary protein 4	*Mus musculus*	Mmus.MUP4	M16358	VIII
Major urinary protein	*Mus musculus*	Mmus.MUP	M28649	VIII
Submaxillary alpha-2u-globulin	*Rattus norvegicus*	Rnor.a2g3	J00738	IX
Alpha-2u-globulin (L type)	*Rattus norvegicus*	Rnor.a2g1	M26836	IX
Odorant-binding protein	*Bos taurus*	Btau.OBP	P07435	X
Allergen BDA20	*Bos taurus*	Btau.alle	L42867	X
Odorant-binding protein 1	*Rattus norvegicus*	Rnor.OBP1	J03093	X
Pheromone carrier Aphrodisin	*Cricetus cricetus*	Ceri.Aphr	X65238	X
Prosthate protein Probasin	*Mus musculus*	Mmus.Pbas	AF005204	X
Prosthate protein Probasin	*Rattus norvegicus*	Rnor.Pbas	M27156	X
Lactation protein Trichosurin	*Trichosurus vulpecula*	Tvul.Lip	U40376	XI
Allergen f2	*Canis familiaris*	Cfam.f2p	AF027178	XI
Zeaxanthin epoxidase	*Oryza sativa*	*Os*ZEP	BAB39765	XII
Zeaxanthin epoxidase	*Arabidopsis thaliana*	*At*ZEP	BAB08942	XII
Alpha-1 acid glycoprotein	*Mus musculus*	Mmus.a1GP	M27009	XII
Alpha-1 acid glycoprotein	*Rattus norvegicus*	Rnor.a1GP	J00696	XII
Alpha-1 acid glycoprotein	*Oryctolagus cuniculus*	Ocun.a1GP	X58727	XII
Alpha-1 acid glycoprotein	*Homo sapiens*	Hsap.a1GP	M13692	XII
Violaxanthin de-epoxidase	*Triticum aestivum*	*Ta*VDE	AAK38177	XII
Violaxanthin de-epoxidase	*Arabidopsis thaliana*	*At*VDE	AAL34211	XII
Late lactation protein	*Trichosurus vulpecula*	Tvul.LLP	U34287	XIII
Vomeronasal secretory protein 1	*Mus musculus*	Mmus.VNSP1	D38580	XIII
Epididymal secretory protein	*Mus musculus*	Mmus.Lcn13	AAR11375	XIII
Vomeronasal secretory protein 2	*Mus musculus*	Mmus.VNSP2	D38581	XIII
von Ebner's gland protein	*Sus scrofa*	Sscr.VEG	S77587	XIII
Allergen f1	*Canis familiaris*	Cfam.f1p	AF027177	XIII
von Ebner's gland protein	*Homo sapiens*	Hsap.VEG	S77587	XIII
von Ebner's gland protein 2	*Rattus norvegicus*	Rnor.VEG2	X74806	XIII
Epididymal secretory protein	*Homo sapiens*	Hsap.Lcn5	AAQ81974	XIV/M
Epididymal retinoic acid binding protein	*Mus musculus*	Mmus.ERBP	AAD09351	XIV
Epididymal secretory protein	*Mus musculus*	Mmus.Lcn8	AF082221	XIV

* Accession numbers are not yet available for proteins reconstructed in this study.

Supplemental Table III. FASTA files of lipocalins and lipocalin-like sequences used in the alignment presented in suppl. fig. 7.

>Mmus.a1GP
QNPEHVNITIGDPITNETLSWLSDKWFFIGAAVLNPDYRQEIQKTQMVFFNLTPNLINDTMELREYHT
IDDHCVYNSTHLGIQRENGTLSKYVGGVKIFADLIVLKMHGAFMLAFDLKDEKKRGLSLNAKRPDITP
ELREVFQKAVTHVGMDESEIIFVDWKKDRCSQQEKQQLELEKETKKDPEEGQA

>Rnor.a1GP
IQNPEPANTLGIPITNETLKWLSDKWFYMGAAFRDPVFKQAVQTIQTEYFYLTPNLINDTIELREFQT
TDDQCVYNFTHLGVQRENGTLSKCAGAVKIFAHLIVLKKHGTFMLAFNLTDENRGLSFYAKKPDLSPE
LRKIFQQAVKDVGMDESEIVFVDWTKDKCSEQQKQQLELEKETKKETKKDP

>Hsap.a1GP
QIPLCANLVPVPITNATLDQITGKWFYIASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIFLREYQTR
QDQCIYNTTYLNVQRENGTISRYVGGQEHFAHLLILRDTKTYMLAFDVNDEKNWGLSVYADKPETTKE
QLGEFYEALDCLRIPKSDVVYTDWKKDKCEPLEKQHEKERKQEEGES

>Ocun.a1GP
QDPACANFSTSPITNATLDQLSHKWFFTASAFRNPKYKQLVQHTQAAFFYFTAIKEEDTLLLREYITT
NNTCFYNSSIVRVQRENGTLSKHDGIRNSVADLLLLRDPGSFLLVFFAGKEQDKGMSFYTDKPKASPE
QLEEFYEALTCLGMNKTEVVYTDWTKDLCEPLEKQHEEERKKEKAES

>Hsap.RBP
ERDCRVSSFRVKENFDKARFSGTWYAMAKKDPEGLFLQDNIVAEFSVDETGQMSATAKGRVRLLNNWD
VCADMVGTFTDTEDPAKFKMKYWGVASFLQKGNDDHWIVDTDYDTYAVQYSCRLLNLDGTCADSYSFV
FSRDPNGLPPEAQKIVRQRQEELCLARQYRLIVHNGYCDGRSERNLL

>Ggal.RBP
ERDCRVSSFKVKENFDKNRYSGTWYAMAKKDPEGLFLQDNVVAQFTVDENGQMSATAKGRVRLFNNWD
VCADMIGSFTDTEDPAKFKMKYWGVASFLQKGNDDHWVVDTDYDTYALHYSCRELNEDGTCADSYSFV
FSRDPKGLPPEAQKIVRQRQIDLCLDRKYRVIVHNGFCS

>Xlae.RBP
EKNCRVDNFEVMKDFNKERYAGVWYAVAKKDPEGLFLLDNIAANFKIEDNGKTTATAKGRVRILDKLE
LCANMVGTFIETNDPAKYRMKYHGALAILERGLDDHWVVDTDYTTYAITYACRRRNLDGTCRDSYSFV
FSRDINGLPSESQRIVRRRQEQLCLDRKYRVVVHNGYCETN

>Omyc.RBP1
SDCQVSNIQVMQNFDRSRYTGRWYAVAKKDPVGLFLLDNVVAQFSVDESGKVTATAHGRVIILNNWEM
CANMFGTFEDTPDPAKFKMRYWGAASYLQTGNDDHWVIDTDYDNYAIHYSCREVDLDGTCLDGYSFIF
SRHPTGLRPEDQKIVTDKKKEICFLGKYRRVGHTGFCESS

>Ggal.Purp
QTCAVDSFSVKDNFDPKRYAGKWYALAKKDPEGLFLQDNISAEYTVEEDGTMTASSKGRVKLFGFWVI
CADMAAQYTVPDPTTPAKMYMTYQGLASYLSSGGDNYWVIDTDYDNYAITYACRSLKEDGSCDDGYSL
IFSRNPRGLPPAIQRIVRQKQEEICMSGQFQPVLQSGAC

>*Tatil-1*
MAAKKSGSEMGVVLGLDVARYMGRWYEIASFPNFFQPRDGRDTRATYELMEDGATVHVLNETWSKGKR
DFIEGTAYKADPASEEAKLKVKFYVPPFLPIIPVVGDYWVLYVDDDYQYALVGEPRRKSLWILCRKTH
IEEEVYNQLLEKAKEEGYDVAKLHKTPQSDPPPESDAAPTDSKGTWWFKSLFGK

>*At*til
MTEKKEMEVVKGLNVERYMGRWYEIASFPSRFQPKNGVDTRATYTLNPDGTIHVLNETWSNGKRGFIE
GSAYKADPKSDEAKLKVKFYVPPFLPIIPVTGDYWVLYIDPDYQHALIGQPSRSYLWILSRTAQMEEE
TYKQLVEKAVEEGYDISKLHKTPQSDTPPESNTAPEDSKGVWWFKSLFGK

>*Ta*til-2
MAAMKVVRNLDLERYMGRWYEIACFPSRFQPKDGANTRATYTLGPDGAVKVLNETWTDGRRGHIEGTA
FRADPAGDEAKLKVRFYVPPFLPVFPVTGDYWVLHVDDAYQFALVGQPSRNYLWILCRQPQMDSEGVY
EELVERAKEEGYDVSKLRKTPHPEPTPESQDAPKDGGLWWIKSLFGK

>*At*ZEP
CITGDRINGLVDGISGTWYVKFDTFTPAASRGLPVTRVISRMTLQQILARAVGEDVIRNESNVVDFED
SGDKVTVVLENGQRYEGDLLVGADGIWSKVRNNLFGRSEATYSGYTCYTGIADFIPADIESVGYRVFL
GHKQYFVSSDVGGGKMQWYAFHEEPAGGADAPNGMKKRLFEIFDGWCDNVLDLLHATEEEAILRRDIY
DRPGFTWGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSFQLALELDEAWKQSVETTTPVDVVSSLKRYE
ESRRLRVAIIHAMARMAAIMASTYKAYLGVGLGPLSFLTKFRVPHPGRVGGRFFVDIAMPSMLDWVLG
GNSEKLQGRPPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCVSETLCLTKDEDQPCIV
GSEPDQDFPGMRIVIPSSQVSKMHARVIYKDGAFFLMDLRSEHGTYVTDNEGRRYRATPNFPARFRSS
DIIEFGSDKKAAFRVKVIRKTPKSTRKNESNNDKLLQTA

>*Os*ZEP
CVTGDRINGLVDGISGSWYIKFDTFTPAAERGLPVTRVISRMTLQQILARAVGDDAILNDSHVVDFID
DGNKVTAILEDGRKFEGDLLVGADGIWSKVRKVLFGQSEATYSEYTCYTGIADFVPPDIDTVGYRVFL
GHKQYFVSSDVGAGKMQWYAFHKEPAGGTDPENGKNKRLLEIFNGWCDNVVDLINATDEEAILRRDIY
DRPPTFNWGKGRVTLLGDSVHAMQPNLGQGGCMAIEDGYQLAVELEKSWQESAKSGTPMDIVSSLRRY
EKERILRVSVIHGLARMAAIMATTYRPYLGVGLGPLSFLTKLRIPHPGRVGGRFFIKYGMPLMLSWVL
GGNSTKLEGRPLSCRLSDKANDQLRRWFEDDDALEQAMGGEWYLLPTSSGDSQPIRLIRDEKKSLSIG
SRSDPSNSTASLALPLPQISENHATITCKNKAFYVTDNGSEHGTWITDNEGRRYRRTSELPCPFPSLG
CH

>*Ta*VDE
PDPSALVKNFNMADFRGKWYISSGLNPTFDTFDCQLHEFRLEGDRLVANLAWRIPTPDTGFFTRGAVQ
RFVQDSSQPAILYNHDNEYLHYQDDWYILSSKIENKDDDYIFVYYRGRNDAWDGYGGAVVYTRSKELP
ETIVPELERATKSVGRDFSTFIRTDNTCGAEPPLADRIERTVEKGEKLIVDEVKEIEGEIEGEVKELE
REEETLVKRLADGIMEVKQDVMNFFQGLSKEEMEILDQLNLEATEVEELFSRSLPIRKLR

>*At*VDE
PDPSVLVQNFNISDFNGKWYITSGLNPTFDAFDCQLHEFHTEGDNKLVGNISWRIKTLDSGFFTRSAV
QKFVQDPNQPGVLYNHDNEYLHYQDDWYILSSKIENKPEDYIFVYYRGRNDAWDGYGGAVVYTRSSVL
PNSIIPELEKAAKSIGRDFSTFIRTDNTCGPEPALVERIEKTVEEGERIIVKEVEEIEEEVEKEVEKV
GRTEMTLFQRLAEGFNELKQDEENFVRELSKEEMEFLDEIKMEASEVEKLFGKALPIRKVR

>*Ta*CHL
MMTKGMTAKNFDPVRYSGRWFEVASRKGGFAGQGQEDCHCTQGVYTFDEKAGAIKVETFCVHGSPDGY
ITGIRGKVQCLSQEDMAGAETDLEREEMISSKCFLRFPTLPFIPKLPYDVLATDYDNYAVVSGAKDTS
FIQIYSRTPNPGPEFIEKYKSYAAGFGYDLSKIKDTPQDCEVSSDQLAEMMSMPGMDQALTNQFPDLK
LKSSVAFDPFTSVTQTLKKLAEVYFK

>*At*CHL
MMMMRGMTAKNFDPVRYSGRWFEVASLKRGFAGQGQEDCHCTQGVYTFDMKESAIRVDTFCVHGSPDG
YITGIRGKVQCVGAEDLEKSETDLEKQEMIKEKCFLRFPTIPFIPKLPYDVIATDYDNYALVSGAKDK

GFVQVYSRTPNPGPEFIAKYKNYLAQFGYDPEKIKDTPQDCEVTDAELAAMMSMPGMEQTLTNQFPDL
GLRKSVQFDPFTSVFETLKKLVPLYFK

>Mmus.ApoD
QNFHLGKCPSPPVQENFDVKKYLGRWYEIEKIPASFEKGNCIQANYSLMENGNIEVLNKELSPDGTMN
QVKGEAKQSNVSEPAKLEVQFFPLMPPAPYWILATDYENYALVYSCTTFFWLFHVDFFWILGRNPYLP
PETITYLKDILTSNGIDIEKMTTTDQANCPDFL

>Hsap.ApoD
QAFHLGKCPNPPVQENFDVNKYLGRWYEIEKIPTTFENGRCIQANYSLMENGKIKVLNQELRADGTVN
QIEGEATPVNLTEPAKLEVKFSWFMPSAPYWILATDYENYALVYSCTCIIQLFHVDFAWILARNPNLP
PETVDSLKNILTSNNIDVKKMTVTDQVNCPKLS

>Sscr.VEG
AQEFPAVGQPLQDLLGRWYLKAMTSDPEIPGKKPESVTPLILKALEGGDLEAQITFLIDGQCQDVTLV
LKKTNQPFTFTAYDGKRVVYILPSKVKDHYILYCEGELDGQEVRMAKLVGRDPENNPEALEEFKEVAR
AKGLNPDIVRPQQSETCSPGGN

>Cfam.flp
QDTPALGKDTVAVSGKWYLKAMTADQEVPEKPDSVTPMILKAQKGGNLEAKITMLTNGQCQNITVVLH
KTSEPGKYTAYEGQRVVFIQPSPVRDHYILYCEGELHGRQIRMAKLLGRDPEQSQEALEDFREFSRAK
GLNQEILELAQSETCSPGGQ

>Rnor.VEG2
AQAFPTTEENQDVSGTWYLKAAAWDKEIFTPDKKFGSVSVTPMKIKTLEGGNLQVKFTVLISGRCQEM
STVLEKTDEPGKYTAYSGKQVFTVYSIPSAVEDHYIFYYEGKIHRHHFQIAKLVGRNPEINQEALEDF
QNAVRAGGLNPDNIFTFIPKQSETCPLGSN

>Hsap.VEG
AHHLLASDEEIQDVSGTWYLKAMTVDREFPEMNLESVTPMTLTTLEGGNLEAKVTMLISGRCQEVKAV
LEKTDEPGKYTADGGKHVAYIIRSHVKDHYIFYCEGELHGKPVRGVKLVGRDPKNNLEALEDFEKAAG
ARGLSTESILIPRQSETCSPGSD

>Mmus.VNSP1
QDSSFLAFNNGNFSGKWFLKALVSEDDIPINKVSPMLILVLNNGDIELSITHMIYDQCLEVTTILEKT
DVPGQYLAFEGKTHLQVQLSSVKGHYMLYCDGEIEGMRFLMTQLIGRDPQENLEALEEFKVFTQIKGL
VAENLVILEQMEKCEPESFYELPSRPSE

>Mmus.VNSP2
LQTYDDLPFISEEDKLSGVWFIKATVSQRREVEGETLVAFPIKFTCPEEGTLELRHTLASKGECINVG
IRLQRTEEPGQYSAFWGHTLFYIYDLPVKDHYIIYCESHPFQKISQFGYLIGKYPEENQDTLEVFKEF
IQHKGFLQEKIGVPEQRDRCIPIHDSAHQDHKC

>Mmus.MUP4
HAEEATSKGQNLNVEKINGEWFSILLASDKREKIEEHGSMRVFVEHIHVLENSLAFKFHTVIDGECSE
IFLVADKTEKAGEYSVMYDGFNTFTILKTDYDNYIMFHLINEKDGKTFQLMELYGRKADLNSDIKEKF
VKLCEEHGIIKENIIDLTKTNRCLKARE

>Mmus.MUP5
EEASSERQNFNVEKINGKWFSILLASDKREKIEEHGTMRVFVEHIDVLENSLAFKFHTVIDEECTEIY
LVADKTEKAGEYSVTYDGFNTFTILKTDYDNYIMFHLINKKDEENFQLMELFGREPDLSSDIKEKFAK
LCEEHGIVRENIIDLSNANRCLQARE

>Mmus.MUP
HAEEASSTGRNFNVEKINGEWHTIILAFDKREKIEDNGNFRLFLEQIHVLENSLVLKFHTVRDEECSE
LSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLINENDGETFQLMGLYGREPDLSSDIKERF
AQLCEKHGILRENIIDLSNANRCLQARE

>Mmus.mMUP
HAEESSSMERNFNVEQISGYWFSIAEASYEREKIEEHGSMRAFVENITVLENSLVFKFHLIVNEECTE
MTAIGEQTEKAGIYYMNYDGFNTFSILKTDYDNYIMIHLINKKDGKTFQLMELYGREPDLSLDIKEKF
AKLCEEHGIIRENIIDLTNVNRCLEARE

>Rnor.a2g1
HAEEASSTRGNLDVDKLNGDWFSIVVASDKREKIEENASMRVFMQHIDVLENSLGFKFRIKENGECRE
LYLVAYKTPEDGEYFVEYDGGNTFTILKTDYDRYVMFHLINFKNGETFQAMVLYGRTKDLSSDIKEKF
AKLCEAHGITRDNIIDLTKTDHCLQARG

>Rnor.a2g3
HAEEASFERGNLDVDKLNGDWFSIVVASDKREKIEENGSMRVFVQHIDVLENSLGFTFRIKENGVCTE
FSLVADKTAKDGEYFVEYDGENTFTILKTDYDNYVMFHLVNVNNGETFQLMELYGRTKDLSSDIKEKF
AKLCVAHGITRDNIIDLTKTDRCLQARG

>Hsap.Lcn9
QEFDPHTVMQRNYNVARVSGVWYSIFMASDDLNRIKENGDLRVFVRNIEHLKNGSLIFDFEYMVQGEC
VAVVVVCEKTEKNGEYSINYEGQNTVAVSETDYRLFITFHLQNFRNGTETHTLALYETCEKYGLGSQN
IIDLTNKDPCYSKHYRSPPRPPMRW

>Cfam.f2p
QEGNHEEPQGGLEELSGRWHSVALASNKSDLIKPWGHFRVFIHSMSAKDGNLHGDILIPQDGQCEKVS
LTAFKTATSNKFDLEYWGHNDLYLAEVDPKSYLILYMINQYNDDTSLVAHLMVRDLSRQQDFLPAFES
VCEDIGLHKDQIVVLSDDDRCQGSRD

>Tvul.Lip
LQPECSRSEEDLSDEKERKWEQLSRHWHTVVLASSDRSLIEEEGPFRNFIQNITVESGNLNGFFLTRK
NGQCIPLYLTAFKTEEARQFKLNYYGTNDVYYGSSKPNEYAKFIFYNYHDGKVNVVANLFGRTPNLSN
EIKKRFEEDFMNRGFRRENILDISEVDHC

>Ccri.Aphr
QDFAELQGKWYTIVIAADNLEKIEEGGPLRFYFRHIDCYKNCSEMEITFYVITNNQCSKTTVIGYLKG
NGTYQTQFEGNNIFQPLYITSDKIFFTNKNMDRAGQETNMIVVAGKGNALTPEENEILVQFAHEKKIP
VENILNILATDTCPE

>Rnor.OBP1
HHENLDISPSEVNGDWRTLYIVADNVEKVAEGGSLRAYFQHMECGDECQELKIIFNVKLDSECQTHTV
VGQKHEDGRYTTDYSGRNYFHVLKKTDDIIFFHNVNVDESGRRQCDLVAGKREDLNKAQKQELRKLAE
EYNIPNENTQHLVPTDTCNQ

>Mmus.Pbas
VMSLKKKIDGPWQTIYLAASTMEKINEGSPLRTYFRHICVGRRSNQVYLYFFIKKGTKCQLYKVIGRK
KQEVYYAQYEGSIAFMLKMVNEKILLFHYFNKNRRNDVTRVAGVLAKGKLNKEEMTEFMNLVEEMGIE
EENVQRIMDTDNCPSKIRISITD

>Rnor.Pbas

MMTDKNLKKKIEGNWRTVYLAASSVEKINEGSPLRTYFRRIECGKRCNRINLYFYIKKGAKCQQFKIV
GRRSQDVYYAKYEGSTAFMLKTVNEKILLFDYFNRNRRNDVTRVAGVLAKGRQLTKDEMTEYMNFVEE
MGIEDENVQRVMDTDTCPNKIRIR

>Btau.alle
AQETPAEIDPSKIPGEWRIIYAAADNKDKIVEGGPLRNYYRRIECINDCESLSITFYLKDQGTCLLLT
EVAKRQEGYVYVLEFYGTNTLEVIHVSENMLVTYVENYDGERITKMTEGLAKGTSFTPEELEKYQQLN
SERGVPNENIENLIKTDNCPP

>Btau.BLB
LIVTQTMKGLDIQKVAGTWYSLAMAASDISLLDAQSAPLRVYVEELKPTPEGDLEILLQKWENGECAQ
KKIIAEKTKIPAVFKIDALNENKVLVLDTDYKKYLLFCMENSAEPEQSLACQCLVRTPEVDDEALEKF
DKALKALPMHIRLSFNPTQLEEQCHI

>Sscr.BLB
VEVTPIMTELDTQKVAGTWHTVAMAVSDVSLLDAKSSPLKAYVEGLKPTPEGDLEILLQKRENDKCAQ
EVLLAKKTDIPAVFKINALDENQLFLLDTDYDSHLLLCMENSASPEHSLVCQSLARTLEVDDQIREKF
EDALKTLSVPMRILPAQLEEQCRV

>Hsap.PP14
VPAMDIPQTKQDLELPKLAGTWHSMAMATNNISLMATLKAPLRVHITSLLPTPEDNLEIVLHRWENNS
CVEKKVLGEKTGNPKKFKINYTVANEATLLDTDYDNFLFLCLQDTTTPIQSMMCQYLARVLVEDDEIM
QGFIRAFRPLPRHLWYLLDLKQMEEPCRF

>Mmus.A1mg
DPASTLPDIQVQENFSESRIYGKWYNLAVGSTCPWLSRIKDKMSVQTLVLQEGATETEISMTSTRWRR
GVCEEITGAYQKTDIDGKFLYHKSKWNITLESYVVHTNYDEYAIFLTKKSSHHHGLTITAKLYGREPQ
LRDSLLQEFKDVALNVGISENSIIFMPDRGECVPGDREVEPTSIAR

>Hsap.A1mg
PVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIMDRMTVSTLVLGEGATEAEISMTSTRWRK
GVCEETSGAYEKTDTDGKFLYHKSKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTITAKLYGRAPQ
LRETLLQDFRVVAQGVGIPEDSIFTMADRGECVPGEQEPEPILIPRV

>Xlae.A1mg
SPIQPEDNIQIQENFDLQRIYGKWYDIAIGSTCKWLKHHKEKFNMGTLELSDGETDGEVRIVNTRMRH
GTCSQIVGSYQKTETPGKFDYFNARWGTTIQNYIVFTNYNEYVIMQMRKKKGSETTTTVKLYGRSPDL
RPTLVDEFRQFALAQGIPEDSIVMLPNNGECSPGEIE

>Ssal.A1mg
VPVLPEPLFPIQDNFDLTKFMGKWHDIAIGSTCPWMQRHKGDAAIGTLELQASGTEDKVSMTRSMKKH
GKCEQISGDYELTATPGRLTYHIAKWGADVDAYVVDTNYDEYAIVMLSKQKTGGEKTKSAKLYSRTME
LPPTILEDFRRLVREQGMADDTIIIKQNKGECVPGTEPVAAEPQPEITAP

>Ppla.A1mg
LPVLPEPLYPTQENFDLTRFVGTWHDVALTSSCPHMQRNRADAAIGKLVLEKDTGNKLKVTRTRLRHG
TCVEMSGEYELTSTPGRIFYHIDRWDADVDAYVVHTNYDEYAIIIMSKQKTSGENSTSLKLYSRTMSV
RDTVLDDFKTLVRHQGMSDDTIIIKQNKGDCIPGEQVEEAPSQPEPK

>Hsap.C8GC
QKPQRPRRPASPISTIQPKANFDAQQFAGTWLLVAVGSACRFLQEQGHRAEATTLHVAPQGTAMAVST
FRKLDGICWQVRQLYGDTGVLGRFLLQARGARGAVHVVVAETDYQSFAVLYLERAGQLSVKLYARSLP

VSDSVLSGFEQRVQEAHLTEDQIFYFPKYGFCEAADQFHVLDEVRR

>Ocun.C8GC
RWAQKPRGAPSAISAIQPKANFDAQQFAGTWLLAAVGSACHFLQEQGHRAEATALHVAPQGAAMAVST
FRKLDGICWQVSQRYGATGVPGRFLLPARGPRGAVHVVAAETDYHSFAVLYLERARQLSVKLYVRSLP
VSDSVLGAFEQRVAQANLTQDQVLFFPTYGFCEAADQFHILDEVRR

>Xlae.cpl1
SLWVGAEVQVQPDFQKEKVLGKWYGIGLASNSNWFKDRKSHMKMCTTIITPTADGNLEVTATYPKMDR
CETKSMTYFKTEQLGGFRAKSPRYGSEHDMRVVETNYDEYILMYTVKTKGSETNQIVSLFGRDKDLRP
ELLDKFQNFAKSQGLADDNIIILPHTDQCMTEA

>Bmar.lip
DVPIQPDFQEDKILGKWYGIGLASNSNWFQSKKQQLKMCTTVITPTADGNLDVVATFPKLDRCEKKSM
TYIKTEQPGRFLSKSPRYGSDHVIRVVESNYDEYTLMHTIKTKGNEVNTIVSLFGRRKTLSPELLDKF
QQFAKEQGLTDDNILILPQTDSCMSEV

>Mmus.PGDS
QTPAQGHDTVQPNFQQDKFLGRWYSAGLASNSSWFREKKAVLYMCKTVVAPSTEGGLNLTSTFLRKNQ
CETKIMVLQPAGAPGHYTYSSPHSGSIHSVSVVEANYDEYALLFSRGTKGPGQDFRMATLYSRTQTLK
DELKEKFTTFSKAQGLTEEDIVFLPQPDKCIQE

>Hsap.PGDS
QAAPEAQVSVQPNFQQDKFLGRWFSAGLASNSSWLREKKAALSMCKSVVAPATDGGLNLTSTFLRKNQ
CETRTMLLQPAGSLGSYSYRSPHWGSTYSVSVVETDYDQYALLYSQGSKGPGEDFRMATLYSRTQTPR
AELKEKFTAFCKAQGFTEDTIVFLPQTDKCMTEQ

>Mmus.NGAL
QDSTQNLIPAPSLLTVPLQPDFRSDQFRGRWYVVGLAGNAVQKKTEGSFTMYSTIYELQENNSYNVTS
ILVRDQDQGCRYWIRTFVPSSRAGQFTLGNMHRYPQVQSYNVQVATTDYNQFAMVFFRKTSENKQYFK
ITLYGRTKELSPELKERFTRFAKSLGLKDDNIIFSVPTDQCIDN

>Hsap.NGAL
QDSTSDLIPAPPLSKVPLQQNFQDNQFQGKWYVVGLAGNAILREDKDPQKMYATIYELKEDKSYNVTS
VLFRKKKCDYWIRTFVPGCQPGEFTLGNIKSYPGLTSYLVRVVSTNYNQHAMVFFKKVSQNREYFKIT
LYGRTKELTSELKENFIRFSKSLGLPENHIVFPVPIDQCIDG

>Hsap.Lcn12
KVLQAQTPTPLPLPPPMQSFQGNQFQGEWFVLGLAGNSFRPEHRALLNAFTATFELSDDGRFEVWNAM
TRGQHCDTWSYVLIPAAQPGQFTVDHGVEPGADREETRVVDSDYTQFALMLSRRHTSRLAVLRISLLG
RSWLLPPGTLDQFICLGRAQGLSDDNIVFPDVTGNMVHLQACWAVGTGPAGMSLVDPRGAGPSVYPGS
SAPACAQGSPGSWVPVLNPGSEPPPAAPGPLSWATSSHPGSPVPGHLLPPQVPCPGPPPPAPPAPGPL
SRPTSSHPGSPVLGYLLPPQVPCPGPSPPSGSPVLGHLLPSPIPAHKELGLIPGGALDLSSLPWVAAP
A

>Mmus.Lcn12
QILESQISAMSQGFPQMTSFQSDQFQGEWFVLGLADNTFRREHRALLNFFTTLFELKEKSQFQVTNSM
TRGKHCNTWSYTLIPATKPGQFTRDNRGSGPGADRENIQVIETDYITFALVLSLRQTSSQNITRVSLL
GRNWRLSHKTIDKFICLTRTQNLTKDNFLFPDLSDWLPDPQVC

>*Dh*Lip
KKEEMPVVEKIELDKYLGKWYEIARKPFLFQKKCYSNVSAKYSLNDNANINVDNSCYSKDGKLRQAIG

EAFTQNPPFNSKLKVSFLPKAIRFLPIGRGDYWILKIDDNYQTVLVGGPSRKYMWILSRSQNHDEIVV
QDYLDYAKEIGFDVSDIIMTKQTNE

>Cfre.OML
CSSPTPPKGVTVVNNFDAKRYLGTWYEIARFDHRFERGLDKVTATYSLRDDGGINVINKGYNPDREMW
QKTEGKAYFTGDPSTAALKVSFFGPFYGGYNVIALDREYRHALVCGPDRDYLWILSRTPTISDEMKQQ
MLAIATREGFEVNKLIWVKQPGA

>Ecol.OML
CSSPTPPRGVTVVNNFDAKRYLGTWYEIARFDHRFERGLEKVTATYSLRDDGGLNVINKGYNPDRGMW
QQSEGKAYFTGAPTRAALKVSFFGPFYGGYNVIALDREYRHALVCGPDRDYLWILSRTPTISDEVKQE
MLAVATREGFDVSKFIWVQQPGS

>Vcho.Lpro
MEILIGATCLGMPESVKPVSDFELNNYLGKWYEVARLDHSFERGLSQVTAEYRVRNDGGISVLNRGYS
EEKGEWKEAEGKAYFVNGSTDGYLKVSFFGPFYGSYVVFELDRENYSYAFVSGPNTEYLWLLSRTPTV
ERGILDKFIEMSKERGFDTNRLIYVQLQ

>Ddis.Lip
ILGGVTYAYNSFKRYIPEGVHAVKPFYPEKYVGKWYEIARLYTYFEKDLDKITAEYSINKDGSITVVN
SGYNYKKKKRENAKGIAYFVNGSDEGMLKVSFFGPFYSGYNVIAIDPDYKYALIAGQSFDYMWILSKE
PTIPEKIKNSYLELAKSVGYDITKLIWSKQNENEN

>Dmel.Lip
AVWVAHAQVPFPGKCPDVKLLDTFDAEAYMGVWYEYAAYPFAFEIGKKCIYANYSLIDNSTVSVVNAA
INRFTGQPSNVTGQAKVLGPGQLAVAFYPTQPLTKANYLVLGTDYESYAVVYSCTSVTPLANFKIVWI
LTRQREPSAEAVDAARKILEDNDVSQAFLIDTVQKNCPRLDGNGTGLAGEDGLDVDDFVSTTVPNAIE
KA

>Gmel.Gall
VHEGKCPDFKPVDNFNLTAYQGVWYEISKTPNDAEKNGKCGQAEYKLEGEVVKVKNSHVVDGVQKYVE
GTAKFAEDANKSAKLLVTLTYGAVNRESPLNVIATDYQNYAIAYTCKYDEKSKSHNDSIWILSRAKKL
EGDAKTAVDNYLKEHAKEIDASKLVQTDFSEEACKFTSTSAVTEPQTKKQ

>Pbra.Bbp
NVYHDGACPEVKPVDNFDWSNYHGKWWEVAKYPNSVEKYGKCGWAEYTPEGKSVKVSNYHVIHGKEYF
IEGTAYPVGDSKIGKIYHKLTYGGVTKENVFNVLSTDNKNYIIGYYCKYDEDKKGHQDFVWVLSRSKV
LTGEAKTAVENYLIGSPVVDSQKLVYSDFSEAACKVNN

>Msex.IcyA
GDIFYPGYCPEVKPVDDFDLSAFAGAWHEIAKLPLENENEGKCTVAEYKYDGKKASVYNSFVINGVKE
YMEGDLEIAPDAKLTKQGKYVMTFKFGPRVVVQVPWVLATDYKNYAINYNCNYHPDKKAHSIHAWVLS
RNKVLEGNTKEVVDNVLKTFSHLIDASKFMSNEFSEAACQYSTTYSLTGPDRH

>Hgam.CRC2
DGIPSFVTAGKCASVANQDNFDLRRYAGRWYQTHIIENAYQPVTRCIHSNYEYSTNDYGFKVTTAGFN
PNDEYLKIDFKVYPTKEFPAAHMLIDAPSVFAAPYEVIETDYETYSCVYSCITTDNYKSEFAFVFSRT
PQTSGPAVEKTAAVFNKNGVEFSKFVPVSHTAECVYRA

>Hgam.CRC1
DKIPDFVVPGKCASVDRNKLWAEQTPNRNSYAGVWYQFALTNNPYQLIEKCVRNEYSFDGKQFVIKST
GIAYDGNLLKRNGKLYPNPFGEPHLSIDYENSFAAPLVILETDYSNYACLYSCIDYNFGYHSDFSFIF

SRSANLADQYVKKCEAAFKNINVDTTRFVKTVQGSSCPYDTQKTL

>Mmus.Lcn13
AQEAPPDDLVDYSGIWYAKAMVHNGTLPSHKIPSIVFPVRIIALEEGDLETTVVFWNNGHCREFKFVM
KKTEEPGKYTAFHNTKVIHVEKTSVNEHYIFYCEGRHNGTSSFGMGKLMGRDSGENPEAMEEFKNFIK
RMNLRLENMFVPEIGDKCVESD

>Tvul.BL
IQAIENIHSKEELVVEKLIGPWYRVEEAKAMEFSIPLFDMNIKEVNRTPEGNLELIVLEQTDSCVEKK
FLLKKTEKPAEFEIYIPSESASYTLSVMETDYDNYILGCLENVNYREKMACAHYERRIEENKGMEEFK
KIVRTLTIPYTMIEAQTREMCRV

>Mmus.Lcn11
LQDFHPEQVTGPWHTLKLASTDRSLIEEGGAYRCFMTDIVLLDNGNLNVTYFHRKDGKCVKEFYIAEK
TDTPGQYTFEYQGRNSLTFVHVTEDFAIMDLENQSEGGTTIVIEFHGRSLSTDELG

>*Gv*Blc
DSQPIETVAEVDFNRYDGRWYELARTPNIFQIGCTCVTANYSVLSESSISVFNTCNRFRPRGNLVTID
GVAVVADPNAPGKLLITFEGSPVAEDYWIIDLVEDPNNSAGDYAFAAIGGPNRDFIFIISRKPALETY
QDVLAYQGIVKRLQAQHFPVDALNSTPQPTSCTYKSQSLPGGL

>Btau.OBP
KNAQEEEAEQNLSELSGPWRTVYIGSTNPEKIQENGPFRTYFRELVFDDEKGTVDFYFSVKRDGKWKN
VHVKATKQDDGTYVADYEGQNVFKIVSLSRTHLVAHNINVDKHGQTTELTGLFVKLNVEDEDLEKFWK
LTEDKGIDKKNVVNFLENEDHPHPE

>*Mg*Lip
DTSSVPNTVPSLWDGECFYPTPDIGFDTKSYLGRWYQVAGTVAPFTASCKCIYAQYALNDNGTIQVNN
TCEAGGRAVNILGTAEPADPGYGAKGALRVQFPGQPGPACSGPNYVVQDYTGDFALVQTYNFSTLFVL
SRNQHPEEAVLDAWIKRAGALGSDLSDVIKNDQTNCSFT

>Same.Laz
AQETMGCADRSAINDFNATLYMGKWYEYAKMGSMPYEEGGVCVTAEYSMSSNNITVVNSMKDNTTHEV
NTTTGWAEFASELHTDGKLSVHFPNSPSVGNYWILSTDYDNYSIVWSCVKRPDSAASTEISWILLRSR
NSSNMTLERVEDELKNLQLDLNKYTKTEQSAKYCA

>Tvul.LLP
DDVAFSAFTPSEGTYYVQVIAVDKEFPEEEIPRDMSPLTIMYLDDGKMEARFTMKKDDNCEEINIMLE
KTADPRKITMNRRLRYTCAAVRTSKQKHWILVCPREFQGETIRMAKLVGPNTDKNPKALEDFYRFIYR
ERFDKRRIITPKQTEACAPEHA

>Ecab.p19p
RRPHALHMGPGDPNFDEKLVKGKWFSVALASNEPKFIAKDTDMKFFIHKIQVTPESLQFHFHRKVRGM
CVPTMMTAHKTKKKFQYTVNHSGHKTIFLEKVDPKHFVIFCAHSMKHGKETVVVTLFSRTPTVSPDVM
WMFKKYCKTHGIHTSNIVDLTQTDRCLHARH

>Ggal.QS-21
AATVPDSSEVAGKWYIVALASNTDSFLREKGKMKMVMARISFLGEDELEVSYAAPSPKGCRKWETTFK
KTSDDGELYYSEEAEKTVEVLDTDYKSYAVIFATRVKDGRTLHMMRLYSRSREVSPTAMAIFRKLARE
RNYTDEMVAVLPSQEECSVDEV

>Hsap.Lcn5

QAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVTLTPENNLRTLSSQHGLGGCDQSVM
DLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQLEFGDEPFNTVELYSLTETASQEAMGLFTKW
SRSLGFLSQ

>Mmus.Lcn8
ESTRVELVPEKIAGFWKEVAVASDQKLVLKAQRRVEGLFLTFSGGNVTVKAVYNSSGSCVTESSLGSE
RDTVGEFAFPGNREIHVLDTDYERYTILKLTLLWQGRNFHVLKYFTRSLENEDEPGFWLFREMTADQG
LYMLARHGRCAELLKEGLV

>Mmus.ERBP
TEAAVVKDFDVNKFLGFWYEIALASKMGAYGLAHKEEKMGAMVVELKENLLALTTTYYNEGHCVLEKV
AATQVDGSAKYKVTRISGEKEVVVVATDYMTYTVIDITSLVAGAVHRAMKLYSRSLDNNGEALNNFQK
IALKHGFSETDIHILKHDLTCVNALQSGQI

>Lviv.ESP
DIPVVPNFDAQKTVGKWHPIGMASKLPEVPEYEQKISPMDHMVELTDGDMKLTANYMDGVCKEATAML
KHTDKPGVFKFTGGEIRMMDIDYEKYLIMYMKKSTFEAMYLSARGSDVGDDIKEKFKKLVLEQNFPEA
HIKYFNAEQCTPTAA

>*Py*Lip
RKCPNPATVPALDVAAYTGRWYQIGVTAEFAERQEDNKPCVTADYRLTGPTVEVINCKQDVPANRSSG
AIVGCAQAVAFPGKKEDPGKLGVQFPGAPFPAPYWVINLAGSKEDGYRVAVVYSCTSTGSFFSQGLFL
LSRTPKLRYGVFEAVYWYVRVLARGIRFQKGNEFKLTPQGKSCTYRGDEGAKVVFQ

Supplemental Table IV. FASTA files of plant lipocalins and lipocalin-like sequences.

Temperature-Induced Lipocalins

>*Ta*TIL-1
MAAKKSGSEMGVVLGLDVARYMGRWYEIASFPNFFQPRDGRDTRATYELMEDGATVHVLNETWSKGKR
DFIEGTAYKADPASEEAKLKVKFYVPPFLPIIPVVGDYWVLYVDDDYQYALVGEPRRKSLWILCRKTH
IEEEVYNQLLEKAKEEGYDVAKLHKTPQSDPPPESDAAPTDSKGTWWFKSLFGK

>*Ta*TIL-2
MAAMKVVRNLDLERYMGRWYEIACFPSRFQPKDGANTRATYTLGPDGAVKVLNETWTDGRRGHIEGTA
FRADPAGDEAKLKVRFYVPPFLPVFPVTGDYWVLHVDDAYQFALVGQPSRNYLWILCRQPQMDEGVYE
ELVERAKEEGYDVSKLRKTPHPEPTPESQDAPKDGGLWWIKSLFGK

>*Hv*TIL-1
MAVKKIGSEMGVVLGLDVARYMGRWYEIASFPNFFQPRDGRDTRATYELMEDGATVHVLNETWSKGKR
DYIEGTAYKADPASDEAKLKVKFYVPPFLPIIPVVGDYWVLYVDDDYQYALVGEPRRKSLWILCRKTH
IEEEVYNQLLEKAKEEGYDVAKLHKTPQSDPPPEGDAAPTDSKGAWWFKSLFGK

>*Hv*TIL-2
MAAMKVVRNLDLERYMGRWYEIACFPSRFQPKDGANTRATYTLGPDGAVKVLNETWTDGRRGHIEGTA
FRADDAGDEAKLKVRFYVPPFLPVFPVTGDYWVLHVDDAYQYALVGQPSRNYLWILCRQPRMDEGVYN
ELVERAKEEGYDVSKLRRTPHPEPTPESQDAPKDGGLWWIRSLFGK

>*Os*TIL-1
MAAAAVEKKSGSEMTVVRGLDVARYMGRWYEIASLPNFFQPRDGRDTRATYALRPDGATVDVLNETWT
SSGKRDYIKGTAYKADPASDEAKLKVKFYLPPFLPVIPVVGYWVLYVDDDYQYALVGEPRRKDLWILC
RQTSMDDEVYGRLLEKAKEEGYDVEKLRKTPQDDPPPESDAAPTDTKGTWWFKSLFGK

>*Os*TIL-2
MKVVRNLDLERYMGRWYEIACFPSRFQPRDGTNTRATYTLAGDGAVKVLNETWTDGRRGHIEGTAYRA
DPVSDEAKLKVKFYVPPFLPIFPVVGDYWVLHVDDAYSYALVGQPSLNYLWILCRQPHMDEEVYGQLV
ERAKEEGYDVSKLKKTAHPDPPPETEQSAGDRGVWWIKSLFGR

>*At*TIL
MTEKKEMEVVKGLNVERYMGRWYEIASFPSRFQPKNGVDTRATYTLNPDGTIHVLNETWSNGKRGFIE
GSAYKADPKSDEAKLKVKFYVPPFLPIIPVTGDYWVLYIDPDYQHALIGQPSRSYLWILSRTAQMEEE
TYKQLVEKAVEEGYDISKLHKTPQSDTPPESNTAPEDSKGVWWFKSLFGK

>*Le*TIL
MATKVMEVVKNLDLKRYMGRWYEIASFPSRFQPKDGVDTRATYTLNSDGTVHVLNETWCNGKRGFIEG
TAYKADPNSDEAKLKVRFYVPPFLPIIPVTGDYWVLYIDEDYQYALIGQPSRRYLWILSRQTRLDDEI
YNQLVEKAKEEGYDVSKLHKTPQSDSPPDSEDSPKDTKGIWWIKSILGK

>*Le*TIL'
MTTKEMEVVKNLDVEKYMGRWYEIASFPSRNQPKDGVNTRATYTLNQDGTVHVLNETWSGGKRGSIEG
TAYKADPKSDEAKLKVKFYIPPFLPIIPIVGDYWVLYIDDDYQYALIGQPSKKYLWILCRQPHLDEEI
YNQLVEKAKEVGYDVSKLHKTPQADPPPDGEDAPKDTKGFWWIKSILGK

>*So*TIL

MAAAEGKKSGGQMTVVRGLDVARYMGRWYEIASFPSFFQPRDGRDTRATYRLLEDGATVHVLNETWSK
GKRDYIEGTAYKADASSDEAKLKVKFYLPPFLPIIPVVGDYWVLYVDDDYQYALVGEPRRKNLWILCR
KTSIDEEVYNQLVERAKEEGYDVSKLHRTPQDDPPPESDAAPTDTKGVWWFKSLFGK

>*Zm*TIL-1
MAAEEGEKAKSGGGGQQMTVVRGLDVARYMGRWYEIASFPSFFQPRDGRDTRATYRLLEDGATVHVLN
ETWSKGKRDYIEGTAYKADPGSDEAKLKVKFYLPPFLPIVPVVGDYWVLYVDDDYQYALVGEPRRKNL
WILCRKTSIDEDVYNQLVERAKEEGYDVSKLHRTPQDDPPPESDAAPTDTKGVWWFKSLFGK

>*Zm*TIL-2
MAMQVVRNLDLERYAGRWYEIACFPSRFQPKTGTNTRATYTLNPDGTVKVVNETWADGRRGHIEGTAW
RADPASDEAKLKVRFYVPPFLPLIPVTGDYWVLHIDADYQYALVGQPSRNYLWILCRQPHMDESVYKE
LVERAKEEGYDVSKLRKTAHPDPPPESEQSPRDGGMWWVKSIFGK

>*Sb*TIL-1
MAAEAGKTTAATKSGGGGQIMTVVRGLDVARYMGRWYEIASFPSFFQPRDGRDTRATYRLLEDGATVH
VLNETWSKGKRDYIEGTAYKADPNSDEAKLKVKFYLPPFLPVIPVVGDYWVLYVDDDYQYALVGEPRR
KNLWILCRKTSIDEEVYNQLVERAKEEGYDVSKLHRTPQDDPPPESDAAPTDTKGVWWFKSLFGK

>*Sb*TIL-2
MAAAAMRVVRDLDLERYAGRWYEIACFPSTFQPKTGTNTRATYTLNPDDRTVKVLNETWTDGGGRRGH
IEGTAWRADPASDEAKLKVRLYVPPFLPVFPVTGDYWVLHVDADYQYALVGQPSRKYLWILCRQPQMD
ESVYNELVERAKEEGYDVSKLRKTAHPDPPPESEQSPGDRGVWWIKSIFGK

>*Gm*TIL
MANKEMEVVKGLDLQRYMGRWYEIASFPSRNQPKDGENTRATYTLRNDGTVQVLNETWSNGKRGYIQG
TAYKVDPKSDEAKFKVKFYIPPFLPIIPINGDYWVLFTDDEYQYALIGQPSRNYLWILSRKPHLDDEI
YNELVQRAKNVGYDVSKLRKTPQSDPPPEEEGPDDTKGIWWLKSIFGK

>*Gm*TIL'
MANNEMQVERGLDLERYMGRWYEIASFPSRNQPKDGVNTRATYTLRNDGTVQVLNETWSNGKRGHIEG
TAFKSNRTSDEAKFKVKFYVPPFLPIIPVTGDYWVLFIDGDYQYALIGQPSRNCLWILSRKPHLDDEI
YNKLVQRAKDVGYDVSKLHKTPQSDPPPEEEGPQDTKGIWWLKSILGK

>*Pp*TIL
MGGEKDLNVVQNVDLKRYQGRWYEIASIPSRFQPSTGTNSRATYALKEDQTIHVLNETWVSGKRSYIE
GKAWKADAASPDAKLKVRFLVPPFFPIFPVTGDYWVMKLDENYQWALIGQPSRRYLWVLSRTPELSDE
IYNQLLEHATNEGYDVSKLHKTQQIPEIGEEGTSNSENTDRAGVWWLKSLFGK

>*Pt*TIL
MGKEDLQVVKGLDLQRYMGVWYEIASMPSFFQPKNGINTRATYSLNKDSTVHVLNETFVDGKKSSIEG
SAYKVDPKSEDAKFKVKFMVPPFFPIIPVYGNYWVLLLDEDYQWALIGEPSLKYLWVLCRQRQLDEAI
YNRLLEHARQEGYDVGRLHKTTQNDDPETEAPKDKGFWWIKALLGK

>*Tr*TIL
MGGEKDLNVVQNVDLTRYQGRWYEIASNPTRFQPSRGSNSRATYTLQEDQTVEVLNETWVNNKRSYIT
GKAWKADPASPDAKLKVRFMVPPFLPVIPVTGDYWVMKLDADYQWALVGVPDRTSLWVLSRTQEMSEE
TYKELVEHAANEGYDVSKLHKTEQNPEVGEGEEESTDRAGAWWVKSLFGK

>*Vv*TIL

MAKKEMEVVRGIDLQRYMGRWYEIASFPSFFQPKNGINTRATYTLEADGTTVRVLNETWSDGKRSYIE
GTAYKADPKSDQAKLKVKFYVPPFLPIIPVVGDYWVLFLDEEYQYALIGQPSRKYLWILCRQTHMDEE
IYEMLVEKAKEVGYDVSKLRKTTQTDPPPEGEGPQDTKGIWWIKSIFGK

>*Mt*TIL
MANKEMDVARGVDLKRYMGRWYEIACFPSRFQPSDGKNTRATYTLRDDGTVNVLNETWSGGKRSYIEG
TAYKADPNSDEAKLKVKFYVPPMLPIIPVTGDYWVLHLDHDYHYALIGQPSRNYLWILCRQPHLDEEI
YNELVQKAKEEGYDVSKLRKTPQSDTPPEQEGPEDTKGIWWFKSLFGK

>*Mt*TIL′
MGNTVGKDKEVVKGVDLERYMGRWYEIASFPSFFQPKNGENTRATYTLNSDGTVHVLNETWNNGKRTS
IEGSAYKADPKSDEAKLKVKFYVPPFLPIIPAVGDYWILYLDEDYQYALIGGPTNKFLWILSRQPHLD
ETIYNQLVEKAKEEGYDVSKLHKTPQSDPPPE

>*St*TIL
MTTKEMEVVKNLDVEKYMGRWYEIASFPSRNQPKDGVNTRATYTLNQDGTVHVLNETWSGGKRGSIEG
TAYKVDPKSDEAKLKVKFYIPPFLPIIPIVGDYWVLYIDDDYQYALIGQPSKKYLWILCRQPHLDEEI
YNQLVEKAKEVGYDVSKLHKTPQADPPPDGEDAPKDDTKGIWWIKSILGK

>*St*TIL′
MATKVMEVVKNLDLKRYMGRWYEIASFPSRFQPKDGVDTRATYTLNSDGTVHVLNETWCNGKRGFIEG
TAYKADPNSDEAKLKVRFYVPPFLPIIPVTGDYWVLYIDEDYQYALIGQPSRRYLWILSRRTCLDDEI
YNQLVEKAKEEGYDVSKLHKTPQSDSPPESEDSPEDTKGIWWIKSILGK

>*Bn*TIL
MTTEKKEMEVVKGLDLERYMGRWYEIASFPSRFQPKNGADTRATYTLNPDGTVKVLNETWDGGKRGFI
QGSAFKTDPKSDEAKFKVRFYVPPFLPIIPVTGDYWVLYIDPEYQHAVIGQPSRSYLWILSRTAHVEE
ETYKQLVEKAVEQGYDVSKLRKTAQSDTPPESDAAPDDTKGIWWIKSIFGK

>*Prp*TIL
MAKKTMDVVKGLDLQRYMGRWYEIASFPSRFQPKNGENTRATYTLRDDGTVNVLNETWSDGKRSSIEG
TAYKADPSSEEAKLKVKFYVPPFLPIIPVVGDYWVLFIDEDYQYALIGQPSRNYLWILSRQPRLDDEI
YNQLVQRAKDEEYDVSKLHKTPQSETPPEGEEGPKDTKGIWWFKSLLGK

>*Pa*TIL
MAKKTMDVVKGLDLQRYMGRWYEIASFPSRFQPKNGENTRATYTLRDDGTVNVLNETWSDGKRSSIEG
TAYKADPSSEEAKLKVKFYVPPFLPIIPVVGDYWVLFIDEDYQYALIGQPSRNYLWILSRQPRLDDEI
YNQLVQRAKDEEYDVSKLHKTPQSETPPEGEEGPKDTKGIWWFKSLLGK

>*Mc*TIL
MAQKAKEMVVVKGLELGRYMGRWYEIASFPSRFQPRDGENTRATYTLRDDGIVDVLNETWSLGKRSYI
QGTAYKADPNSDEAKLKVKFYVPPFLPIIPVTGDYWVLFIDDDYQYALIGQPSRNYLWILCRTPHMDE
SVYNELVQKAVEEGYDVNKLHKTPQADPPPEGNQAPEDTKGVWWFKSLIGK

>*Mc*TIL′
MAHKSKEMVVVRGLDLERFMGRWYEIASFPSFFQPRDGENTRATYTLNDDGTVHVLNETWSHGKRDAI
EGTAYKADPKSDEAKLKVKFYVPPFLPIIPVTGNYWVLFIDDDYQYALIGEPLRKYLWILCRKTNMDE
SIYEELVQKAVEEGYDVKKLHKTPQADPPPESSDQTPKDKGGWWIKSLFGK

>*Ga*TIL

MSQKTMEVVKNLDIKRYMGRWYEIASFPSRFQPRNGVNTRATYTLNEDGTVHVLNETFTDGKRGFIEG
TAYKADPQSDEAKLKVKFYVPPFLPIIPIVGDYWVLHLDDDYQYALIGQPSRNYLWVLCRQTHMDDEI
YNQLVQKAKDEGYDVSKLHKTPQSDPPPEGDDTPKDAKGIWWIKSLLGM

>*Ga*TIL'
MEVVKNLDIQRYMGKWYEIASFPSFFQPKKGENTSAFYTLKEDGTVHVLNETFVNGKKDSIEGTAYKA
DPKSDEAKLKVKFYVPPFLPIIPVTGDYWVLYIDEDYQYVLVGGPTKKYLWILCRQKHMDEEIYNMLE
QKAKDLGYDVSKLHKTPQSDSTPEGEHVPQEKGFWWIKSLFGK

>*Cs*TIL
MASKKEMEVVRGLDIKRYMGRWYEIASFPSRNQPKNGADTRATYTLNEDGTVHVRNETWSDGKRGSIE
GTAYKADPKSDEAKLKVKFYVPPFFPIIPVVGNYWVLYIDDNYQYALIGEPTRKYLWILCREPHMDEA
IYNQLVEKATSEGYDVSKLHRTPQSDNPPEAEESPQDTKGIWWIKSIFGK

>*Pb*TIL
MATKKEMEVVKGVDLKRYMGRWYEIASFPSRFQPKNGVNTRATYTLNEDGTVHVLNETWNDGKRGYIE
GSAYKADPNSDEAKLKVKFYVPPFLPIIPVVGDYWVLSLDEDYQYALIGQASRKYLWILCRKTHMEDD
IYNQLVEKAKEEGYDVEKLHKTPQTDPPPEEEGPKDTKGIWWFQSILGK

>*Pb*TIL'
MATKKEMEVVKGVDLKRYMGRWYEIASFPSRFQPKNGVNTRATYTLNEDGTVHVLNETWNDGKRGSIE
GSAYKADPNSDEAKLKVKFYVPPFLPIIPVVGDYWILYLDDDYQYALIGQPSRSYLWILCRKTHMEDE
IYNQLVEKAKEEGYDVGKLHKTPQTDPPPEEEGPKDTKGIWWIKSILGK

>*Pot*TIL
MATKKVMEVVKGVDLKRYMGRWYEIASFPSRFQPKNGVNTRATYTLNEDGTIHVLNETWNDGKRGSIE
GSAYKADPNSDEAKLKVKFYVPPFLPIIPVVGDYWILYLDDDYQYALIGQPSRSYLWILCRKTHMEDE
IYNQLVEKAKEEGYDVGKLHKTPQTDPPPEEEGPKDTKGIWWIKSILGK

>*PotxPotr*TIL
MATKKEMEVVKGVDLKRYMGRWYEIASFPSRFQPKNGVNTRATYTLNEDGTVHVLNETWNDGKRGSIE
GSAYKADPNSDEAKLKVKFYVPPFLPIIPVVGDYWILYLDDDYQYALIGQPSRSSLWILCRKTHMEDE
IYNQLVEKAKEEGYDVGKLHKTPQTDPPPEEEGPKDTKGIWWIKSILGK

>*Ls*TIL
MSKKAMEVVKGIDLQRYMGRWYEIASFPSRFQPKDGINTRATYTLKDDGTVNVLNETWSGGKRGFIEG
TAYKADPKSDEAKLKVKFYVPPFLPIIPVTGDYWVLYLDDDYQYALIGQPSRSYLWILCRQTHLDDEI
YNQLVQKATEEGYDVSKLKKTTQTEPPPESEDAPADTKGIWWFKSLFGK

CHloroplastic Lipocalins

>*At*CHL
MILLSSSISLSRPVSSQSFSPPAATSTRRSHSSVTVKCCCSSRRLLKNPELKCSLENLFEIQALRKCF
VSGFAAILLLSQAGQGIALDLSSGYQNICQLGSAAAVGENKLTLPSDGDSESMMMMMMRGMTAKNFDP
VRYSGRWFEVASLKRGFAGQGQEDCHCTQGVYTFDMKESAIRVDTFCVHGSPDGYITGIRGKVQCVGA
EDLEKSETDLEKQEMIKEKCFLRFPTIPFIPKLPYVIATDYDNYALVSGAKDKGFVQVYSRTPNPGPE
FIAKYKNYLAQFGYDPEKIKDTPQDCEVTDAELAAMMSMPGMEQTLTNQFPDLGLRKSVQFDPFTSVF
ETLKKLVPLYFK

>*Sb*CHL

MVLVVLGCSPASPRPACTPNSRRRCSATRQKIIRCSLNEETLLSKHGVVSRQLISCLAASLVFVSPPS
QAIPAETFAHPGLCQIATVAAIDNASVPLKFDNPSDDGSAGMMMRGMTAKNFDPVRYSGRWFEVASLK
RGFAGQGQEDCHCTQGVYSFDEKSRSIQVDTFCVHGGPDGYITGIRGRVQCLSEEDMSSAETDLERQE
MIREKCFLRFPTLPFIPKEPYDVLATDYDNYAVVSGAKDTSFIQIYSRTPNPGPEFIEKYKSYVANFG
YDPSKIKDTPQDCEYMSSDQIALMMSMPGMNEALTNQFPDLKLKAPVALNPFTSVFDTLKKLLELYFK

>*Os*CHL
MVLALLLGSSSSSLAAPHPACSSRRKCRPAGRNNFRCSLHDKVPLNAHGVLSTKLLSCLAASLVFISP
PCQAIPAETFVQPKLCQVAVVAAIDKAAVPLKFDSPSDDGGTGLMMKGMTAKNFDPIRYSGRWFEVAS
LKRGFAGQGQEDCHCTQGVYSFDEKSRSIQVDTFCVHGGPDGYITGIRGRVQCLSEEDMASAETDLER
QEMIKGKCFLRFPTLPFIPKEPYDVLATDYDNYAVVSGAKDTSFIQIYSRTPNPGPEFIEKYKSYAAN
FGYDPSKIKDTPQDCEVMSTDQLGLMMSMPGMTEALTNQFPDLKLSAPVAFNPFTSVFDTLKKLVELY
FK

>*Hv*CHL
MALLPLVGFPSFPFPACPSRRTCGPASRMNFRCCVQERVPAVRNDGISKHLLSCLAASLVFISTPSQA
VPADTFARPSLCQVAVVAAIDKAAVPLKFDGPSDDGMMMMTKGMTAKNFDPVRYSGRWFEVASLKRGF
AGQGQEDCHCTQGVYTFDEKAGAIKVETFCVHGSPDGYITGIRGKVQCLSQEDMASAETDLEKEEMIS
SKCFLRFPTLPFIPKLPYDVLATDYDNYAVVSGAKDTSFIQIYSRTPNPGPEFIEKYKSYAAGFGYDP
SKIKDTPQDCEVSSDQLAQMMSMPGMDEALTNQFPDLKLKSSVAFDPFTSVTQTLKKLEVYFK

>*In*CHL
MASHNLFVAQSSPVLLPSHHPSKPRGVSGKMIVRCTLEQTAPSKARAKHLVSGLAASIVFLTQVNSVV
AADLSYQNNICQLASAADNLPSLPLDGGDDNGGMLMMMRGMTAKNFDPTRYAGRWFEVASLKRGFAGQ
GQEDCHCTQGIYTFDVNAAAIQVDTFCVHGGPDGYITGIRGKVQCLSEDETLKTATDLEKQEMIKGKC
YLRFPTLPFIPKEPYDVIATDYDNFAIVSGAKDKSFVQIYSRTPDPGPEFIEKYKAYLADFGYDPSKI
KDTPQDCEVMSNSQLSAMMSMAGMQQALNNQFPELELKAPVAFNPFTSVFDTLKKLVELYFK

>*St*CHL
MFCYNLVAQQSPPILVQSQYPSKPRGLPVKVVSACCTECPIFRKVEVKHVISGLATSILCLSPSNMAF
AADLPHYNSVFQLANVADSMPTLPLEKENDGGKLMMMRGMTAKDFDPIRYSGRWFEVASLKRGFAGQG
QEDCHCTQGIYTVDMNAPAIQVDTFCVHGGPDGYITGIRGRVQCLNEEDKEKDETDLERQEMIREKCY
LRFPTLPFIPKEPYDVIATDYDNFALVSGAKDKSFVQIYSRTPNPGPEFIEKYKNYLASFGYDPSKIK
DTPQDCEVKTTSQLSAMMSMSGMQQALNNQFPDLELRRPVQFNPFTSVFETLKKLAELYFK

>*Ta*CHL
MALLPLLGSPSFPFPASRPARPSRRKCGPAARMNFRCSAEERAPVRNSGISKHVLSCLAASLLFISPP
SQAVPADTFARPSLCQVAVVAAIDKGAVPLKFDAPSDDAMMMMTKGMTAKNFDPVRYSGRWFEVASRK
GGFAGQGQEDCHCTQGVYTFDEKAGAIKVETFCVHGSPDGYITGIRGKVQCLSQEDMAGAETDLEREE
MISSKCFLRFPTLPFIPKLPYDVLATDYDNYAVVSGAKDTSFIQIYSRTPNPGPEFIEKYKSYAAGFG
YDLSKIKDTPQDCEVSSDQLAEMMSMPGMDQALTNQFPDLKLKSSVAFDPFTSVTQTLKKLAEVYFK

>*Gm*CHL
MVELLLRASPPPPHSSYLRQCRTVSGRTLVKCSLEVPSKVLTKHVLSGLAASLIFISPANQTIAADLS
RAPNNICQLASASENAVTSPFENEKGSNLMMMRGMTAKDFDPVRYSGRWFEVASLKRGFAGQGQEDCH
CTQGVYTFDREAPSIQVDTFCVHGGPNGFITGIRGRVQCLSEEDLGKTETQLEKQEMIKEKCYLRFPT
LPFIPKEPYDVIATDYDNFSLVSGAKDQSFIQIYSRTPNPGPEFIEKYKSYLANYGYDPSKIKDTPQD
CEVMSNSQLAAMMSMSGMQQALTNQFPDLGLNAPIELNPFTSVFDTLKKLLEPYFK

>*So*CHL
GLCQIATVAAIDSASVTLKFDNPSDDGSAGMMMRGMTAKNFDPVRYSGRWFEVASLKRGFAGQGQEDC
HCTQGVYSFDEKSRSIQVDTFCVHGGPDGYITGIRGRVQCLSEEDMASAETDLERQEMIREKCFLRFP

TLPFIPKEPYDVLATDYDNYAVVSGAKDTSFIQIYSRTPNPGPEFIEKYKSYVANFGYDPSKIKDTPQ
DCEYMSSDQIALMMSMPGMNEALTNQFPDLKLKAPVALNPFTSVFDTLKKLLELYFK

Violaxanthin De-Epoxidases

>*At*VDE
MAVATHCFTSPCHDRIRFFSSDDGIGRLGITRKRINGTFLLKILPPIQSADLRTTGGRSSRPLSAFRS
GFSKGIFDIVPLPSKNELKELTAPLLLKLVGVLACAFLIVPSADAVDALKTCACLLKGCRIELAKCIA
NPACAANVACLQTCNNRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVPRKSDLGEFPAPDPSVLV
QNFNISDFNGKWYITSGLNPTFDAFDCQLHEFHTEGDNKLVGNISWRIKTLDSGFFTRSAVQKFVQDP
NQPGVLYNHDNEYLHYQDDWYILSSKIENKPEDYIFVYYRGRNDAWDGYGGAVVYTRSSVLPNSIIPE
LEKAAKSIGRDFSTFIRTDNTCGPEPALVERIEKTVEEGERIIVKEVEEIEEEVEKEVEKVGRTEMTL
FQRLAEGFNELKQDEENFVRELSKEEMEFLDEIKMEASEVEKLFGKALPIRKVR

>*Nt*VDE
MALAPHSNFLANHETIKYYVGSKLPGHKRFSWGWEDYFGSIVVAKICSSRRIPRYFRKSPRICCGLDS
RGLQLFSHGKHNLSPAHSINQNVPKGNSGCKFPKDVALMVWEKWGQFAKTAIVAIFILSVASKADAVD
ALKTCTCLLKECRLELAKCISNPACAANVACLQTCNNRPDETECQIKCGDLFENSVVDEFNECAVSRK
KCVPRKSDVGDFPVPDPSVLVQKFDMKDFSGKWFITRGLNPTFDAFDCQLHEFHTEENKLVGNLSWRI
RTPDGGFFTRSAVQKFVQDPKYPGILYNHDNEYLLYQDDWYILSSKVENSPEDYIFVYYKGRNDAWDG
YGGSVLYTRSAVLPESIIPELQTAAQKVGRDFNTFIKTDNTCGPEPPLVERLEKKVEEGERTIIKEVE
EIEEEVEKVRDKEVTLFSKLFEGFKELQRDEENFLRELSKEEMDVLDGLKMEATEVEKLFGRALPIRK
LR

>*Os*VDE_jap
MAARPEVVSALSPPAGGGAMGGVRYHRCCPPRAYLWRKGDHLPLHHAKISARCSEIKAHTVLQGSDAL
SSIREWSRSHLVTMTGLVACAVLVVPSADAVDALKTCTCLLKECRIELAKCIANPSCAANVACLNTCN
NRPDETECQIKCGDLFENTVVDEFNECAVSRKKCVPQKSDVGEFPVPDPSALVKNFNMADFNGKWYIS
SGLNPTFDTFDCQLHEFRVEGDKLIANLTWRIRTPDSGFFTRTAIQRFVQDPAQPAILYNHDNEFLHY
QDDWYIISSKVENKEDDYIFVYYRGRNDAWDGYGGAVLYTRSKVVPESIVPELERAAKSVGRDFSTFI
RTDNTCGPEPPLVERIEKTVEQGEKTIIREVQEIEGEIEGEVKELEEEEVTLFKRLTDGLMEVKQDLM
NFFQGLSKEEMELLDQMNMEATEVEKVFSRALPIRKLR

>*Os*VDE_ind
MAARPEVVSALSPPAGGGAMGGVRYHRCCPPRAHLWRKGDHLPLHHAKIPARCSEIKVHTVLQASDAL
SSTREWSRSHLVTMTGLVACAVLVVPSADAVDALKTCTCLLKECRIELAKCIANPSCAANVACLNTCN
NRPDETECQIKCGDLFENTVVDEFNECAVSRKKCVPQKSDVGEFPVPDPSALVKNFNMADFNGKWYIS
SGLNPTFDTFDCQLHEFRVEGDKLIANLTWRIRTPDSGFFTRTAIQRFVQDPAQPAILYNHDNEFLHY
QDDWYIISSKVENKEDDYIFVYYRGRNDAWDGYGGAVLYTRSKVVPESIVPELERAAKSVGRDFSTFI
RTDNTCGPEPPLVERIEKTVEQGEKTIIREVQEIEGEIEGEVKELEEEEVTLFKRLTDGLMEVKQDLM
NFFQGLSKEEMELLDQMNMEATEVEKVFSRALPIRKLR

>*Cs*VDE
MALSAHLIYLSNDGSIGLCARRQLTCERLRRRVADPCCVVNVKMQPNRRIPKYFGLLRSYRMPCGLES
KYSNLLSCGSMKISSVCENSTSIPEEKGIFEFQMEVIMSVLKSQLIRVAAVMACIFLVIPAADAVDAL
KTCSRLLKECGVELAKCIANPSCAANVACLQACNNRPDETECQIKCGDLFENSVVDEFNECAVSRKKC
VPQKSDVGEFPVPHPNVLVRNFNMKDFSGKWFITSGLNPTFDAFDCQLHEFHMESNKLLGNSTWRIRT
PDGGFFTRSAVQRFVQDPTQPAILYNHDNEYLHYQDDWYILSSKIENKPDDYVFVYYRGRNDAWDGYG
GAVVYTRSAVLPNSIIPELEKAAQSVGRDFSKFIRTDNSCGPEPPLVERLEKTVEEGERTIIREVEEI
EGEVEKTEMNLFGRLLEGFKELQQDEENFLRELSKEEMDILSELKMEASEVEKLFGQALPLRKLR

>*So*VDE
MALVARSICVSYDEIAGICNNVSHRNFKKWVQWKNPFLFQDDARRNIRFNDRKLSCTKFIGASEKLQH
SKSPKSGLISCGWEVNSSKVVSNAVIPKKWNLLKLKVVEVTAIVACTFFVMSSAQAVDALKTCTCLLK
ECRIELAKCIANPSCAANVACLQTCNNRPDETECQIKCGDLFANKVVDEFNECAVSRKKCVPQKSDVG
EFPVPDPSVLVKSFNMADFNGKWFISSGLNPTFDAFDCQLHEFHLEDGKLVGNLSWRIKTPDGGFFTR
TAVQKFAQDPSQPGMLYNHDNAYLHYQDDWYILSSKIENQPDDYVFVYYRGRNDAWDGYGGAFLYTRS
ATVPENIVPELNRAAQSVGKDFNKFIRTDNTCGPEPPLVERLEKTVEEGERTIIKEVEQLEGEIEGDL
EKVGKTEMTLFQRLLEGFQELQKDEEYFLKELNKEERELLEDLKMEAGEVEKLFGRALPIRKLR

>*Ta*VDE
MLPRQCFKHVFPAEGSSSSILHGPGSRGAGSRGRTTLNFHRCCVRASLWRTDHLHISTARSSEIKVHT
LLQVPDVFNTIKSWSKLQLVTVTGLAACVVLLVPSAGATDALKTCTCLLKECRIELAKCIANPSCAAN
VACLNTCNNRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVPKKSDVGEFPVPDPSALVKNFNMAD
FRGKWYISSGLNPTFDTFDCQLHEFRLEGDRLVANLAWRIPTPDTGFFTRGAVQRFVQDSSQPAILYN
HDNEYLHYQDDWYILSSKIENKDDDYIFVYYRGRNDAWDGYGGAVVYTRSKELPETIVPELERATKSV
GRDFSTFIRTDNTCGAEPPLADRIERTVEKGEKLIVDEVKEIEGEIEGEVKELEREEETLVKRLADGI
MEVKQDVMNFFQGLSKEEMEILDQLNLEATEVEELFSRSLPIRKLR

>*Ls*VDE
MALSLHTVFLCKEEALNLYARSPCNERFHRSGQPPTNIIMMKIRSNNGYFNSFRLFTSYKTSSFSDSS
HCKDKSQICSIDTSFEEIQRFDLKRGMTLILEKQWRQFIQLAIVLVCTFVIVPRVDAVDALKTCACLL
KECRIELAKCIANPSCAANVACLQTCNNRPDETECQIKCGDLFENSVVDQFNECAVSRKKCVPRKSDV
GEFPVPDRNAVVQNFNMKDFSGKWYITSGLNPTFDAFDCQLHEFHMENDKLVGNLTWRIKTLDGGFFT
RSAVQTFVQDPDLPGALYNHDNEFLHYQDDWYILSSQIENKPDDYIFVYYRGRNDAWDGYGGSVIYTR
SPTLPESIIPNLQKAAKSVGRDFNNFITTDNSCGPEPPLVERLEKTAEEGEKLLIKEAVEIEEEVEKE
VEKVRDTEMTLFQRLLEGFKELQQDEENFVRELSKEEKEILNELQMEATEVEKLFGRALPIRKLR

>*Hs*VDE
MLSRQCFKHVFPAEGSSSSILHGPGSRGAASRSRTALNFHRCCVRASLWRADHLNIRTAGSSEVKVYT
SLQVPDVFNSIKSWSKLQLVTVTGLAACVVLLVPSAGATDALKTCTCLLKECRIELAKCIANPSCAAN
VACLNTCNNRPDETECQIKCGDLFENSVVDEFNECAVSRKKCVPKKSDVGEFPVPDPSSALVKNFNMA
DFTGKWYISSGLNPTFDTFDCQLHEFRVEGDRLVANLAWRIPTPDTGFFTRGAVQRFVQDPSQPAILY
NHDNEYLHYQDDWYILSSKIENKEDDYIFVYYRGRNDAWDGYGGAVVYTRSKELPETIVPELERAAKS
VGRDFSTFIRTDNTCGAEPPLADRIERTVEKGEKLIVDEVKEIEGEIEGEVKELEREEETLVKRLADG
IMEVKQDVMNFFQGLSKEEMEILDQLNLEATEVEELFSRSLPIRKLR

Zeaxanthin Epoxidases

>*At*ZEP_col
MGSTPFCYSINPSPSKLDFTRTHVFSPVSKQFYLDLSSFSGKPGGVSGFRSRRALLGVKAATALVEKE
EKREAVTEKKKKSRVLVAGGGIGGLVFALAAKKKGFDVLVFEKDLSAIRGEGKYRGPIQIQSNALAAL
EAIDIEVAEQVMEAGCITGDRINGLVDGISGTWYVKFDTFTPAASRGLPVTRVISRMTLQQILARAVG
EDVIRNESNVVDFEDSGDKVTVVLENGQRYEGDLLVGADGIWSKVRNNLFGRSEATYSGYTCYTGIAD
FIPADIESVGYRVFLGHKQYFVSSDVGGGKMQWYAFHEEPAGGADAPNGMKKRLFEIFDGWCDNVLDL
LHATEEEAILRRDIYDRSPGFTWGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSFQLALELDEAWKQSV
ETTTPVDVVSSLKRYEESRRLRVAIIHAMARMAAIMASTYKAYLGVGLGPLSFLTKFRVPHPGRVGGR
FFVDIAMPSMLDWVLGGNSEKLQGRPPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCV
SETLCLTKDEDQPCIVGSEPDQDFPGMRIVIPSSQVSKMHARVIYKDGAFFLMDLRSEHGTYVTDNEG
RRYRATPNFPARFRSSDIIEFGSDKKAAFRVKVIRKTPKSTRKNESNNDKLLQTA

>*At*ZEP_ler

MGSTPFCYSINPSPSKLDFTRTHVFSPVAKQFYLDLSSFSGRSGGGLSVFRSRKTLLGVKAATALVEK
EEKREAVTEKKKSRVLVAGGGIGGLVFALAAKKKGFDVLVFEKDLSAIRGEGKYRGPIQIQSNALAAL
EAIDIEVAEQVMEAGCITGDRINGLVDGISGTWYVKFDTFTPAASRGLPVTRVISRMTLQQILARAVG
EDVIRNESNVVDFEDSGDKVTVVLENGQRYEGDLLVGADGIWSKVRNNLFGRSEATYSGYTCYTGIAD
FIPADIESVGYRVFLGHKQYFVSSDVGGGKMQWYAFHEEPAGGADAPNGMKKRLFEIFDGWCDNVLDL
LHATEEEAILRRDIYDRSPGFTWGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSFQLALELDEAWKQSV
ETTTPVDVVSSLKRYEESRRLRVAIIHAMARMAAIMASTYKAYLGVGLGPLSFLTKFRVPHPGRVGGR
FFVDIAMPSMLDWVLGGNSEKLQGRPPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCV
SETLCLTKDEDQPCIVGSEPDQDFPGMRIVIPSSQVSKMHARVIYKDGAFFLMDLRSEHGTYVTDNEG
RRYRATPNFPARFRSSDIIEFGSDKKAAFRVKVIRKTPKSTRKNESNNDKLLQTA

>*At*ZEP_?
MGSTPFCYSINPSPSKLDFTRTHVFSPVSKQFYLDLSSFSGKPGGVSGFRSRRALLGVKAATALVEKE
EKREAVTDKKKKSRVLVAGGGIGGLVFALAAKKKGFDVLVFEKDLSAMRGEGKYRGPIQIQSNALAAL
EAIDIEVAEQVMEAGCITGDRINGLVDGISGTWYVKFDTFTPAGVTGLPVTRVISRMTLQQILARAVG
EDVIRNESNVVDFEDSGDKVTVVLENGQRYEGDLLVGADGIWSKVRNNLFGRSEATYSGYTCYTGIAD
FIPADIESVGYRVFLGHKQYFVSSDVGGGKMQWYAFHEEPAGGADAPNGMKKRLFEIFDGWCDNVLDL
LHATEEEAILRRDIYDRSPGFTWGKGRVTLLGDSIHAMQPNMGQGGCMAIEDSFQLALELDEAWKQSV
ETTTPVDVVSSLKRYEESRRLRVAIIHAMARMAAIMASTYKAYLGVGLGPLSFLTKFRVPHPGRVGGR
FFVDIAMPSMLDWVLGGNSEKLQGRPPSCRLTDKADDRLREWFEDDDALERTIKGEWYLIPHGDDCCV
SETLCLTKDEDQPCIVGSEPDQDFPGMRIVIPSSQVSKMHARVIYKDGAFFLMDLRSEHGTYVTDNEG
RRYRATPNFPARFRSSDIIEFGSDKKAAFRVKVIRKTPKSTRKNESNNDKLLQTA

>*Cu*ZEP
MVSSMFYNSVNLSTAVFSRTHFPVPVYKHSCIEFSRYDHCINYKFRTGTSGQSKNPTQMKAAVAESPT
NNSDSENKKLRILVAGGGIGGLVFALAAKRKGFEVLVFEKDMSAIRGEGQYRGPIQIQSNALAALEAI
DLDVAEEVMRAGCVTGDRINGLVDGISGSWYIKFDTFTPAAEKGLPVTRVISRMTLQQILAKAVGDEI
ILNESNVIDFKDHGDKVSVVLENGQCYAGDLLIGADGIWSKVRKNLFGPQEAIYSGYTCYTGIADFVP
ADIESVGYRVFLGHKQYFVSSDVGAGKMQWYAFHKEPAGGVDDPEGKKERLLKIFEGWCDNVVDLILA
TDEEAILRRDIYDRTPIFTWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDGYQLAVELEKACKKSNESK
TPIDIVSALKSYERARRLRVAVIHGLARSAAVMASTYKAYLGVGLGPLSFLTKFRIPHPGRVGGRFFI
DLAMPLMLSWVLGGNSSKLEGRSPCCKLSDKASDNLRTWFRDDDALERAMNGEWFLVPSGSENVVSQP
IYLSGSHENEPYLIGSESHEDFPRTSIVIPSAQVSKMHARISYKDGAFYLIDLQSEHGTYVTDNEGRR
YRVSSNFPARFRPSDTIEFGSDKKAIFRVKVIGTPPNNNSERKEAGEILQAV

>*Cr*ZEP
MLASTYTPCGVRQVAGRTVAVPSSLVAPVAVARSLGLAPYVPVCEPSAALPACQQPSGRRHVQTAATL
RADNPSSVAQLVHQNGKGMKVIIAGAGIGGLVLAVALLKQGFQVQVFERDLTAIRGEGKYRGPIQVQS
NALAALEAIDPEVAAEVLREGCITGDRINGLCDGLTGEWYVKFDTFHPAVSKGLPVTRVISRLTLQQI
LAKAVERYGGPGTIQNGCNVTEFTERRNDTTGNNEVTVQLEDGRTFAADVLVGADGIWSKIRKQLIGE
TKANYSGYTCYTGISDFTPADIDIVGYRVFLGNGQYFVSSDVGNGKMQWYGFHKEPSGGTDPEGSRKA
RLLQIFGHWNDNVVDLIKATPEEDVLRRDIFDRPPIFTWSKGRVALLGDSAHAMQPNLGQGGCMAIED
AYELAIDLSRAVSDKAGNAAAVDVEGVLRSYQDSRILRVSAIHGMAGMAAFMASTYKCYLGEGWSKWV
EGLRIPHPGRVVGRLVMLLTMPSVLEWVLGGNTDHVAPHRTSYCSLGDKPKAFPESRFPEFMNNDASI
IRSSHADWLLVAERDAATAAAANVNAATGSSAAAAAAADVNSSCQCKGIYMADSAALVGRCGATSRPA
LAVDDVHVAESHAQVWRGLAGLPPSSSSASTAAASASAASSAASGTASTLGSSEGYWLRDLGSGRGTW
VNGKRLPDGATVQLWPGDAVEFGRHPSHEVFKVKMQHVTLRSDELSGQAYTTLMVGKIRNNDYVMPES
RPDGGSQQPGRLVTA

>*Csp*ZEP
MHARSSLGPRARAGARAPAVCHVAALAACRPAASTPPQPLCATISNSAGSARAAIMPQQGRGLGSGLV
RCPTTPVATRASSVASAPASQPPAADSMKRPLRVLIAGAGIGGLVLAVALIKKGFHVTVFERDMTAIR

GEGKYRGPIQIQSNALGALEAIDPSIADEVMDEGCITGDRVNGLCDGVTGDWYVKFDTFHPAVSKGLP
VTRVISRVTLQNILAKAVLRYGGPDTIMSNSHVVGFEESNNGVSVTLENGDVHRGDILVGADGIWSKI
RKAILGETEANYSQYTCYTGISDFTPADIDIVGYRVFLGNSQYFVSSDVGGGKMQWYGFHKEPAGGTD
PEGQRKARLLDIFGHWNDNVVDLIKATPEEDIMRRDIFDRPPVFKWSEGRTVLLGDSVHAMQPNLGQG
GCMAIEDAYELANNLSDGMDAAGQQPAHLDVKKAFSTYQSHRMIRASAIHGMAGMAAFMASTYKAYLG
EGLPGPLQQLTKLKIHHPGRVVGRLVMNLTMPQVLGWVLGGNTENLDKSRVGHCRIADQPKAFHESQF
SYLMENDEAIIQSSHADWMLMTSREAGSGSSDSNARVDATADATSTSECKGIYIGDEPSIIGRKSESA
DLSINDGQVAPQHARVWRTETSSVSGRDVVAYEYHVQDLGSDAGTWLNGRPMPRGGTCQLHAGDVLEF
GQSPSKEVYRVKMQHVSLRNDKLNGHAFTTLVVGAHEGHGDKHMIMA

>*Le*ZEP
MYSTVFYTSVHPSTSVLSRKQLPLLISKDFSAELYHSLPCRSLENGHINKVKGVKVKATIAEAPVTPT
EKTDSGANGDLKVPQKKLKVLVAGGGIGGLVFALAAKKRGFDVLVFERDLSAIRGEGQYRGPIQIQSN
ALAALEAIDLDVAEDIMNAGCITGQRINGLVDGISGNWYCKFDTFTPAVERGLPVTRVISRMTLQQIL
ARAVGEEIIMNESNVVDFEDDGEKVTVVLENGQRFTGDLLVGADGIRSKVRTNLFGPSEATYSGYTCY
TGIADFVPADIDTVGYRVFLGHKQYFVSSDVGGGKMQWYAFYNEPAGGADAPNGKKERLLKIFGGWCD
NVIDLLVATDEDAILRRDIYDRPPTFSWGRGRVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELEKA
CSRSAEFGSPVDIISSLRSYESARKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTQYRIPHPG
RVGGRVFIDLGMPLMLSWVLGGNGDKLEGRIKHCRLSEKANDQLRKWFEDDDALERATDAEWLLLPAG
NGSSGLEAIVLSRDEDVPCTVGSISHTNIPGKSIVLPLPQVSEMHARISCKDGAFFVTDLRSEHGTWV
TDNEGRRYRTSPNFPTRFHPSDVIEFGSDKAAFRVKAMKFPLKTSERKEEREAVEAA

>*Nt*ZEP
MYSTVFYTSVHPSTSAFSRKQLPLLISKDFPTELYHSLPCSRSLENGQIKKVKGVVKATIAEAPATIP
PTDLKKVPQKKLKVLVAGGGIGGLVFALAAKKRGFDVLVFERDLSAIRGEGQYRGPIQIQSNALAALE
AIDMDVAEDIMNAGCITGQRINGLVDGVSGNWYCKFDTFTPAVERGLPVTRVISRMTLQQNLARAVGE
DIIMNESNVVNFEDDGEKVTVTLEDGQQYTGDLLVGADGIRSKVRTNLFGPSDVTYSGYTCYTGIADF
VPADIETVGYRVFLGHKQYFVSSDVGGGKMQWYAFHNEPAGGVDDPNGKKARLLKIFEGWCDNVIDLL
VATDEDAILRRDIYDRPPTFSWGKGRVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELDKALSRSAE
SGTPVDIISSLRSYESSRKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTKFRIPHPGRVGGRF
FIDLGMPLMLSWVLGGNGEKLEGRIQHCRLSEKANDQLRNWFEDDDALERATDAEWLLLPAGNSNAAL
ETLVLSRDENMPCNIGSVSHANIPGKSVVIPLPQVSEMHARISYKGGAFFVTDLRSEHGTWITDNEGR
RYRASPNFPTRFHPSDIIEFGSDKKAAFRVKVMKFPPKTAAKEERQAVGAA

>*Np*ZEP
MYSTVFYTSVHPSTSAFSRKQLPLLISKDFPTELYHSLPCSRSLENGQIKKVKGVVKATIAEAPATIP
PTDLKKVPQKKLKVLVAGGGIGGLVFALAAKKRGFDVLVFERDLSAIRGEGQYRGPIQIQSNALAALE
AIDMDVAEDIMNAGCITGQRINGLVDGVSGNWYCKFDTFTPAVERGLPVTRVISRMTLQQNLARAVGE
DIIMNESNVVNFEDDGEKVTVTLEDGQQYTGDLLVGADGIRSKVRTNLFGPSDVTYSGYTCYTGIADF
VPADIETVGYRVFLGHKQYFVSSDVGGGKMQWYAFHNEPAGGVDDPNGKKARLLKIFEGWCDNVIDLL
VATDEDAILRRDIYDRPPTFSWGKGRVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELDKALSRSAE
SGTPVDIISSLRSYESSRKLRVGVIHGLARMAAIMASTYKAYLGVGLGPLSFLTKFRIPHPGRVGGRF
FIDLGMPLMLSWVLGGNGEKLEGRIQHCRLSEKANDQLRNWFEDDDALERATDAEWLLLPAGNSNAAL
ETLVLSRDENMPCNIGSVSHANIPGKSVVIPLPQVSEMHARISYKGGAFFVTDLRSEHGTWITDNEGR
RYRASPNFPTRFHPSDIIEFGSDKKAAFRVKVMKFPPKTAAKEERQAVGAA

>*Os*ZEP
MALLSATAPAKTRFSLFSHEEAQHPHPHALSACCGGGASGKRQRARARVAAAMRPADAAASVAQAASP
GGGGEGTRRPRVLVAGGGIGGLVLALAARRKGYEVTVFERDMSAVRGEGQYRGPIQIQSNALAALAAI
DMSVAEEVMREGCVTGDRINGLVDGISGSWYIKFDTFTPAAERGLPVTRVISRMTLQQILARAVGDDA
ILNDSHVVDFIDDGNKVTAILEDGRKFEGDLLVGADGIWSKVRKVLFGQSEATYSEYTCYTGIADFVP
PDIDTVGYRVFLGHKQYFVSSDVGAGKMQWYAFHKEPAGGTDPENGKNKRLLEIFNGWCDNVVDLINA

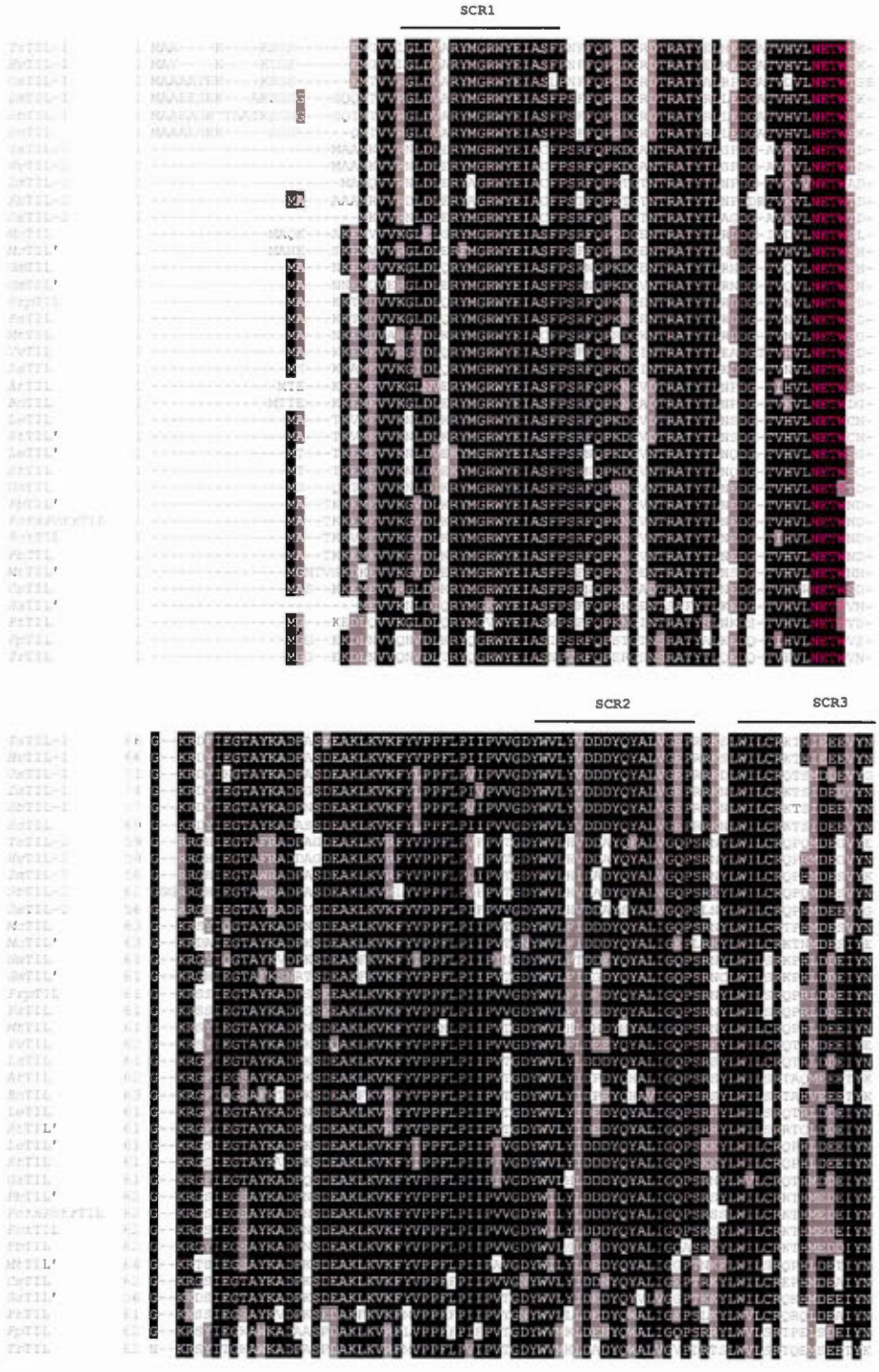
TDEEAILRRDIYDRPPTFNWGKGRVTLLGDSVHAMQPNLGQGGCMAIEDGYQLAVELEKSWQESAKSG
TPMDIVSSLRRYEKERILRVSVIHGLARMAAIMATTYRPYLGVGLGPLSFLTKLRIPHPGRVGGRFFI
KYGMPLMLSWVLGGNSTKLEGRPLSCRLSDKANDQLRRWFEDDDALEQAMGGEWYLLPTSSGDSQPIR
LIRDEKKSLSIGSRSDPSNSTASLALPLPQISENHATITCKNKAFYVTDNGSEHGTWITDNEGRRYRR
TSELPCPFPSLGCH

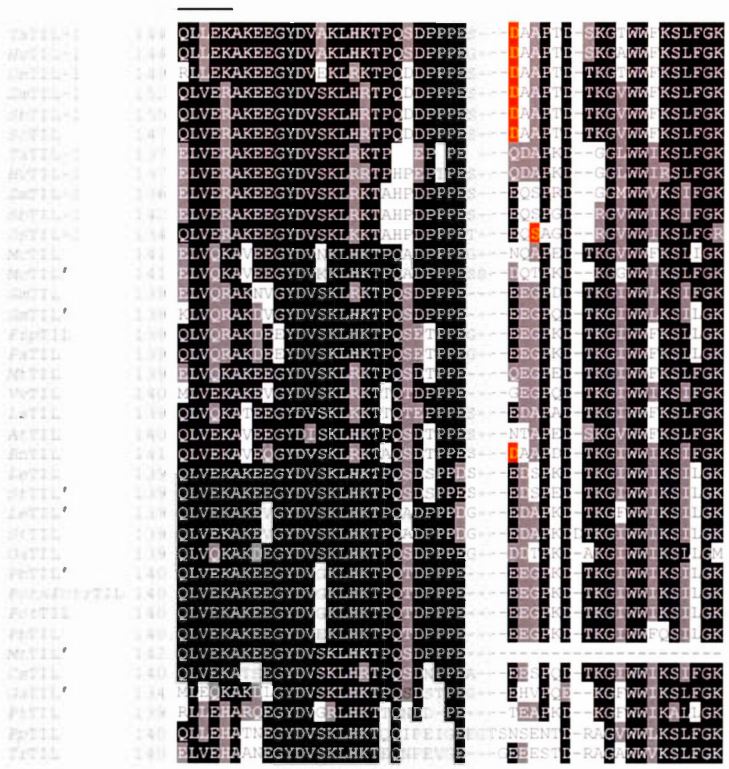
>*Pa*ZEP
MASTLFYNSMNLSAAVFSRTHFPIPINKDFPLEFSPCIHTDYHLRSRTRSGQKKCLTEVRATVASPTE
VPSAPASTQPKKLRILVAGGGIGGLVFALAAKKKGFDVVVFEKDLSAVRGEGQYRGPIQIQSNALAAL
EAIDMDVAEEVMRVGCVTGDRINGLVDGVSGTWYVKFDTFTPAVERGLPVTRVISRIALQQILARAVG
EEIIINDSNVVNFEDLGDKVNVILENGQRYEGDMLVGADGIWSKVRKNLFGLNEAVYSGYTCYTGIAD
FVPADINSVGYRVFLGHKQYFVSSDVGGGKMQWYAFHKESPGGVDSPNGKKERLLKIFEGWCDNVIDL
LLATEEDAILRRDIYDRTPILTWGKGHVTLLGDSVHAMQPNMGQGGCMAIEDGYQLALELDKAWKKSS
ETGTPVDVASSLRSYENSRRLRVAIIHGMARMAALMASTYKAYLGVGLGPLSFLTKFRIPHPGRVGGR
VFIDKAMPLMLSWVLGGNSSKLEGRSPSCRLSDKASDQLRNWFEDDDALERAIDGEWYLIPCGQDNDA
SQLICLNRDEKNPCIIGSAPHGDVSGISIAIPKPQVSEMHARISYKDGAFYLTDLRSEHGTWIADIEG
KRYRVPPNFPARFRPSDAIEIGSQKVAFRVKVMKSSPGSVEKEGILQAA

>*Ca*ZEP
MYASSARDGIPGKWCNARRKQLPLLISKDFPAELYHSLPCKSLENGHIKKVKGVKATLAEAPATPTEK
SNSEVPQKKLKVLVAGGGIGGLVFALAGKKRGFDVLVFERDISAIRGEGQYRGPIQIQSNALAALEAI
DMDVAEEIMNAGCITGQRINGLVDGISGNWYCKFDTFTPAVERGLPVTRVISRMTLQQILARLQGEDV
IMNESHVVNFADDGETVTVNPELCQQYTGDLLVGADGIRSKVRTNLFGPSELTYSGYTCYTGIADFVP
ADIDTAGYRVFLGHKQYFVSSDVGGGKMQWYAFHNEPAGGVDAPNGKKERLLKIFGGWCDNVIDLSVA
TDEDAILRRDIYDRPPTFSWGKGRVTLLGDSVHAMQPNLGQGGCMAIEDSYQLALELEKAWSRSAESG
SPMDVISSLRSYESARKLRVGVIHGLARMAAIMASAYKAYLGVGLGPLSFITKFRIPHPGRVGGRFFI
DLGMPLMLSWVLGGNGEKLEGRIQHCRLSEKANDQLRNWFEDDDALERATDAEWLLLPAGNSNAALET
LVLSRDENMPCTIGSVSHANIPGKSVVIPLSQVSDMHARISYNGGAFLGTAFRSDHGTWFIDNEGRRY
RVSPNFPMRFHSSDVIVFGSDKAAFRIKAMKFAPKTAAKEDRQAVGAA

Supplemental Figure 1. Alignment of the deduced amino acid sequences of TIL lipocalins.

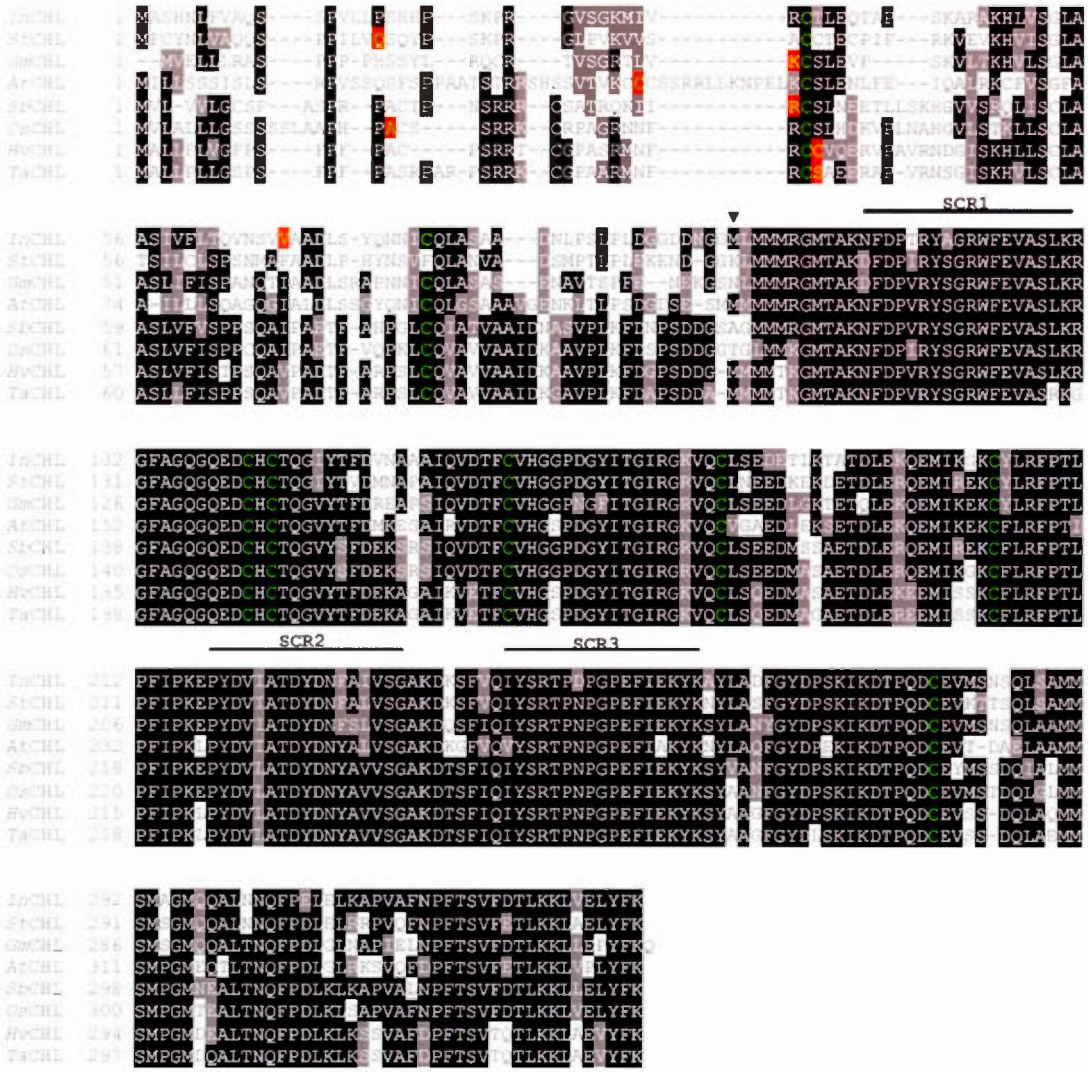
Identical residues are in black and similar residues are in grey. The three SCRs that provide a signature for the lipocalins are indicated above. Conserved N-glycosylation sites are in pink. Putative cleavage sites with a DGPI score over 0.700 are in red with yellow letters.


SCR1
SCR2
SCR3



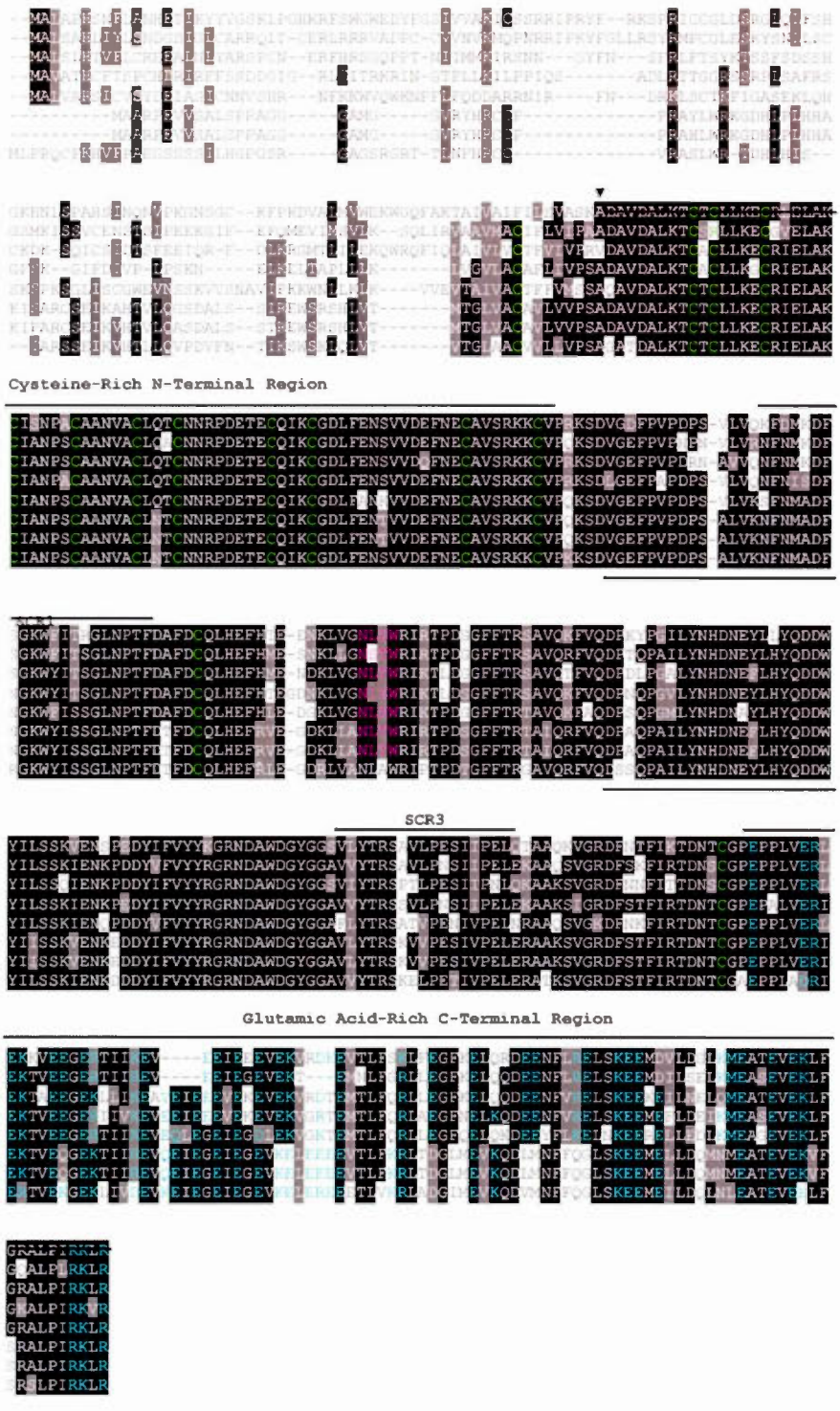
Supplemental Figure 2. Alignment of the deduced amino acid sequences of CHL lipocalins.

Identical residues are in black and similar residues are in grey. The three SCRs that provide a signature for the lipocalins are indicated above. Conserved cysteine residues are in green. Putative chloroplastic transit peptide cleavage sites identified with Signal P / ChloroP software are in red with yellow letters. The triangle indicates the putative chloroplastic transit peptide cleavage site based on sequence comparison.


SCR1
SCR2
SCR3

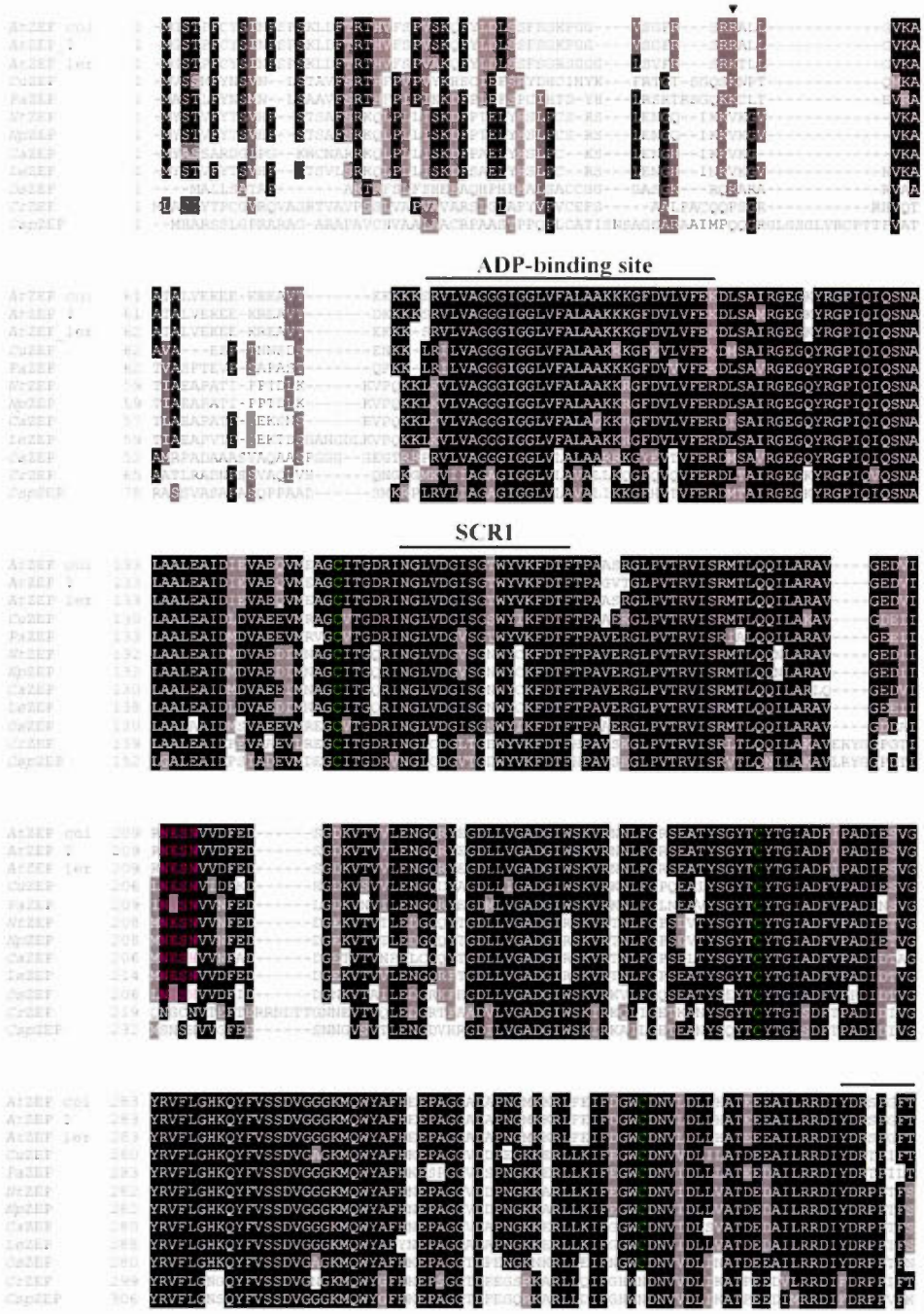
Supplemental Figure 3. Alignment of the deduced amino acid sequences of VDE proteins.

Identical residues are in black and similar residues are in grey. The SCRs that provide a signature for the lipocalins, the cysteine-rich N-terminal region and the glutamic acid-rich C-terminal region are indicated above. Conserved N-glycosylation sites are in pink. Conserved cysteine residues are in green. Conserved glutamic acid residues and other charged residues are in turquoise. The triangle indicates the chloroplastic transit peptide cleavage site.

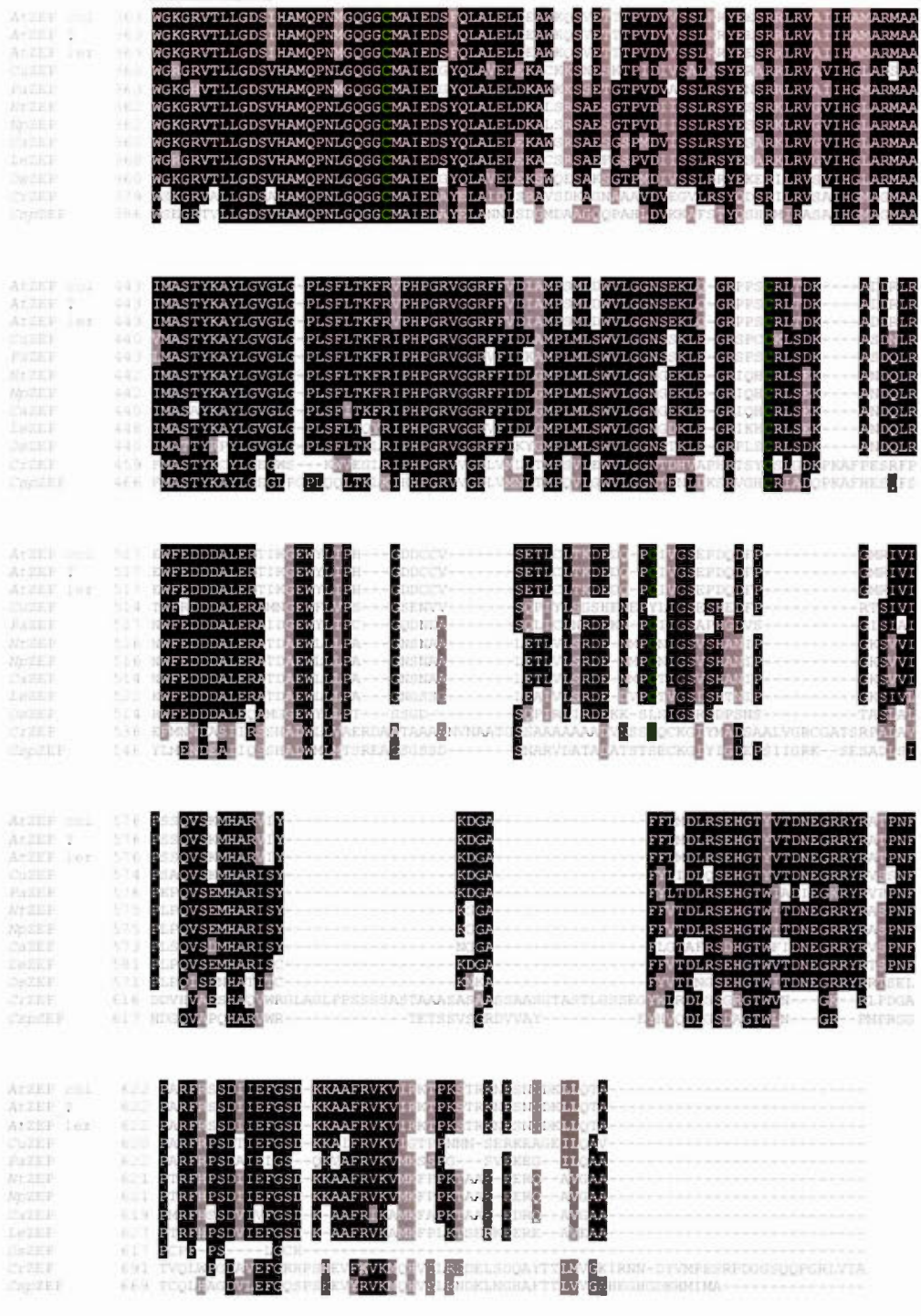

Cysteine-Rich N-Terminal Region
SCR3
Glutamic Acid-Rich C-Terminal Region

Supplemental Figure 4. Alignment of the deduced amino acid sequences of ZEP proteins.

Identical residues are in black and similar residues are in grey. The SCRs that provide a signature for the lipocalins and other conserved motifs are indicated above. Conserved N-glycosylation sites are in pink. Conserved cysteine residues are in green. The triangle indicates the chloroplastic transit peptide cleavage site.

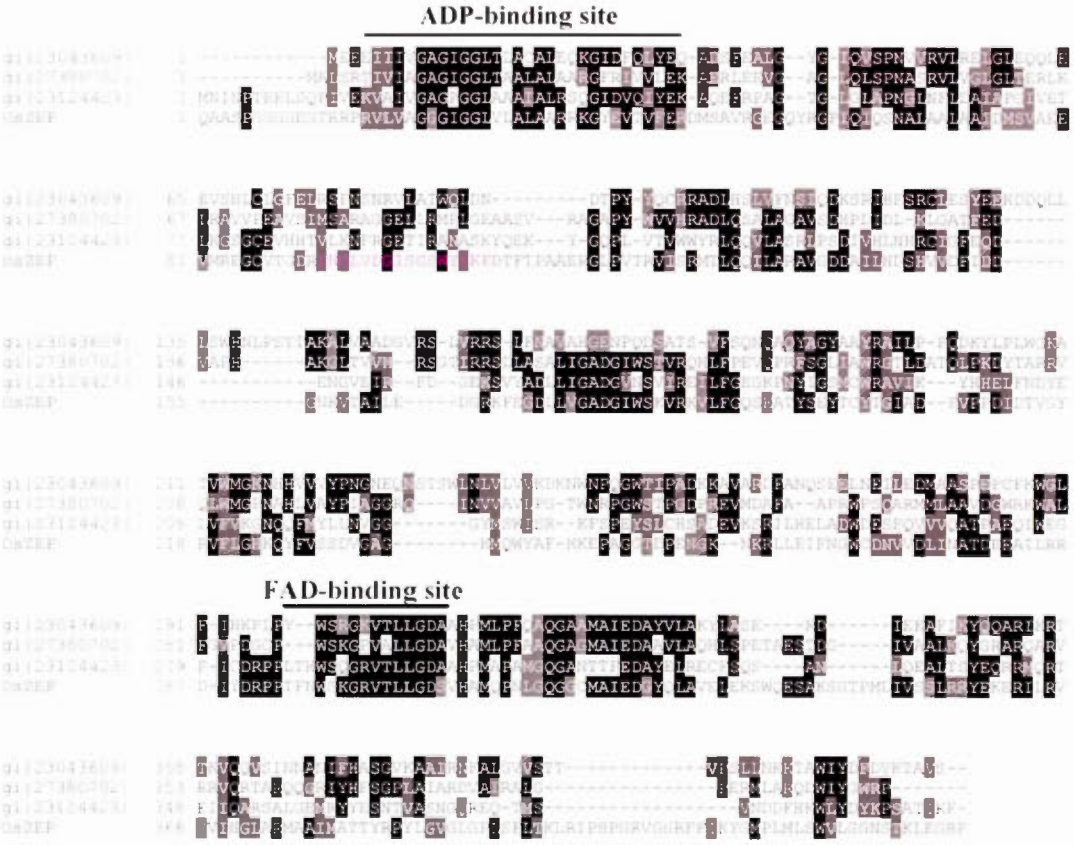

ADP-binding site
SCR1

FAD-binding site



Supplemental Figure 5. Alignment of *Os*ZEP with mono-oxygenases and related FAD-dependent oxydases from bacteria and cyanobacteria.

Identical residues are in black and similar residues are in grey. Two conserved motifs are indicated above. The *Os*ZEP SCR1 is indicated by pink letters.


ADP-binding site
FAD-binding site

Supplemental Figure 6. Cellular localization of the plant TIL lipocalins. (color)

A. Schematic representation of GFP fusions used in the transient expression experiments. N and C are the amino and carboxy termini of the proteins, respectively; 1, 2 and 3 indicate the 3 structurally conserved regions (SCRs).

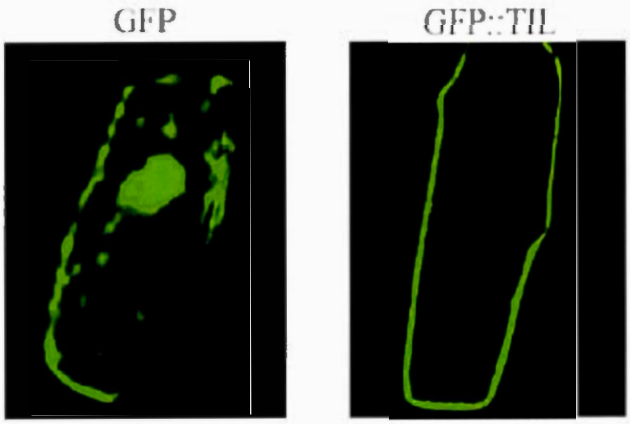
B. Transient expression assays of GFP-TIL fusions. Plasmids carrying the fusions were transformed into onion epidermal cells by microprojectile bombardment. Confocal images of GFP fluorescence were captured 20 hours after transformation. Only the GFP::*At*TIL data is shown since the three constructs gave the same fluorescence pattern. The color figure is shown in Supplemental Fig. 6.

C. Biochemical fractionation analysis. Wheat protein extracts were prepared and subjected to SDS-PAGE and western blot analyses. Upper panels, western blot results obtained using the anti-*Ta*TIL antibody (dil. 1/25,000, 10 sec. exposure for the plasma membrane fractions; dil. 1/2,500, 5 min. exposure for the other fractions). Lower panel, Coomassie Brilliant Blue-stained gel showing the quality of the preparations. Typical protein patterns are observed for each fraction. NA, non-acclimated plants grown for 7 days; CA7, plants grown for 7 days at 24°C then cold acclimated at 4°C for 7 days.

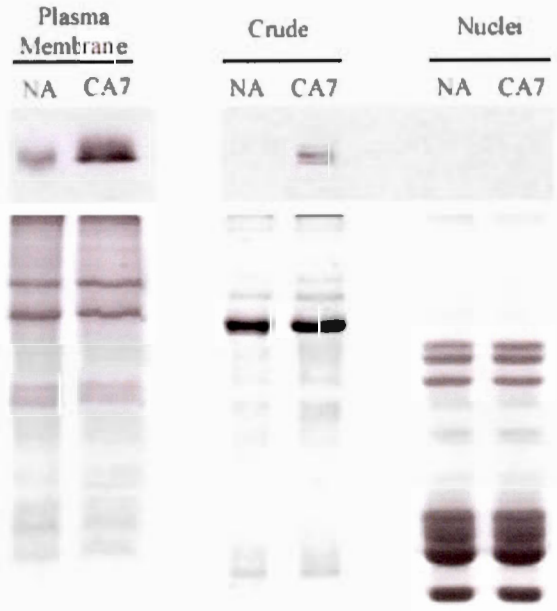
A


GFP TIL
N
1
2
3
C
TIL GFP
N
1
2
3
C
TI GFP IL
N
1
2
3
C

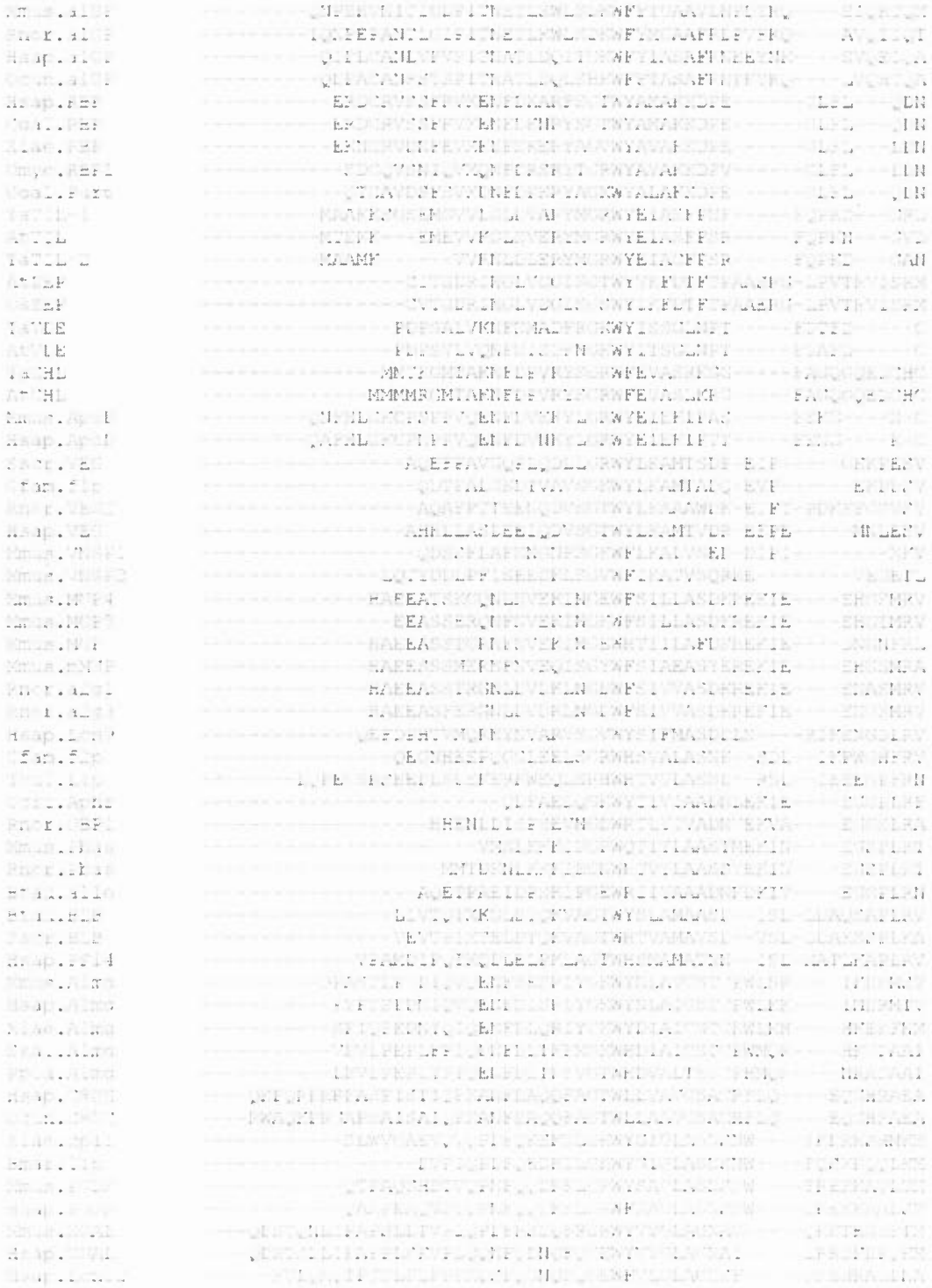
B

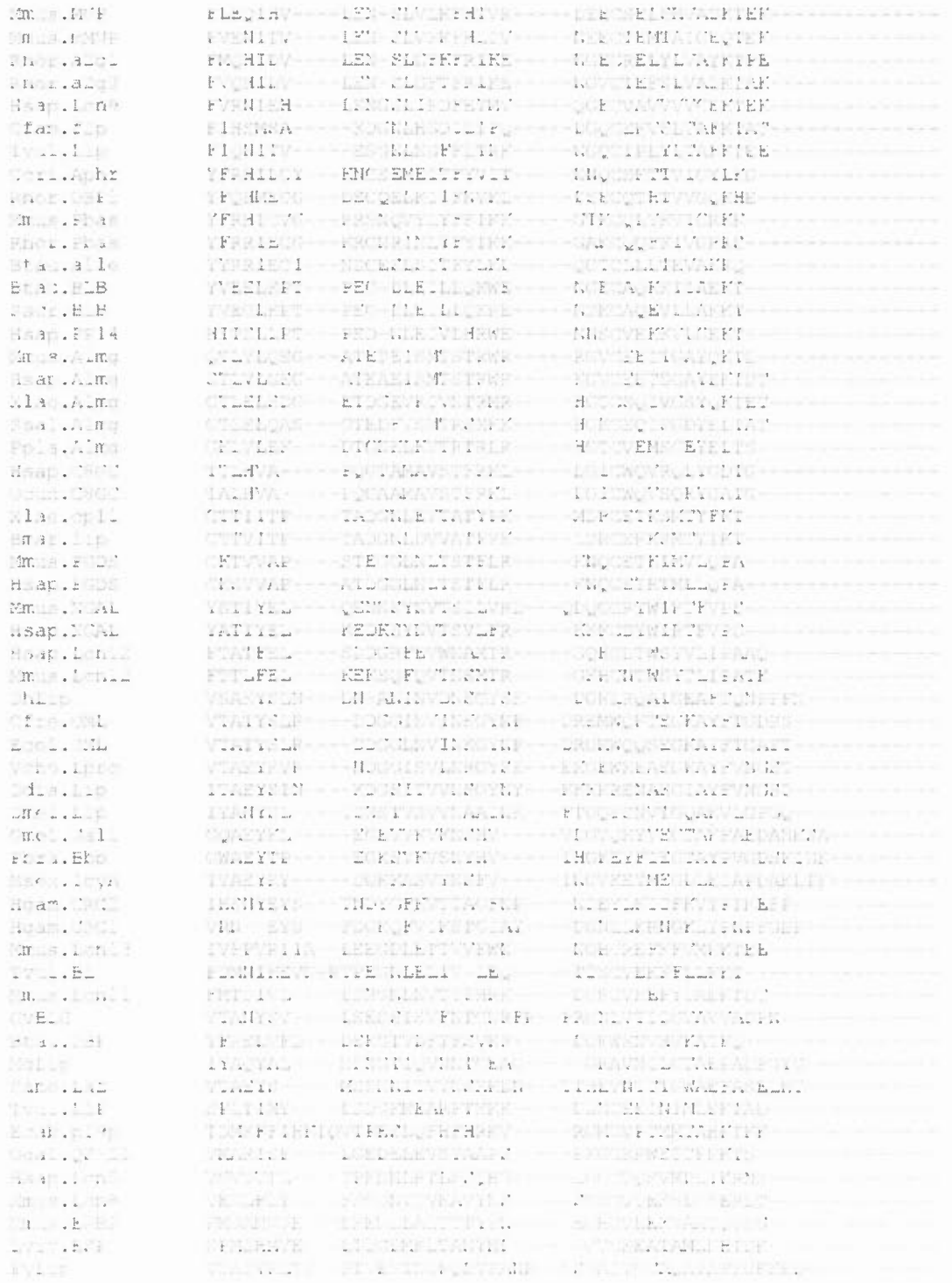

GFP
GFP::TIL

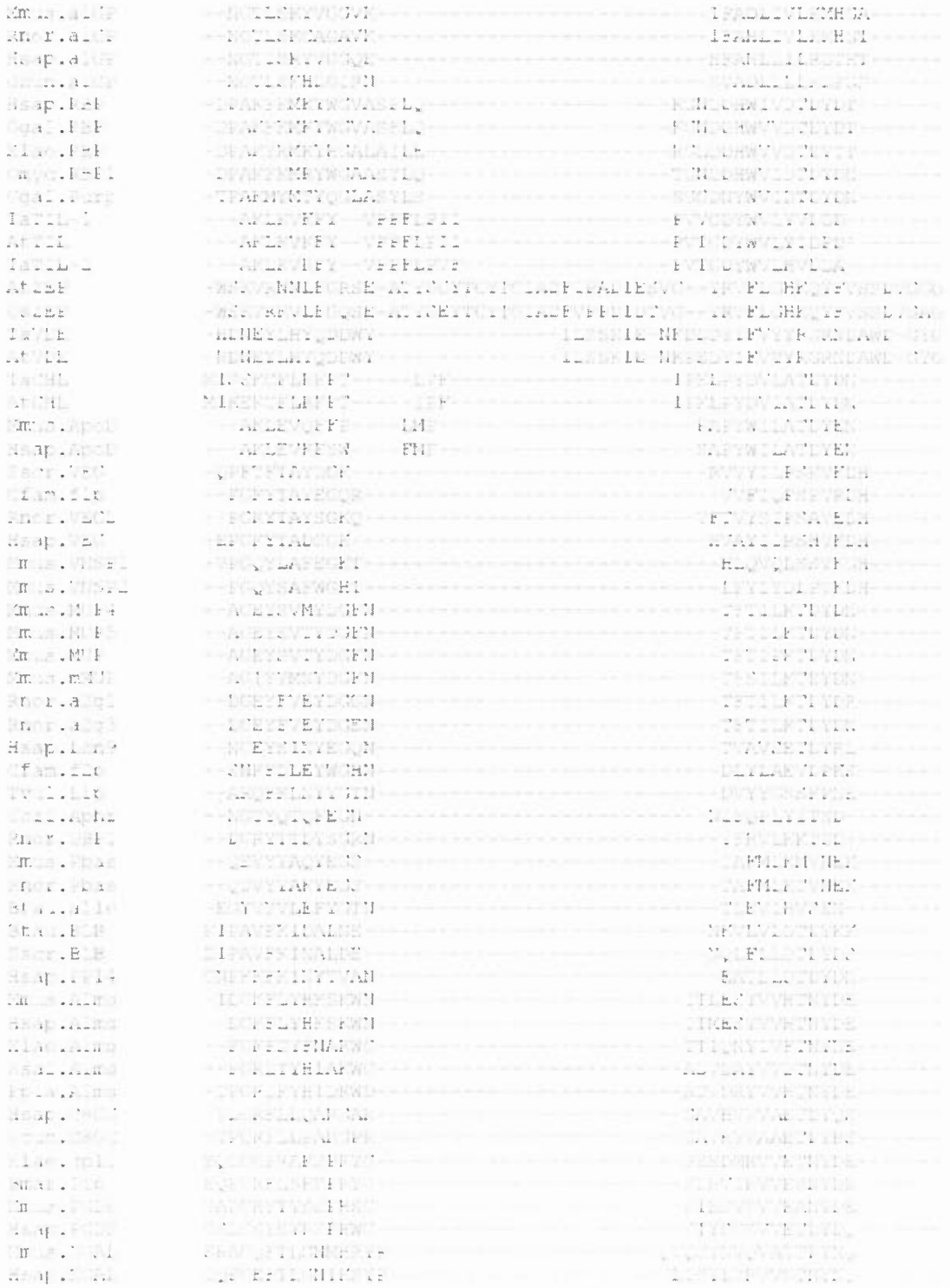
C

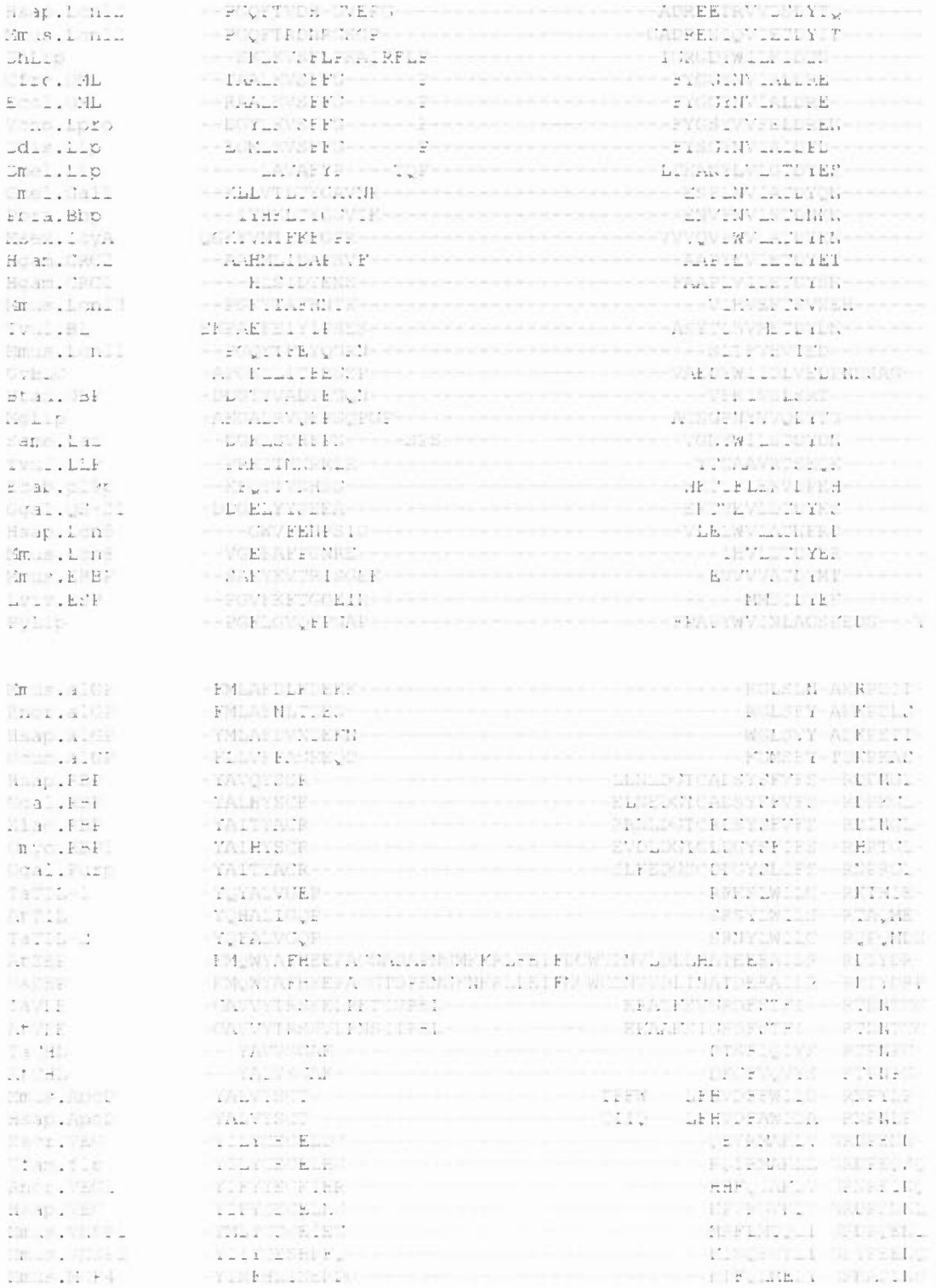

Plasma Membrane
NA
CA7
Crude
NA
CA7
Nuclei
NA
CA7

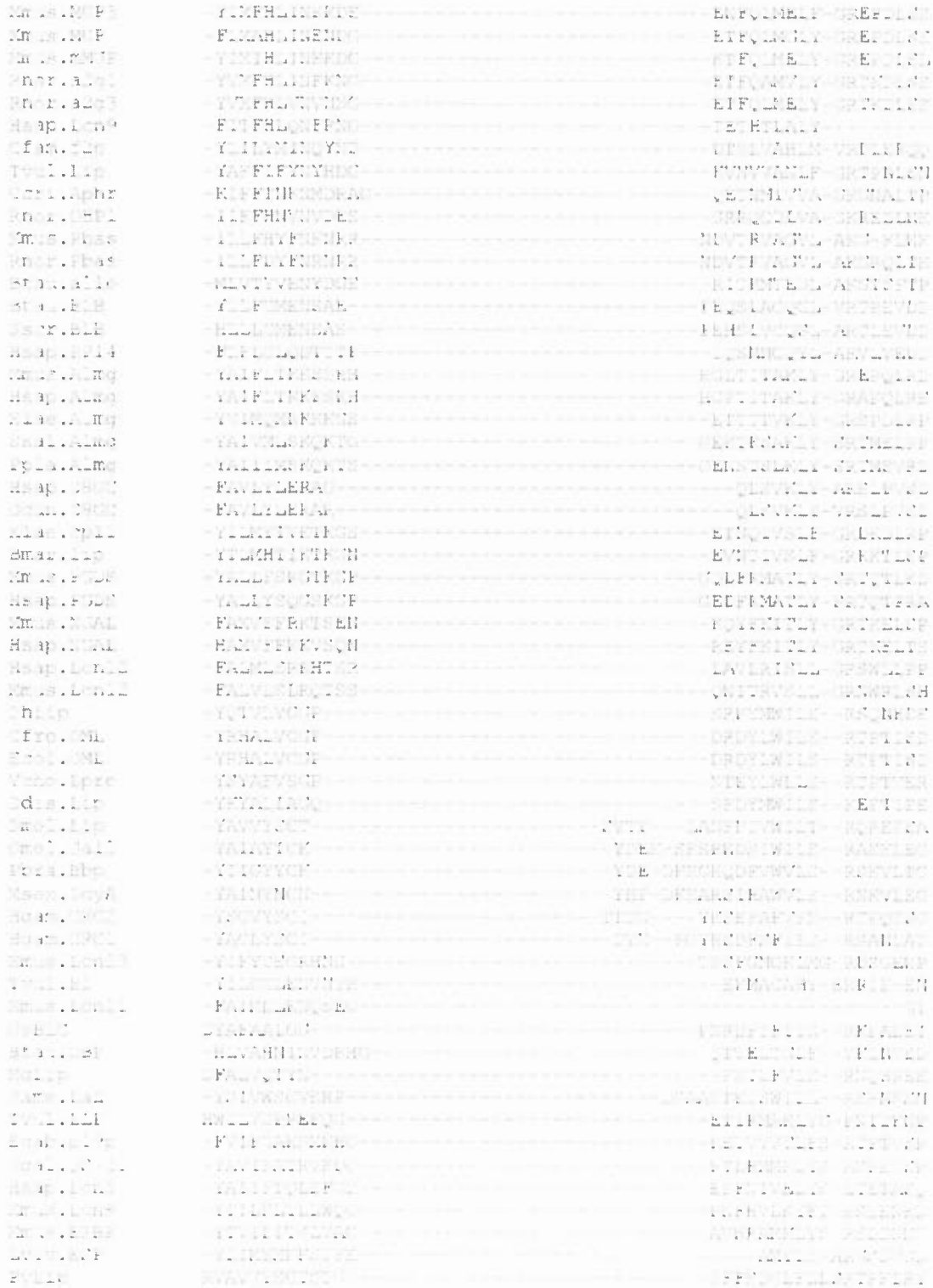
Supplemental Figure 7. Multiple sequence alignment of plant lipocalins, plant lipocalin-like proteins and other selected lipocalins. The alignment was generated using CLUSTAL X v.1.83.

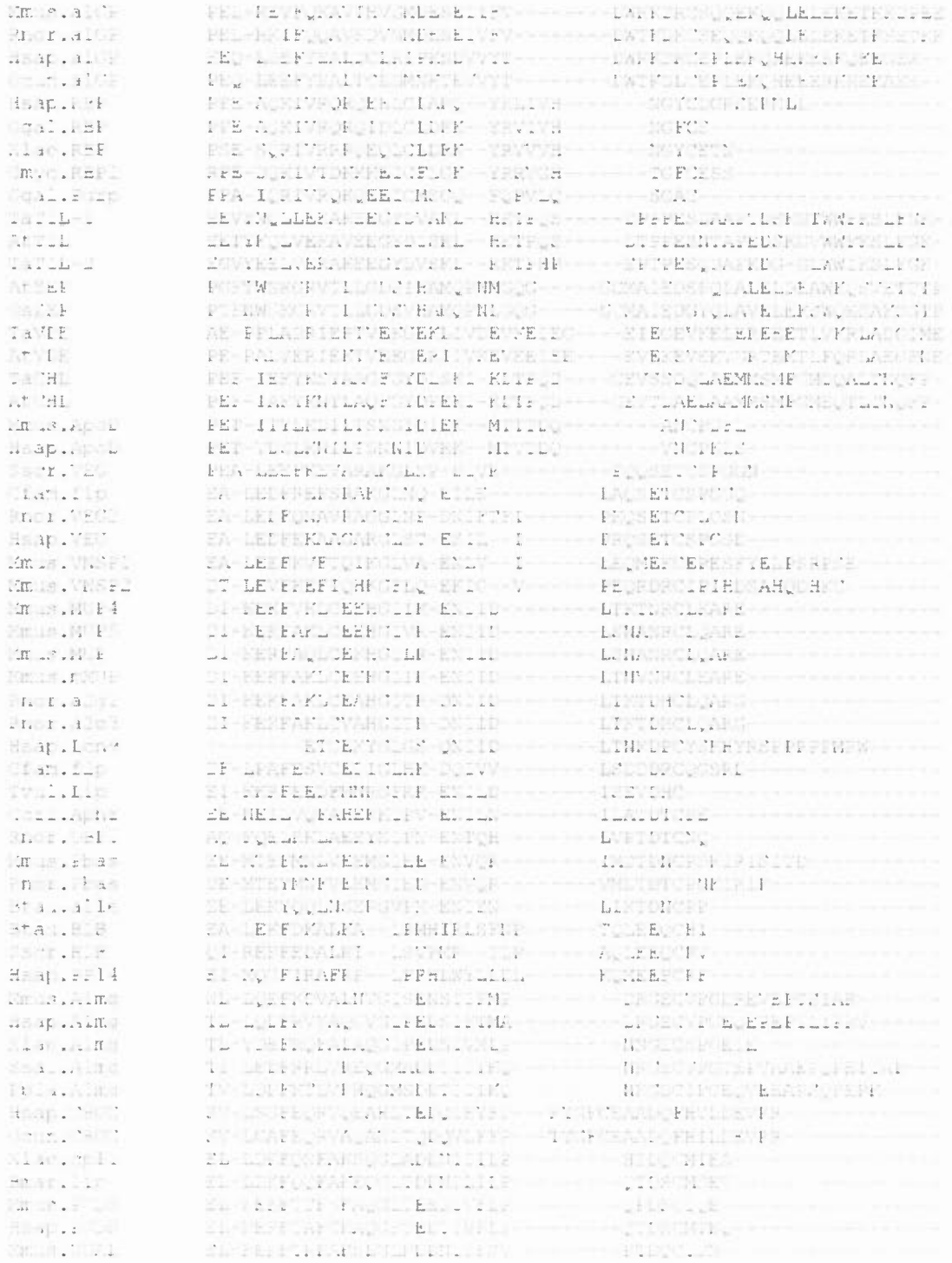


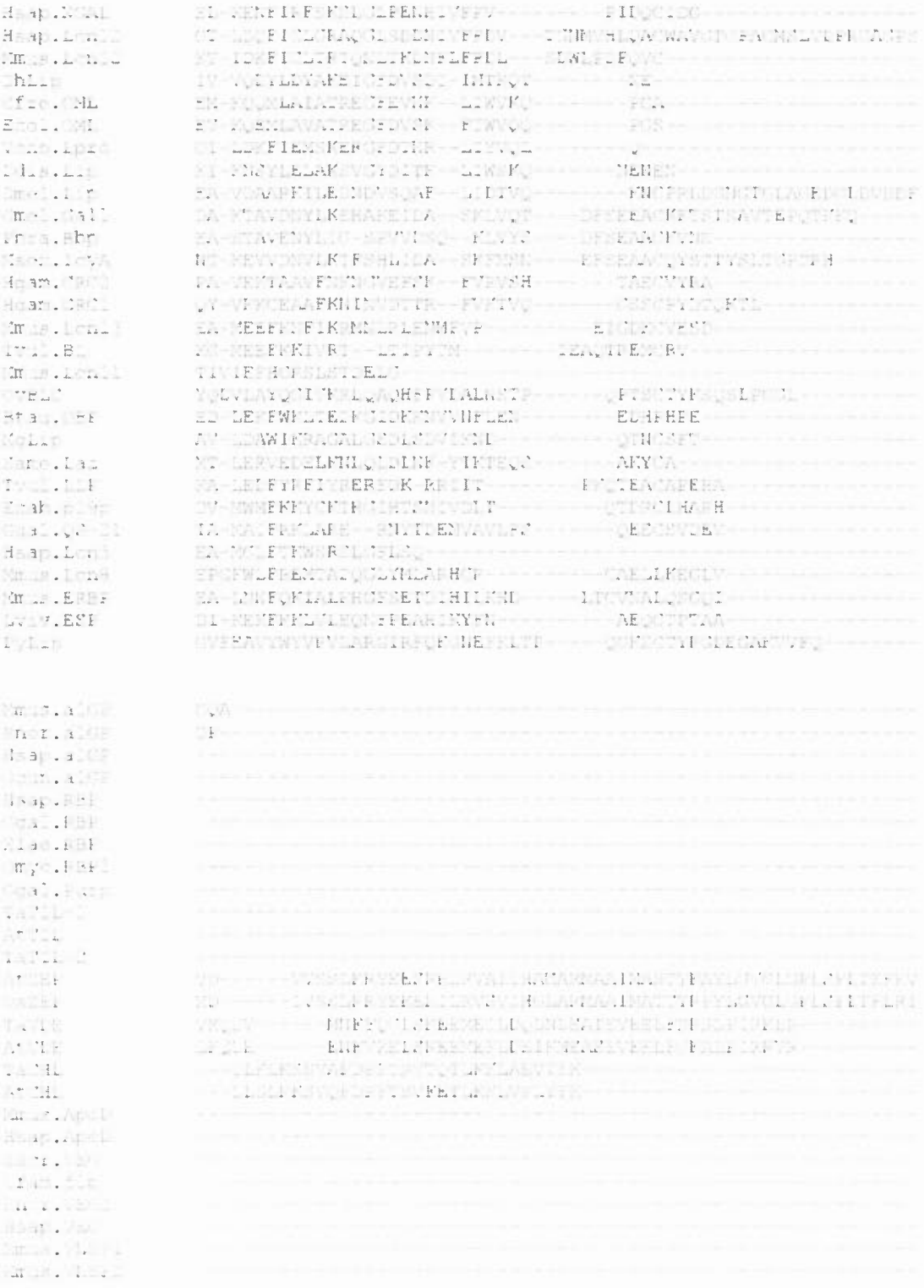


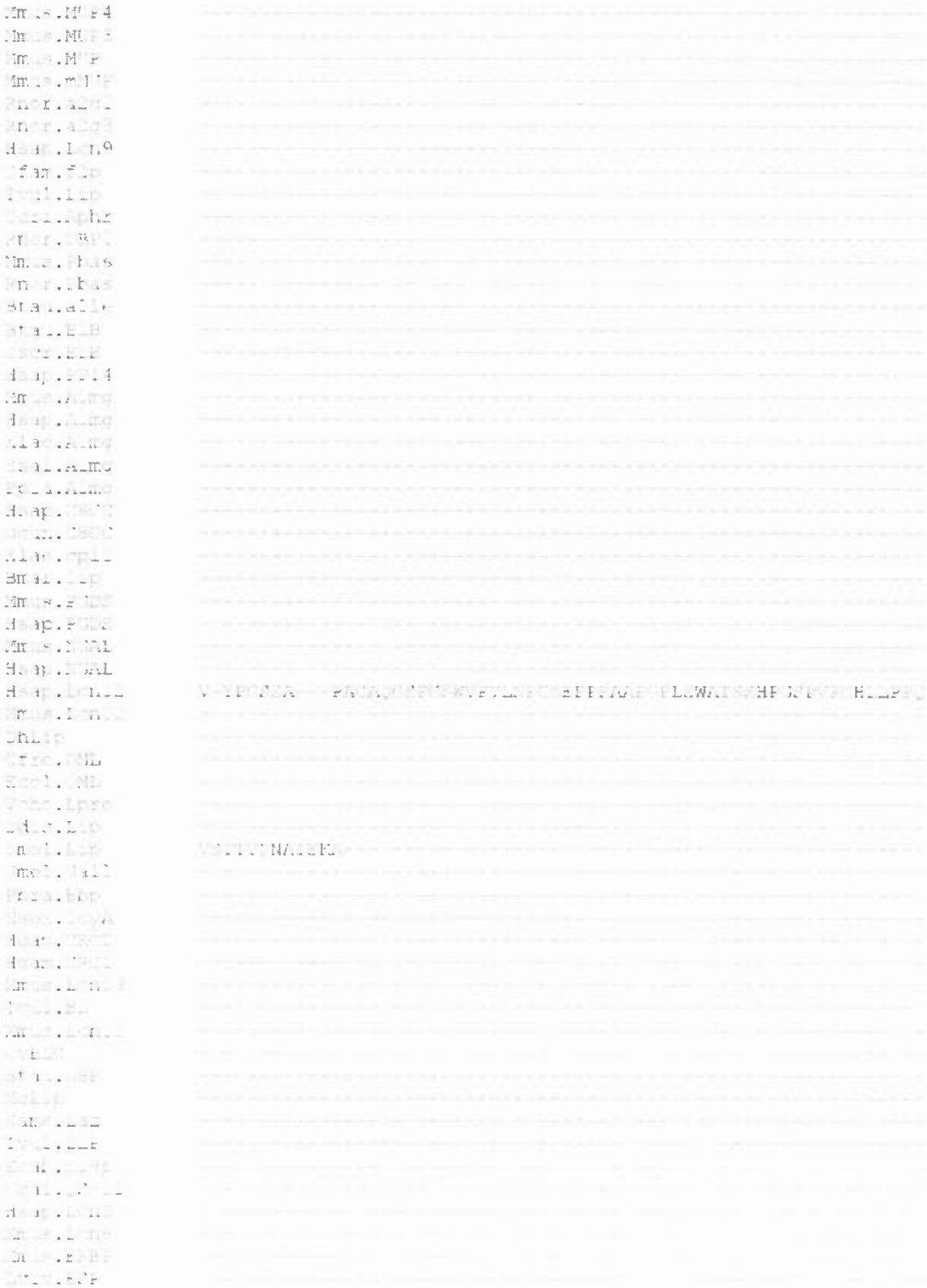


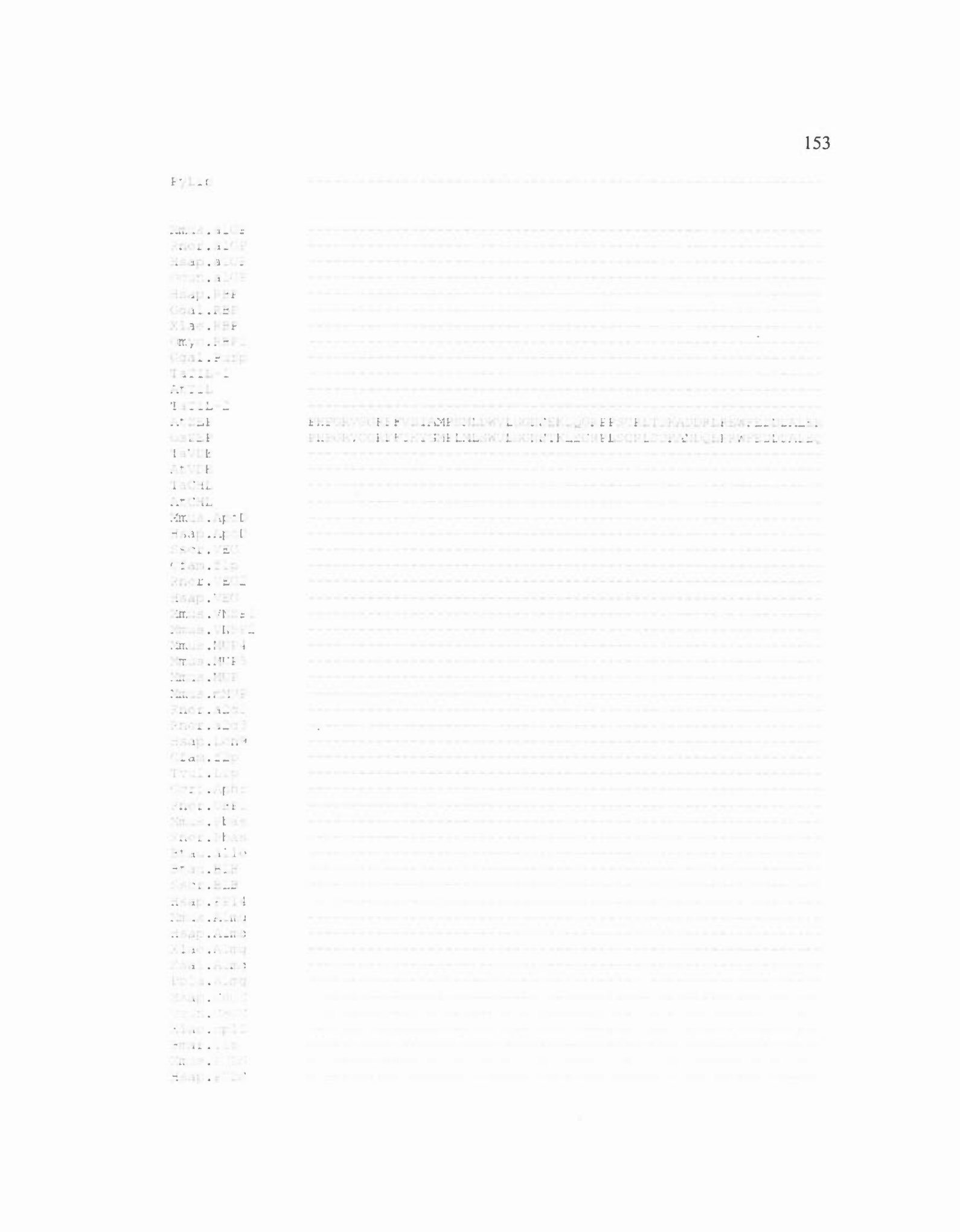


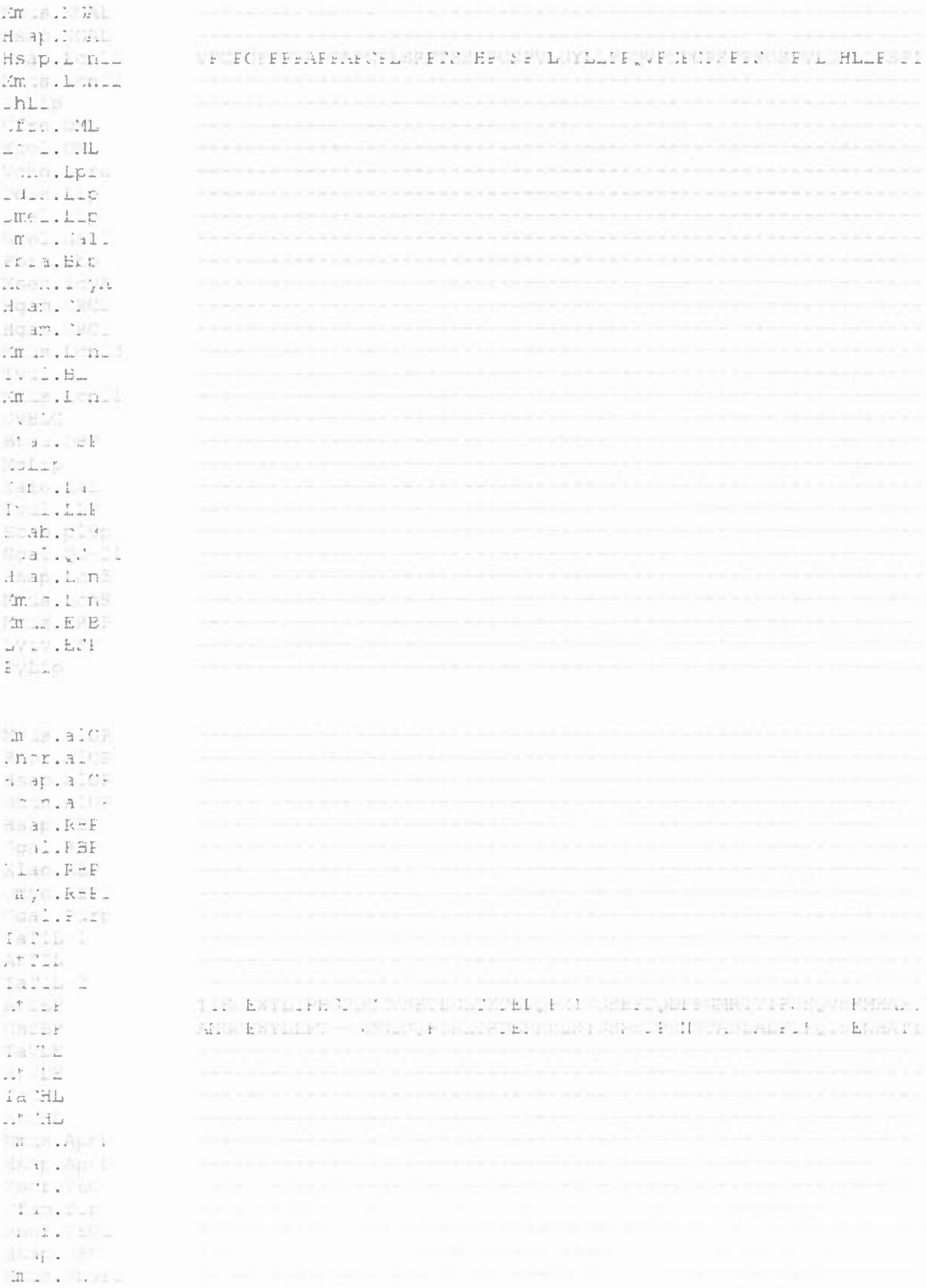












CHAPITRE IV

The ApoD ortholog *At*TIL protects *Arabidopsis* against oxidative stress and delays senescence.

Jean-Benoit F. Charron, Francois Ouellet, Mario Houde, and Fathey Sarhan

Pour les travaux associés à cet article, j'ai élaboré le design expérimental et effectué les analyses phénotypiques. J'ai aussi rédigé le manuscrit et conçu les figures. François Ouellet, a participé à l'élaboration du design expérimental, à la rédaction et plus particulièrement à l'édition finale du manuscrit. Mario Houde a supervisé la partie biopuces à ADN. Il a aussi participé à la rédaction des divers textes concernant cette section.

Résumé

Les stress environnementaux, tels que le froid et le stress thermique, induisent l'accumulation des lipocalines chez les végétaux. Cependant, la fonction cellulaire et le mode d'action de ces protéines demeurent inconnus. L'utilisation de différentes approches de type gain ou perte de fonction nous a permis de démontrer que chez *Arabidopsis*, la lipocaline *At*TIL est impliquée dans la modulation de la tolérance au stress oxydatif. Comparativement à des plantes témoins, les plantes n'accumulant pas *At*TIL sont très sensibles aux baisses soudaines de températures et à l'exposition au paraquat. Ce phénotype peut être renversé lorsque l'accumulation de cette protéine est rétablie. Inversement, la surexpression de *At*TIL augmente la tolérance des plantes à ces deux stress. De plus, l'accumulation de *At*TIL retarde la floraison et la sénescence de la plante. Pour sa part, l'analyse biopuces d'ADN indique que l'absence de cette lipocaline affecte l'expression de 66 gènes. Parmi ceux-ci, on dénombre les facteurs de transcription impliqués dans le contrôle de l'horloge circadienne et plusieurs gènes impliqués dans la balance énergétique de la plante. Ces données suggèrent que *At*TIL affecte une voie métabolique alternative qui module le niveau d'énergie cellulaire dans le but d'accroître la tolérance au stress oxydatif. Des résultats similaires ont été obtenus chez la *Drosophile* surexprimant une lipocaline (tolérance accrue au stress oxidatif et une plus longue longévité), suggèrant que les lipocalines possèdent une fonction conservée entre les espèces.

Mots clés : apolipoprotéine D; *Arabidopsis thaliana*; lipocaline; paraquat, sénescence; stress oxydatif; tolérance au gel

Abstract

The cellular functions of plant lipocalins are still unknown. Here we demonstrate that the *Arabidopsis At*TIL lipocalin is involved in modulating tolerance to oxidative stress. *AtTIL* knock-out plants are very sensitive to sudden drops in temperature and paraquat treatment, and dark-grown plants die shortly after transfer to light. Complementation restores the normal phenotype and overexpression enhances tolerance to the oxidative stress caused by freezing, paraquat and light. Moreover, the accumulation of *At*TIL delays flowering and senescence. Microarray analyses identified several differentially-regulated genes encoding components of oxidative stress, energy balance and circadian clock. These findings support the data in the accompanying two papers showing that overexpression of *ApoD*, a homologue of *At*TIL, enhances tolerance to oxidative stress and increases life span in mice and *Drosophila*..

Introduction

Over 40 lipocalin members have been identified from all kingdoms (*1*). The bacterial lipocalin (Blc), the mammalian apolipoprotein D (ApoD), and the insect Lazarillo protein (GLaz) are among the most studied lipocalins. They show a simple tertiary structure which gives them the ability to bind small, generally hydrophobic, molecules. It was recently shown that GLaz possesses a protective role against oxidative stress conditions and that its absence increases lipid peroxydation, reduces life span and accelerates neurodegeneration in *Drosophila* (*2*). On the other hand, its overexpression protects against the effects of starvation, hypoxia and hyperoxia, and extends the fly's life span (*3*).

Plants also possess lipocalins, which were classified as temperature-induced lipocalins (TILs) and chloroplastic lipocalins (CHLs), and lipocalin-like proteins (*4*). The *TIL* genes are induced by high and low temperature (LT). The TIL properties, their association with the plasma membrane in photosynthetic tissues, and their accumulation in response to temperature stress support the hypothesis that these proteins may act as scavengers of potentially harmful molecules known to be induced by temperature stress and excess light (*5*). In this report, we characterized knock-out, complementation and overexpression lines to determine the cellular and biochemical functions of the *At*TIL lipocalin.

Results and discussion

Four T-DNA lines carrying insertions in the *AtTIL* gene were analyzed to identify knock- out (KO) lines (Fig. 1A and S1) (*6*). The SALK_136775 line carries a single insertion in the first exon of the *AtTIL* gene and shows no detectable *At*TIL expression (Fig. 1B). Complementation of this KO line restored *At*TIL protein accumulation to a level about two-fold higher than the wild-type (WT) plants (Fig. 1B, Comp). Overexpression of the *AtTIL* cDNA resulted in a 4-fold accumulation of the protein compared to the WT (Fig. 1B, OEX). Under normal growth conditions, the downregulation of *AtTIL* expression (KO) has no visible effect on plant growth and development (Fig. 1C). In contrast, the Comp and OEX lines show a delay in flowering and a stay-green phenotype (Fig. 1C and 1D).

The effect of light stress on the different lines was monitored under different fluence rates (Fig. 2). Dark-grown KO plants show a reduction in hypocotyl elongation that is reversed in the Comp line, while OEX plants show longer hypocotyls (Fig. 2A). When dark-grown WT, Comp, OEX and control plants are transfered to normal photoperiod and light conditions, greening occurs and typical development resumes. In contrast, dark-grown KO plants transferred to light do not accumulate chlorophyll and die shortly after (data not shown). This suggests that the absence of lipocalin impairs the plant's ability to adapt to a sudden light exposure. Under a continuous moderate fluence of 100 μmol^{-1} sec^{-2}, the different lines show similar hypocotyl elongation (Fig. 2B), but the KO plants have smaller cotyledons (Fig. 2C, LL). There was no difference in growth and development when plants of the different lines were grown under a normal light-dark cycle (Fig. 2C, LD). These data indicate that in the absence of *AtTIL*, plants cannot tolerate the stress generated by continuous light and that a dark period is needed for the plants to recover.

The different *Arabidopsis* lines are similar after 21 days of growth under control non-acclimated (NA) conditions or after 7 days of cold acclimation (CA) (Fig. 3A). After NA plants were subjected to a freezing test at -6°C, a survival rate of

75% was obtained for the wild-type (WT) and control (vector) plants (Fig. 3B). In contrast, only 50% of the KO plants survived while the survival rate of the Comp and OEX plants was more than 90%. Freezing treatment caused typical damage and necrosis to the leaves of surviving WT and KO plants. On the other hand, leaves of OEX plants showed no damage after freezing and resumed vigourous growth when transferred to normal growth conditions at 20°C (Fig. 3A). The level of *At*TIL protein accumulation in leaves of NA plants is associated with increased FT, indicating that *At*TIL provides protection against damages caused by a sudden drop in temperature (Fig. 3C). This difference in sensitivity to freezing between WT and KO plants is not observed when plants are cold acclimated prior to exposure to freezing temperature (CA -10°C, Fig. 3A). Cold acclimation is therefore sufficient to provide maximal FT, and *At*TIL is unlikely to be a major component of the cold acclimation process.

The protection against freezing damages could be due to a role of lipocalins in scavenging harmful reactive oxygen species generated when plants are exposed to freezing. To test this hypothesis, plants were treated with the oxidant paraquat and necrotic lesions were monitored over 7 days. At the end of this period, KO plants showed more damage than WT plants, whereas OEX plants were more resistant (Fig. 4A and 4B). The protein level is associated with the level of paraquat tolerance, indicating that *At*TIL provides protection against oxidative stress (Fig. 4C).

Expression profiles were determined using the *Arabidopsis* ATH1 Genome Array (Affymetrix) for WT and KO plants grown under NA and CA conditions (7 days at 4°C). No differences in gene expression were observed for all the genes present on the microarray between the CA WT and KO, except for the absence of *AtTIL* transcript in the KO line. However, in the NA KO, a total of 66 genes were differentially regulated by more than 2-fold (51 up and 15 down-regulated) compared to the NA WT (Table S1). Among the 66 genes regulated by the absence of *At*TIL at 20°C, 49 are regulated in a similar manner (up or down) by LT exposure in WT plants while 5 are regulated in an opposite manner. These genes encode known or putative transcription factors or signal transduction proteins, heat shock proteins,

enzymes involved in carbohydrate metabolism or tolerance to oxidative stress, senescence and circadian clock genes. The remaining 12 genes are regulated only in the KO line (* in Table S1; KO-specific genes) and encode genes that are associated with stress responses. This analysis indicates that without exposure to LT, the absence of *AtTIL* mimicks part of the LT response.

AtTIL KO plants show an increased sensitivity to paraquat. This powerful herbicide acts in the chloroplastic electron transport, generating ROS that can diffuse outside the chloroplasts (*7-9*). On the other hand, *AtTIL* OEX plants show improved tolerance to the oxidative stress caused by paraquat. The *At*TIL protein is associated with the plasma membrane and could contribute to scavenging ROS and prevent membrane damage. Another possible mode of action is that *At*TIL could increase oxidative stress tolerance through its role in post-stress recovery. It could play a role in restoring membrane integrity caused by oxidative stress, as suggested for animal lipocalins (*10*).

Several genes related to the disease resistance pathway are overexpressed in the *AtTIL* KO plant. The WRKY54 gene is known for its association with the defense response and involves an increase in salicylic acid (SA), which accumulates during oxidative stress and mediates the induction of defense response genes (*11,12*). Similarly, the SigA binding protein is a homologue of MSK1 which was shown to cause an accumulation of SA (*13*). The At3g55450 protein is most homologous to the BIK1 kinase which is induced by ROS generators such as paraquat and proposed to play a role in regulating optimal SA levels (*14*). Induction of the BIK1 kinase-like gene in *AtTIL* KO plants without exposure to any stress suggests that the ROS level is higher than normal in these plants. The greater susceptibility to paraquat is also indicative that the *AtTIL* KO plants are less capable to fight additional oxidative stress. Furthermore, riboflavin deficiency is associated with oxidative stress (*15*) and is an important component of FAD-requiring enzymes such as glutathione reductase. The overexpression of riboflavin biosynthesis protein in *AtTIL* KO plants suggests

that riboflavin becomes limiting and may be mobilized to fight against oxidative stress.

The microarray analysis suggests that in the KO plant, starch synthesis is reduced while starch catabolism is accelerated to provide additional soluble sugars needed to fight oxidative stress. In addition, the Krebs cycle is inhibited under stress conditions, preventing the efficient utilization of the glycolysis products and thus affecting the respiratory machinery (*16*). To overcome this situation, plants rely on the γ-aminobutyrate (GABA) shunt. Succinic semialdehyde dehydrogenase, the last enzyme of the GABA shunt, is upregulated in *AtTIL* KO plants. The GABA shunt is a bypass of the Krebs cycle which provides succinate and NADH to the respiratory machinery (*17*). In plants, the activity of this pathway is enhanced in response to biotic and abiotic stresses (*18,19*) suggesting a potential role in reducing the impact of oxidative stress in mitochondria (*16*).

Our data demonstrate that the inactivation of *At*TIL affects tolerance of *Arabidopsis* to the oxidative stress caused by freezing, paraquat treatment and light. Overexpression of this gene enhances the tolerance to both freezing stress and paraquat treatment without affecting the plant's growth and development. Results presented in this report and the two accompanying papers by Ganfornina et al. and Muffat et al. suggest that lipocalins may have a universal cellular function. Overexpression of either homologs (*AtTIL* and *ApoD*) could be used to improve the abilitity of organisms to tolerate oxidative stress and delay aging. This function could have important medical and agricultural applications.

Materials and methods

AtTIL lines

Arabidopsis ecotype Columbia (Col-0) was the genetic background of all the lines used. T-DNA knock-out (KO) lines for the *AtTIL* gene were obtained from the Salk Institute Genome Analysis Laboratory (http://signal.salk.edu/). Seeds were sown on agar plates containing 50 μg/ml kanamycin, stratified for 2 days at 4°C and grown at 22°C. Kanamycin-resistant plants were propagated as individual lines on potting medium consisting of two parts *Arabidopsis* growing medium PM-15-13 (LEHLE seeds), one part vermiculite and one part black earth. Based on the segregation analysis of plants grown on antibiotic-containing plates, lines putatively homozygous for the T-DNA insertion were subjected to PCR analysis using T-DNA-specific (LBb1) and *AtTIL*-specific (RP) primers, and the resulting PCR fragments were sequenced to determine the precise insertion sites.

To generate the complementation (Comp) and overexpressing (OEX) lines, the *AtTIL* ORF was first cloned into pRTL2, a vector that contains a double cauliflower mosaic virus (CaMV) 35S promoter and a 35S terminator (*21*). The resulting PRO_{35S}:*AtTIL*:$TERM_{35S}$ cassette was then cloned into the binary plant expression vector pPZP121 (*22*) and electroporated into *Agrobacterium* GV3101. KO and WT *Arabidopsis* plants were transformed with this vector using the floral dipping method (*23*) to generate the Comp and OEX lines, respectively. To generate a negative control, WT plants were transformed with the pPZP121 vector that does not contain the *AtTIL* cassette. Transformed plants of all lines were selected by growth on gentamycin-containing medium. Homozygous plants of the fourth generation after transformation were used for the experiments.

Plant Growth Conditions and Treatments

Surface-sterilized seeds were sown on sterile half strength Murashige and Skoog medium (Sigma-Aldrich). Seeds were stratified for 2 days at 4°C then germinated and grown at 22°C/18°C (day/night) with a 16 hr photoperiod and a photon flux density of 90 μmol m^{-2} sec^{-1}. For soil-based analyses, seeds were sown directly on potting medium, stratified for 3 days at 4°C and grown for 11 days at 20°C under long day conditions (16 hr light/8 hr dark) at a photon flux density of 90 μmol m^{-2} sec^{-1}. Seedlings were then transferred to 3.5 inches or 1.5 inches pots containing potting medium and grown under the same conditions as described above.

For low temperature treament, 3 week-old soil-grown plants were transferred to 4°C for 7 days under the same photoperiod conditions. For paraquat treatment, paraquat (Sigma) was dissolved in water at a concentration of 15 μM and plants were sprayed once with this solution until run off. Control plants were sprayed with distilled water. After spraying, plants were grown for 7 days under normal conditions of temperature and photoperiod.

Hypocotyl Analyses

Seeds were sown on sterile half strength Murashige and Skoog medium (Sigma-Aldrich) and stratified in the dark at 4°C for 6 days. They were then exposed to white light (90 mol m^{2} s^{1}) at 20°C for 60 min and returned to darkness for 23 hr before being exposed to varying fluence rates at 20°C for 7 days (see Fig 3 legend). Hypocotyl length was measured using the Image J 1.36b software (mean ± SE; $n = 40$).

Determination of Freezing Tolerance

A Caltec Scientific Ltd. Model 8-792 Large Capacity Temperature Stress Chamber was used to perform the freezing tolerance tests. This instrument consists of four major component systems: a Sanyo Model MDF-792 24.75 ft^3 capacity ultra-low temperature chest freezer, a custom designed stainless steel plenum box with its integral blower and heater (provides air circulation and heating) and an Omega Engineering Inc. Model CN3002 programmable profile controller (monitors the test-chamber air temperature and controls the heating element). The controlled action of the heater combines with the constant cooling of the freezer to achieve the desired temperature at any given time.

Non-acclimated (NA) and cold-acclimated (CA) soil-grown plants (3 weeks-old) were subjected to the following freezing regime. The temperature was lowered gradually (2°C per hr) to -6°C and maintained at this temperature for 6 hr. The temperature was then gradually increased (2°C per hr) to 4°C. To determine temperature variability in the freezer, temperatures were monitored by 6 independent thermocoupled T probes distributed in the freezer and connected to an Agilent 3497-0A data acquisition/switch unit. Freezing regimes that showed more than 0.5 °C discrepancies between the different probes were rejected. To minimize light stress effects after the freezing treatment, plants were thawed at 4°C for 24 hr in the dark and away from direct light in the growth chamber (20°C) for an additional 24 hr before returning to normal light conditions. Pictures were taken 3 weeks after the freezing treatment. Eighteen plants were frozen per line per assay, and the experiment was repeated 3 times with independent biological replicates.

Protein Isolation and Immmunoblot Analysis

The antibody raised against the wheat *Ta*TIL lipocalin protein (*4*) does not cross react with the *Arabidopsis At*TIL lipocalin, therefore it was necessary to raise a

specific antibody against the latter. The *AtTIL* cDNA was cloned in the pTrc-HIS vector (Invitrogen) and electroporated into the *E. coli* strain DH5α. The recombinant HIS::*At*TIL protein was produced by induction with 1 mM IPTG for 3 hr. The cells were collected, broken by lysozyme treatment and sonication, and HIS-tagged proteins were purified by immobilized metal affinity chromatography on Ni-NTA agarose (Novagen). The purified proteins were more than 90% pure and used to immunize a rabbit to obtain polyclonal antibodies.

Aerial parts of the *Arabidopsis* plants were cut and immediately frozen in liquid nitrogen. One hundred milligrams of leaf material was processed as one sample. Proteins were isolated using Tri Reagent according to the manufacturer's instructions (Sigma). Samples were separated on 12% SDS-PAGE gels and the rabbit anti-*At*TIL antibody (1:10,000) was used for the immunoblot analysis. Detection was performed with a peroxidase-coupled anti-rabbit IgG secondary antibody (1:25,000) and the Western Lightning Chemiluminescence Reagent Plus (Perkin-Elmer).

PCR and Southern Blot Analyses

Genomic DNA was extracted from flower buds of the different *Arabidopsis* lines. PCR analysis was performed according to the SIGnAL protocol using the recommended primers designed with the SIGnAL iSect Primer Design sofware : LBb1 (5' GCGTGGACCGCTTGCTGCAACT 3'), LP (5' CTGGATCCAGAGATGAAGTCG 3') and RP (5' AAGACGTGTATGGTACCGTCG 3'). For Southern analysis, the DNA (5 µg) was digested by *Pci*I (New England Biolabs) and fractionated by electrophoresis on a 0.7% agarose gel. After electrophoresis, the gel was transferred to a positively charged nylon membrane and hybridized with ^{32}P-labeled probes corresponding to the T-DNA of pROK2 (vector used to generate the SALK lines) or the T-DNA of the *35S:AtTIL* vector. The following primer pairs were used for amplification of the

fragments used as probes: 5'- CAGCAAAATCACCAGTAGCACCATTACCAT-3' and 5'- GCGATAGAAAACAAAATATAGCGCGCAAAC-3' for the pROK2 T-DNA, and 5'-ACGAAACGTGGAGCAACGGGAAGAG-3' and 5'-TGCACATACAAATGGACGAACGGATAAACC -3' for the *AtTIL* T-DNA. All filters were washed at high stringency (0.1X SSC, 0.1% SDS), exposed to K screens (Kodak) and analyzed on a Molecular Imager FX (Bio-Rad).

Transcriptome Analysis

Non-acclimated and cold-acclimated WT and KO plants were used for microarray analysis. For each sample, three independent biological replicates of 25 plants were harvested for RNA isolation and samples were sent to the Microarray services platform of the McGill University and Génome Québec Innovation Centre (Montreal QC Canada). RNA quality was assessed on the Agilent Bioanalyzer, and cRNA was synthesized and hybridized on GeneChip *Arabidopsis* ATH1 Genome Arrays from Affymetrix (Santa Clara, CA). All Affymetrix protocols recommended for cDNA synthesis, array hybridization, and chip scanning were followed. Data accumulation and analysis were performed using the Robust Multi-array Average analysis (RMA version 0.2) [http://www.stat.berkeley.edu/~bolstad/RMAExpress/RMAExpress.html], and Excel (Microsoft, Redmond, USA). The RMA software provides background-adjusted, normalized, and log-transformed perfect match values and is recognized as one of the most robust tools for the analysis of microarray data (*24*). An analysis of variance was performed using GraphPad InStat 3 to select genes that are significantly differentially expressed by at least two-fold under the conditions specified for each analysis (4°C vs 22°C and KO vs WT). Functional classification of the genes was done according to their annotation in the TAIR database (http://www.arabidopsis.org/tools/bulk/go/index.jsp) and by literature search performed using the gene name and annotation.

Acknowledgments

We thank Guillaume Brault, Greg Cormack, Nadia Toudji, and Karine Tremblay for technical assistance. This work was supported by grants from Genome Canada, Génome Québec and the Natural Sciences and Engineering Research Council of Canada.

References

1. D. Sanchez, M. D. Ganfornina, G. Gutierrez, A. C. Gauthier-Jauneau, J. L. Risler *et al.*, in *Lipocalins*, B. Åkerström and L. Lögdberg, Eds. (Landes Bioscience, Georgetown TX, 2006).

2. D. Sanchez, B. Lopez-Arias, L. Torroja, I. Canal, X. Wang *et al.*, *Curr. Biol.* **16**, 680-686 (2006).

3. D. W. Walker, J. Muffat, C. Rundel, S. Benzer, *Curr. Biol.* **16**, 674-679 (2006).

4. J. B. Charron, F. Ouellet, M. Pelletier, J. Danyluk, C. Chauve *et al.*, *Plant Physiol.* **139**, 2017-2028 (2005).

5. J. B. F. Charron and F. Sarhan, in *Lipocalins*, B. Åkerström and L. Lögdberg, Eds. (Landes Bioscience, Georgetown TX, 2006) , chap. 5.

6. J. M. Alonso, A. N. Stepanova, T. J. Leisse, C. J. Kim, H. Chen *et al.*, *Science* **301**, 653-657 (2003).

7. P. Broadbent, G. P. Creissen, B. Kular, A. R. Wellburn, P. M. Mullineaux, *Plant J.* **8**, 247-255 (1995).

8. R. Perl-Treves, E. Galun, *Plant Mol. Biol.* **17**, 745-760 (1991).

9. E. W. Tsang, C. Bowler, D. Herouart, W. Van Camp, R. Villarroel *et al.*, *Plant Cell* **3**, 783-792 (1991).

10. R. E. Bishop, *Biochim. Biophys. Acta* **1482**, 73-83 (2000).

11. N. Journot-Catalino, I. E. Somssich, D. Roby, T. Kroj, *Plant Cell* **18**, 3289-3302 (2006).

12. D. Wang, N. Amornsiripanitch, X. Dong, *PLoS. Pathog.* **2**, e123 (2006).

13. E. Andreasson, T. Jenkins, P. Brodersen, S. Thorgrimsen, N. H. Petersen *et al.*, *EMBO J.* **24**, 2579-2589 (2005).

14. P. Veronese, H. Nakagami, B. Bluhm, S. Abuqamar, X. Chen *et al.*, *Plant Cell* **18**, 257-273 (2006).

15. R. S. Rivlin, in *Present knowledge in nutrition*, E. E. Ziegler and L. J. Filer, Eds. (ILSI Press, Washington DC, 1996).

16. N. Bouché, A. Fait, D. Bouchez, S. G. Moller, H. Fromm, *Proc. Natl. Acad. Sci. USA* **100**, 6843-6848 (2003).

17. B. J. Shelp, A. W. Bown, M. D. McLean, *Trends Plant Sci.* **4**, 446-452 (1999).

18. A. W. Bown, B. J. Shelp, *Plant Physiol.* **115**, 1-5 (1997).

19. W. A. Snedden and H. Fromm, in *Plant responses to environmental stresses: from phytohormones to genome reorganization*, H. R. Lerner, Ed. (Marcel Dekker, New-York NY, 1999).

20. D. C. Boyes, A. M. Zayed, R. Ascenzi, A. J. McCaskill, N. E. Hoffman *et al.*, *Plant Cell* **13**, 1499-1510 (2001).

21. M. A. Restrepo, D. D. Freed, J. C. Carrington, *Plant Cell* **2**, 987-998 (1990).

22. P. Hajdukiewicz, Z. Svab, P. Maliga, *Plant Mol. Biol.* **25**, 989-994 (1994).

23. S. J. Clough, A. F. Bent, *Plant J.* **16**, 735-743 (1998).

24. R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis *et al.*, *Biostatistics* **4**, 249-264 (2003).

Figure 1. Modulation of *At*TIL protein level affects development.

(A) Genomic organization of *Arabidopsis* SALK lines carrying T-DNA insertions in the *AtTIL* gene (At5g58070). Boxes 1 and 2, exons. The primers used for PCR analysis of the genomic DNA are indicated (see Fig. S1). **(B)** *At*TIL protein levels in leaf extracts. Top panel: A rabbit anti-*At*TIL antibody was used for immunoblot analysis. Bottom panel: Coomassie Brilliant Blue-stained gel. **(C)** Plants were grown under normal conditions of temperature and photoperiod. Overhead (21 d) or lateral (28, 35, 45 d) views are shown. **(D)** Developmental growth stages expressed according to Boyes *et al.* (2001;(*20*)). 1.10: 10 rosette leaves; 5.10: appearance of the first flower buds; 6.00: opening of the first flower; 6.90: completion of flowering. At least 30 plants per line per assay were monitored, and the experiment was repeated 3 times. WT, wild type Col-0 plants; SALK_XXXXXX, *AtTIL* T-DNA insertion lines from the SALK collection; Comp, SALK_136775 KO plant complemented by overexpression of *AtTIL*; OEX, an *AtTIL* overexpressing line; Vector, Col-0 transformed with a binary vector that does not carry the *AtTIL* cDNA (negative control).

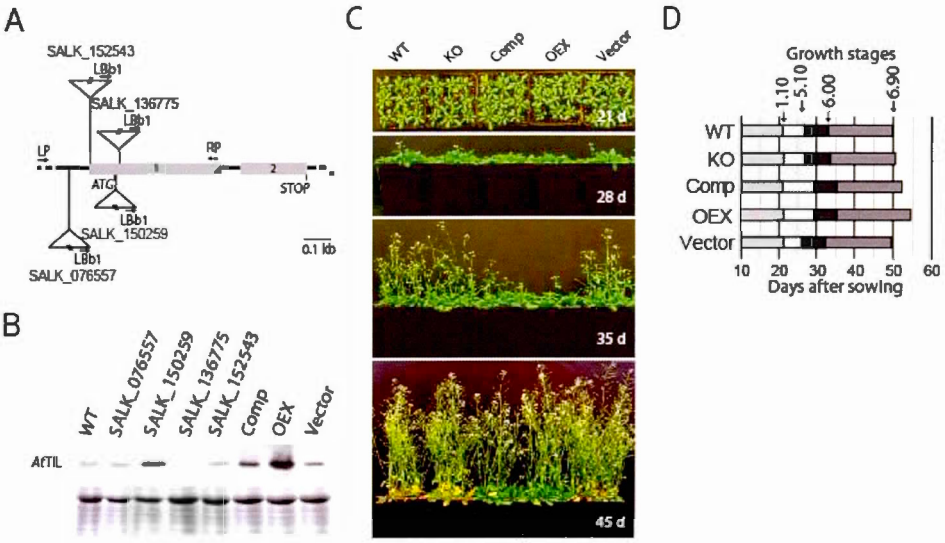

A
SALK_152543
LBb1
SALK_136775
LBb1
LP
RP
ATG
STOP
LBb1
SALK_150259
LBb1
SALK_076557
0.1 kb
B
WT
SALK_076557
SALK_150259
SALK_136775
SALK_152543
Comp
OEX
Vector
AtTIL
C
WT
KO
Comp
OEX
Vector
21 d
28 d
35 d
45 d
D
Growth stages
1.10
5.10
6.00
6.90
WT
KO
Comp
OEX
Vector
10
20
30
40
50
60
Days after sowing

Figure 2. *AtTIL* knock-out plants show reduced hypocotyl elongation.

(A) Representative seedlings grown on plates under dark conditions. **(B)** Fluence response curves of seedlings grown under continuous light. Seeds were stratified at 4°C in the dark for 6 days on agar plates, then transferred at 22°C for 1 hr in the light and 23 hr in the dark. Plants were then exposed 7 days in the dark or under continuous light at various light intensities. Images were captured on a dissecting microscope and hypocotyl length was measured with the Image J imaging software. At least 40 plants per line per assay were monitored, and the experiment was repeated 2 times. **(C)** Representative seedlings grown on plates under continuous light (LL; upper panel) and light-dark (LD; lower panel) conditions at 100 μmol $m^{-2}s^{-1}$.

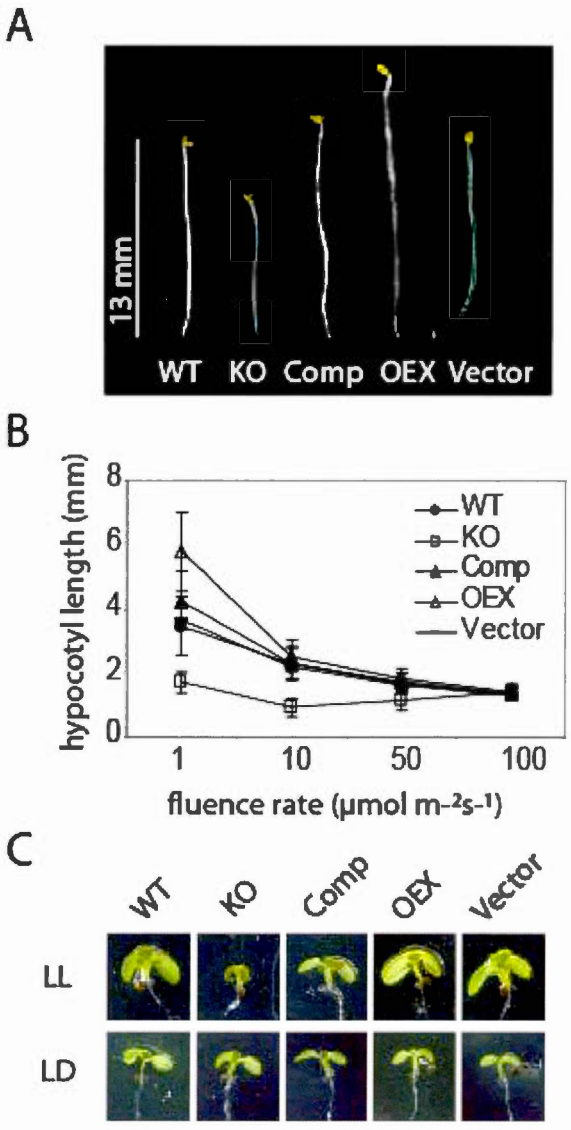

A
13 mm
WT
KO
Comp
OEX
Vector
B
hypocotyl length (mm)
8
6
4
2
0
1
10
50
100
fluence rate (µmol m-2s-1)
WT
KO
Comp
OEX
Vector
C
WT
KO
Comp
OEX
Vector
LL
LD

Figure 3. *At*TIL enhances tolerance to freezing stress.

(A) Plants were grown for 3 weeks at 22°C (NA) or grown at 22°C then transferred at 4°C for 7 days (CA7), and pictures were captured. The same plants were subjected to a freeze test performed at -6°C or -10°C for the NA and CA plants, respectively, and pictures were captured after a recovery period of 3 weeks. **(B)** Survival rate after freezing of NA plants to -6°C, expressed as a ratio of surviving to total plants. Statistical analysis was performed by one-way ANOVA, and the asterisks (*) indicate differences that are significant at the $P<0.001$ level. **(C)** *At*TIL protein accumulation in leaves of NA and CA plants. Proteins were extracted and analyzed by immunoblotting using the anti-*At*TIL antibody. The Coomassie Brilliant Blue-stained gels are shown as loading controls. Results are representative of at least three independent assays involving 18 plants per line per assay.

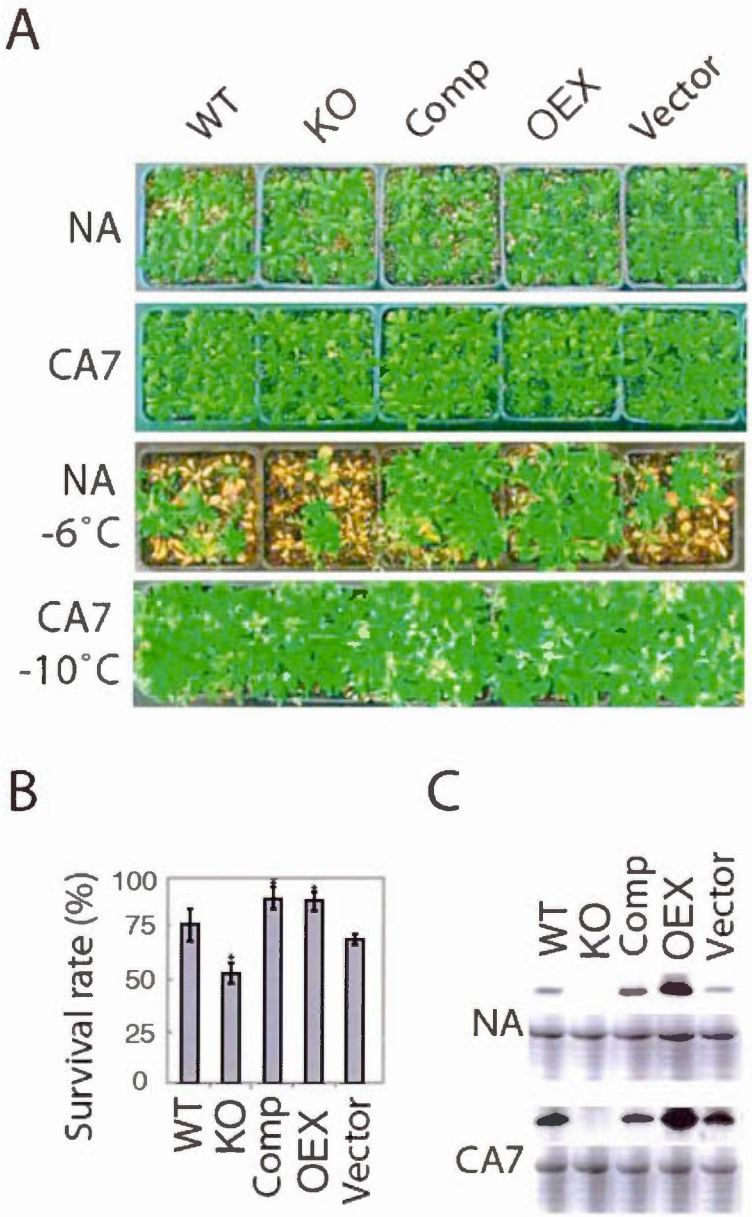

A
WT
KO
Comp
OEX
Vector
NA
CA7
NA
-6˚C
CA7
-10˚C
B
Survival rate (%)
100
75
50
25
0
WT
KO
Comp
OEX
Vector
C
WT
KO
Comp
OEX
Vector
NA
CA7

Figure 4. The level of *At*TIL accumulation influences oxidative stress tolerance of *Arabidopsis*.

(A) Plants were grown under normal conditions for 3 weeks and then sprayed until run off with either water or a 15 μM paraquat solution. Pictures were captured at the indicated time after treatment. Only the paraquat-treated plants are shown since no effect was observed for plants sprayed with water. **(B)** A close-up is shown to better show the necrotic lesions. **(C)** *At*TIL accumulation following paraquat treatment. Protein extracts were prepared from at least three *Arabidopsis* plants and subjected to immunoblot analysis using the anti-*At*TIL antibody. The Coomassie Brilliant Blue-stained gels are shown as loading controls (bottom panels).

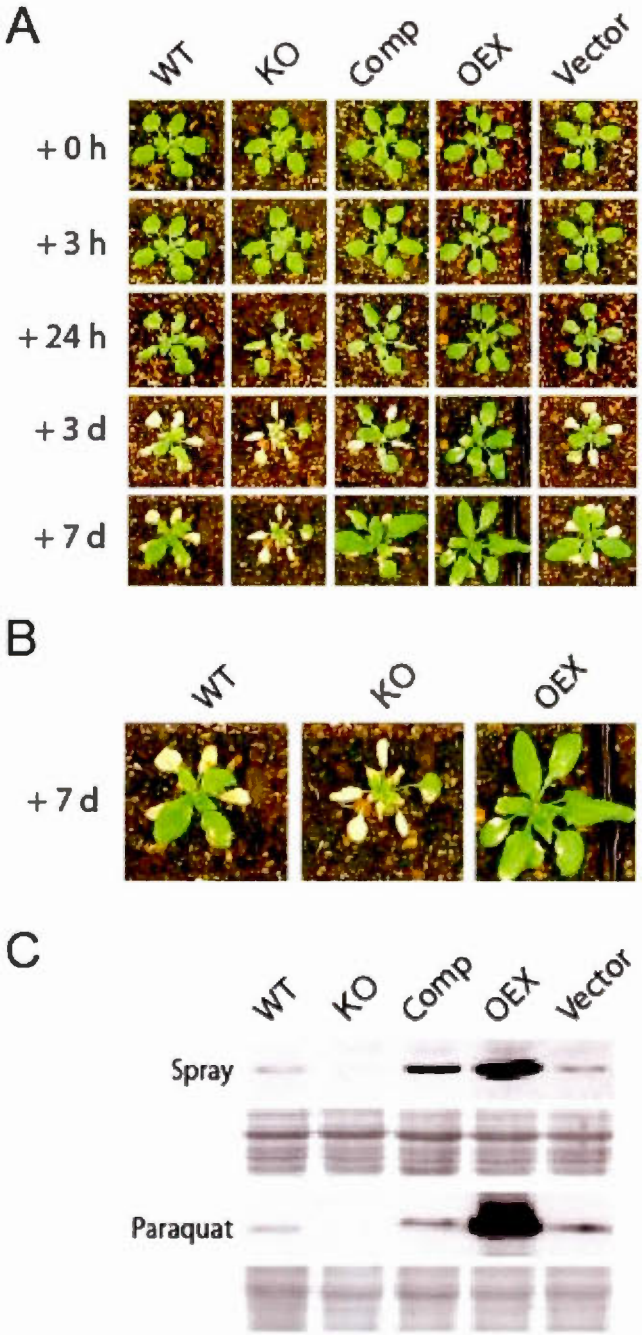

A
WT
KO
Comp
OEX
Vector
+0 h
+3 h
+24 h
+3 d
+7 d
B
WT
KO
OEX
+7 d
C
WT
KO
Comp
OEX
Vector
Spray
Paraquat

Table S1. Genes showing at least two-fold differential expression (induction / repression) in *Arabidopsis AtTIL* knock-out plants.

			Differential expression (absolute values)					
Putative Function and Reference	Locus ID	ProbesetIDs	KO vs WT	KO 4°C vs KO	WT 4°C vs WT	KO 4°C vs WT 4°C	Regul-ation	
Upregulated Genes								
Transcription Factors:								
pseudo-response regulator 5 (APRR5)	At5g24470	249741_at	5.3	6.3	33.8	-1.0		‡
zinc-binding family protein (DUF597)	At1g76590	259977_at	2.9	2.1	5.7	1.1	#	
WRKY family transcription factor	At2g40750	257382_at	2.8	2.2	4.8	1.3	#	
DRE-binding protein (DREB1A) / CRT/DRE-binding factor 3 (CBF3)	At4g25480	254066_at	2.3	4.0	12.9	-1.4		
ABA-responsive element-binding protein (ABRE)	At1g49720	261613_at	2.2	2.4	4.9	1.1		
sigA-binding protein	At3g56710	246293_at	2.2	1.8	3.6	1.1	#	
heat shock factor protein, putative (HSF5) (HSTF5)	At1g67970	259992_at	2.1	2.2	4.0	1.2	#	
gigantea protein (GI)	At1g22770	264211_at	2.0	3.6	6.9	1.1		‡
zinc finger (MYND type) family protein	At5g50450	248502_at	2.1	2.2	4.4	1.1		
zinc finger (C3HC4-type RING finger) family protein	At1g49230	260753_at	2.1	1.4	2.9	1.0	#	
CONSTANS-LIKE 13 zinc finger (B-box type) family protein	At2g47890	266514_at	2.1	3.5	6.5	1.1		
ABI3-interacting protein 1 (AIP1) (TOC1)	At5g61360	247525_at	2.1	3.1	5.4	1.2		‡
Signal Transduction:								
invertase/pectin methylesterase inhibitor family protein	At5g62360	247478_at	5.2	1.3	5.8	1.1	#	†
cold circadian rhythm-RNA binding 2; CCR2	At2g21660	263548_at	4.4	3.2	13.4	1.1		
acid phosphatase class B family protein	At2g39920	267361_at	4.0	-1.1	3.4	1.0	#	
auxin-regulated protein kinase, putative	At2g33830	267461_at	3.2	-1.6	1.7	1.1	*	
peptidyl-prolyl cis-trans isomerase / cyclophilin (CYP2) / rotamase	At2g21130	264019_at	2.3	2.9	6.4	1.1		
ankyrin repeat family protein (ACD6)	At4g14400	245265_at	2.2	2.0	4.5	-1.0		
protein kinase, putative	At3g55450	251789_at	2.2	2.1	4.2	1.1	#	
phosphorylase family protein	At4g24340	254163_s_at	2.1	-3.3	-1.3	-1.2	*	
Stress Reponse, Cell Rescue and Defense:								
DNAJ heat shock N-terminal domain-containing protein	At5g23240	249850_at	6.5	2.1	11.6	1.2	#	
DNAJ heat shock N-terminal domain-containing protein	At1g56300	256221_at	4.4	-1.2	3.1	1.2	#	
hydrophobic protein, low temperature and salt responsive protein	At4g30650	253627_at	2.3	7.8	17.5	1.0		
senescence-associated protein-related (SAG102)	At1g53885	262226_at	2.3	-4.7	-2.1	1.0		
hydrophobic protein, low temperature and salt responsive protein	At4g30660	253581_at	2.1	-2.7	-1.1	-1.2	*	
universal stress protein (USP) family protein	At3g62550	251221_at	2.0	-2.4	-1.3	1.1	*	
heavy-metal-associated domain-containing protein	At1g51090	245749_at	2.0	5.7	11.1	1.0		
pathogenesis-related thaumatin family protein (calcium storage)	At1g20030	261248_at	2.0	-1.7	1.0	1.0	*	
cold-regulated protein (cor15b)	At2g42530	263495_at	2.0	24.0	47.7	1.0		
pathogen and circadian controlled 1 (PCC1)	At3g22231	256766_at	2.0	1.6	2.9	1.0	#	
Metabolism:								
riboflavin biosynthesis protein, putative	At2g22450	264045_at	2.8	1.2	3.2	-1.0	#	††
pyruvate decarboxylase, putative	At5g54960	248138_at	2.6	1.6	3.7	1.1	#	††
preprotein and amino acid transporter	At4g26670	253981_at	2.4	1.8	4.0	1.1	#	
succinate-semialdehyde dehydrogenase (SSADH1)	At1g79440	262892_at	2.2	2.6	5.6	1.0		
aldose 1-epimerase family protein	At3g47800	252387_at	2.2	1.6	3.7	-1.1	#	†
nucellin protein, putative	At4g33490	253331_at	2.1	-1.1	1.7	1.1	*	
auxin-responsive family protein	At5g35735	249719_at	2.1	1.2	2.4	1.0	#	
acyl CoA:diacylglycerol acyltransferase (DGAT)	At2g19450	267280_at	2.1	3.0	5.9	1.1		
isoamylase, putative / starch debranching enzyme, putative	At4g09020	255070_at	2.1	3.0	6.0	1.0		†
2-oxoacid-dependent oxidase, putative (DIN11)	At3g49620	252265_at	2.1	-2.8	-1.2	-1.1	*	
OEP16-like protein	At2g28900	266225_at	2.1	2.7	5.8	-1.0		
starch phosphorylase, putative	At3g46970	252468_at	2.1	1.6	2.9	1.1	#	†
glycoside hydrolase starch-binding domain-containing protein	At5g26570	246829_at	2.0	2.5	4.9	1.0		†
amino acid permease family protein or GABA permease	At2g01170	265790_at	2.0	1.1	2.1	1.0	#	
Expressed Protein								
expressed protein	At4g16146	245319_at	4.0	4.2	14.7	1.2		
expressed protein	At1g53035	261318_at	2.6	5.1	12.0	1.1		
expressed protein	At4g04330	255331_at	2.6	-1.6	1.5	1.1	*	
expressed protein	At2g14560	265837_at	2.5	2.9	6.1	1.2		
expressed protein	At2g15890	265478_at	2.4	-3.0	-1.4	1.1	*	
expressed protein	At1g70420	264314_at	2.1	-1.8	1.6	-1.3	*	
expressed protein	At1g14870	262832_at	2.1	4.0	5.9	1.4		
Downregulated Genes								
Transcription Factors:								
late elongated hypocotyl (LHY)	At1g01060	261569_at	-3.2	6.2	2.1	-1.1		‡
CONSTANS-LIKE 2 (COL2)	At3g02380	258497_at	-2.4	-2.0	-3.2	-1.5	#	‡
Signal Transduction:								
cysteine proteinase, putative	At2g27420	265665_at	-2.3	2.3	1.0	-1.0	*	
calcium-binding RD20 protein (RD20)	At2g33380	255795_at	-2.1	-2.1	-3.7	-1.2	#	
Stress response, Cell Rescue and Defense:								
At TIL, lipocalin	At5g58070	247851_at	-3.6	-1.7	1.5	-8.8	*	
putative membrane related protein CP5	At1g55960	260603_at	-2.2	-2.0	-4.3	-1.0	#	
phosphate-responsive protein, putative (EXO)	At4g08950	255064_at	-2.0	-2.0	-5.3	1.3		

Metabolism:							
dehydrodolichyl diphosphate synthase, putative	At5g58770	247780_at	-3.1	-2.8	-7.6	-1.1	
trehalose-6-phosphate phosphatase (TPPA)	At5g51460	248404_at	-2.8	-2.6	-6.7	-1.1	
starch synthase, putative	At1g32900	261191_at	-2.5	6.4	2.7	-1.1	
carbonic anhydrase family protein	At3g52720	252011_at	-2.2	-3.3	-6.9	-1.1	
glutathione S-trensferase, putative (ERD9)	At1g10370	264436_at	-2.2	4.5	2.1	-1.0	
adenosylmethionine decarboxylase family protein	At5g15950	246490_at	-2.0	6.6	3.5	-1.1	
Expressed Protein							
expressed protein	At3g28270	256603_at	-2.3	-1.9	-4.0	-1.1	#
expressed protein	At1g68440	259856_at	-2.1	-2.5	-4.6	-1.1	

* 12 genes regulated (induced or repressed) only in the KO line (not regulated by exposure of the WT to LT)

\# 22 genes for which the regulation in KO is at least 50% of the regulation seen in WT exposed at 4°C

† genes associated with carbohydrate metabolism

†† genes associated with tolerance to oxidative stress

‡ key regulators of the circadian clock

A typical induction of known cold-regulated genes (*COR15a*, *COR15b*, *COR47*, *CBF3*, *CBF2*) was observed in both the cold-acclimated WT and KO lines (result not shown). Of the 1587 genes regulated more than two-fold after 14 days of cold treatment in the study by Hannah *et al.* (PLoS Genet. **1**, e26 2005), 675 genes (43%) are also regulated after 7 days of cold treatment in our study.

Figure S1. Genomic organization of *Arabidopsis* SALK lines carrying T-DNA insertions in the *AtTIL* gene (At5g58070).

(A) PCR analysis of genomic DNA extracted from the different lines, performed with primers indicated in Fig. 1. **(B)** Southern blot analysis of genomic DNA extracted from the different lines. Left panel, probe detecting the pROK2 T-DNA insertion present in the SALK lines; right panel, probe detecting both a portion of the *AtTIL* transgene T-DNA (including the right border) and the *AtTIL* endogenous gene. WT, wild type Col-0 plants; SALK_XXXXXX, *AtTIL* T-DNA insertion lines from the SALK collection; Comp, SALK_136775 KO plant complemented by overexpression of *AtTIL*; OEX, an *AtTIL* overexpressing line; Vector, Col-0 transformed with a binary vector that does not carry the *AtTIL* cDNA (negative control).

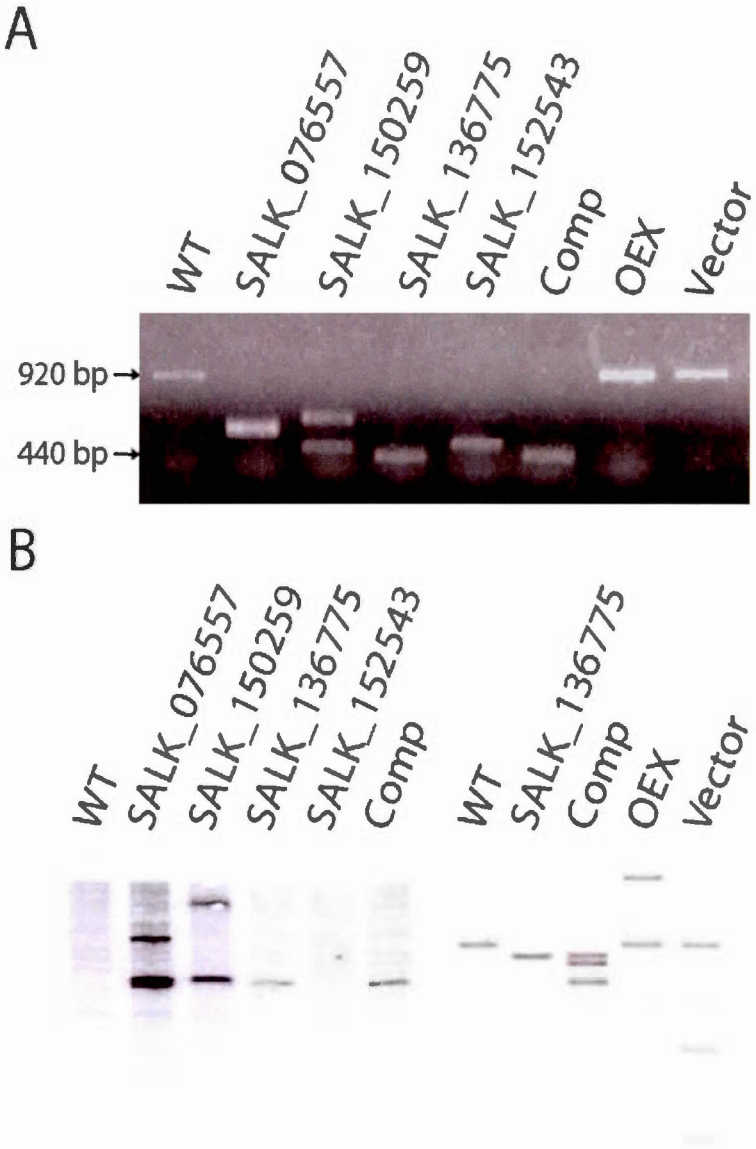

A
WT
SALK_076557
SALK_150259
SALK_136775
SALK_152543
Comp
OEX
Vector
920 bp
440 bp
B
WT
SALK_076557
SALK_150259
SALK_136775
SALK_152543
Comp
WT
SALK_136775
Comp
OEX
Vector

Figure S2. Accumulation of the *At*TIL protein immunodetected with the anti-*At*TIL antibody.

Bottom panel: Coomassie Brilliant Blue-stained gel.

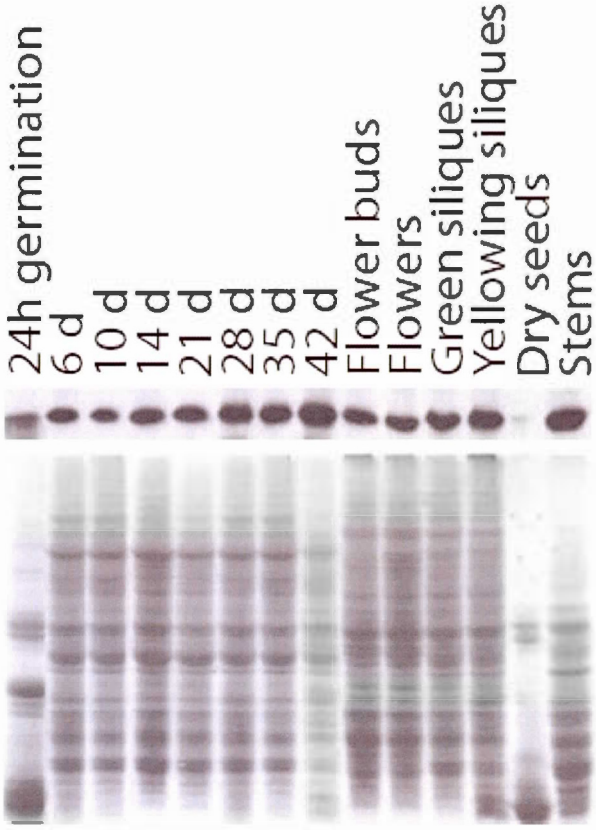
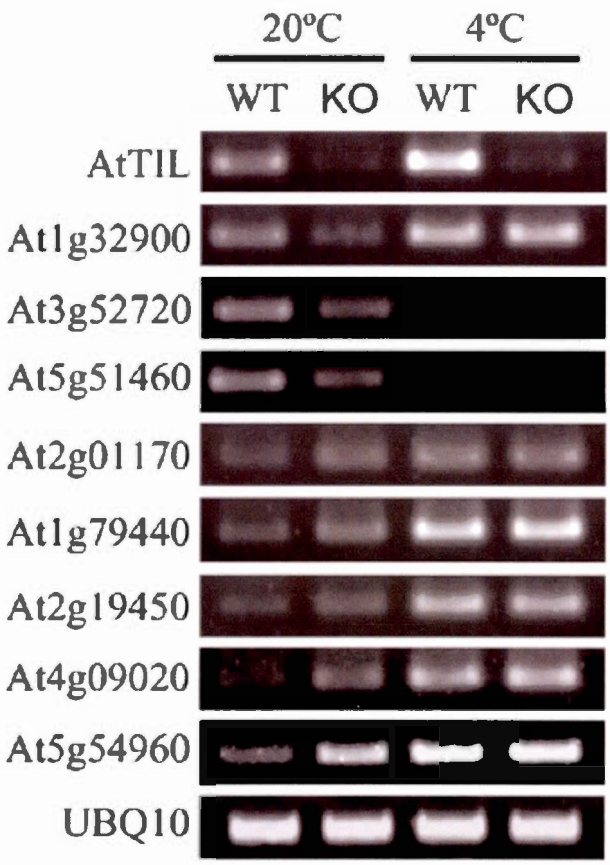

24h germination
6 d
10 d
14 d
21 d
28 d
35 d
42 d
Flower buds
Flowers
Green siliques
Yellowing siliques
Dry seeds
Stems

Figure S3. Validation of microarray data using reverse transcriptase-PCR (RT-PCR). Transcript levels of randomly selected genes were quantitated in the control and cold-treated WT and *AtTIL* KO plants by RT-PCR. UBQ10 was used as a constitutively expressed control transcript. Reactions were performed using three independent biological replicates. Only one replicate is presented. At5g54960, pyruvate decarboxylase; At1g79440, succinate-semialdehyde dehydrogenase; At4g09020, isoamylase; At2g19450, diacylglycerol O-acyltransferase; At2g01170, amino acid permease; At3g52720, carbonic anhydrase; At1g32900, starch synthase; At5g51460, trehalose-6-phosphate phosphatase.


20°C
4°C
WT
KO
WT
KO
AtTIL
At1g32900
At3g52720
At5g51460
At2g01170
At1g79440
At2g19450
At4g09020
At5g54960
UBQ10

CHAPITRE V

Plant lipocalins

Jean-Benoit F. Charron, Fathey Sarhan

Dans **Lipocalins.** ***Éditeurs*****: B. Åkerström, N. Borregaard, D.R. Flower, et J.P. Salier. Landes Bioscience. Georgetown, TX. 2006, pp. 41-48.**

Résumé

Un grand nombre de lipocalines sont présentes chez les animaux, les insectes et les bactéries. Toutefois, les lipocalines de plantes sont encore méconnues. Les premières protéines de type "lipocalin-like" à avoir été découvertes chez les plantes sont deux enzymes du cycle des xanthophylles, la violaxanthine époxidase et la zéaxanthine dé-époxidase. Cependant, l'architecture distinctive de ces protéines a soulevé des doutes quant à leur appartenance à la famille des lipocalines. Récemment, nous avons rapporté l'identification et le clonage des premières vraies lipocalines de plantes chez le blé et *Arabidopsis*. Les protéines ont été nommées "temperature-induced lipocalins" et possèdent les trois régions structurales conservées qui caractérisent les lipocalines. Les analyses de séquences ont révélé que les lipocalines de plantes possèdent une homologie de séquence avec trois lipocalines reliées au point de vue évolutif : l'Apolipoprotéine D de mammifère, la lipocaline Blc bactérienne et la lipocaline d'insecte Lazarillo. L'utilisation d'une approche intégrée de compilation de données, profilage d'expression, localisation cellulaire, analyses phylogénétiques, et prédictions bioinformatiques a démontré que les plantes possèdent deux autres lipocalines : la "temperature-induced lipocalin-2" et la "chloroplastic lipocalin". Ces analyses suggèrent aussi que les lipocalines de plantes sont associées aux stress environnementaux.

Mots clés : apolipoprotéine D; *Arabidopsis thaliana*; *Triticum aestivum* L.; lipocaline; stress abiotiques; chloroplaste; cycle des xanthophylles; membrane plamique

Summary

Lipocalins are widely distributed in animals, insect and bacteria but very little is known about plant lipocalins. The first lipocalin-like proteins reported in plants were the two key enzymes of the xanthophyll cycle, the violaxanthin de-epoxidases and the zeaxanthin epoxidases. However, the peculiar architecture of these proteins raised doubt as of their true belonging to the lipocalin family. We recently reported the identification and cloning of the first true plant lipocalins from wheat and *Arabidopsis*. The encoded proteins were named temperature-induced lipocalins and possess the three structurally-conserved regions that characterize lipocalins. Sequence analyses revealed that these plant lipocalins share significant homology with three evolutionarily-related lipocalins, the mammalian apolipoprotein D, the bacterial lipocalin Blc and the insect Lazarillo protein. Data mining of genomic databases and bioinformatic predictions revealed that plants possess two other lipocalin members: temperature-induced lipocalin-2 and chloroplastic lipocalin. Expression and regulation studies suggest that the plant lipocalins are associated with environmental stresses.

Introduction

Lipocalins are an ancient and functionally diverse family of mostly extracellular proteins.[1] This family has been studied in details in bacteria, invertebrates and vertebrates, and these studies have been summarized in several excellent reviews.[2-5] However, very little is known about plant lipocalins.[6-7] The rapidly expanding area of functional, structural and comparative genomics provides opportunities for the identification of lipocalin homologs in plants. Using an integrated approach of data mining of EST databases, bioinformatics predictions, phylogenetic studies, and structural, cellular localization and expression profiling analyses, we identified novel plant lipocalins Here we describe the molecular characterization and evolution of plant lipocalins and discuss their putative function during plant development under environmental stresses.

Temperature-Induced Lipocalins

The first true plant lipocalins were recently identified from wheat and *Arabidopsis thaliana*.[7] A full length clone was first isolated from a cDNA library prepared from cold-acclimated wheat tissues and named *TaTIL* for *Triticum aestivum* temperature-induced lipocalin. This gene has since been renamed *TaTIL-1*. The open reading frame encodes a protein of 190 amino acids (aa) with a calculated molecular mass of 22 kDa and a theoretical pI of 5.5 (Table 1). A search in the GenBank ESTs database revealed homology (74% identity, 83% similarity) with a predicted putative protein from *Arabidopsis thaliana* that we named *At*TIL for *Arabidopsis thaliana* temperature-induced lipocalin. Sequence analysis of this *Arabidopsis* clone revealed that the cDNA encodes a 186 aa protein. The SCR1 region is located from aa 15 to 31 (GLDVARYMGRWYEIASF) in *Ta*TIL-1 and from aa 12 to 28 (GLNVERYMGRWYEIASF) in *At*TIL, and possesses the two conserved amino acids G and W (Table 1).[8-9] The SCR2 of *Ta*TIL-1 is found in the C-terminal portion of the protein from aa 105 to 119 (YWVLYVDDDYQYALV) while in *At*TIL it is found from aa 101 to 115 (YWVLYIDPDYQHALI). The SCR2 of animal and bacterial lipocalins generally contains a TDY triplet.[8-9] However, in *Ta*TIL-1 and *At*TIL, only the central D is present (Table 1). SCR3 is found in the C-terminal portion of both proteins, from aa 129 to 144 (ILCRKTHIEEEVNQL) in *Ta*TIL-1 and from aa 125 to 140 in *At*TIL (ILSRTAQMEEETYKQL). The conserved R residue that characterizes this fingerprint is present in both sequences (Table 1).[8-9] Further sequence analysis of *Ta*TIL-1 and *At*TIL indicated the presence of a putative N-glycosylation site (Table 1). Putative C-terminal cleavage sites are predicted by several targeting peptide prediction programs (DGPI, PSORT, and SignalP) to be at aa 172 in *Ta*TIL-1 and at aa 168 in *At*TIL.[10-11] Considering this putative cleavage site, the calculated molecular mass of the mature proteins in wheat and *Arabidopsis* is 20 kDa with a pI of 5.2 (Table 1).

The homology search revealed that *Ta*TIL-1 (accession no. AY077702) and its ortholog from *Arabidopsis* (accession no. AY062789) share significant similarity with three evolutionarily related lipocalins: the human apolipoprotein D (ApoD) precursor (accession no. P05090), the *Escherichia coli* outer membrane lipoprotein Blc precursor (accession no. P39281), and the American grasshopper Lazarillo precursor (accession no. P49291). These proteins respectively share 29%, 31%, and 23% identity, and 46%, 54% and 40% similarity with *Ta*TIL-1. Among all lipocalins, Blc, ApoD, and Lazarillo are the only ones known to be anchored to biological membranes.[3] The good similarity between these proteins and the plant TILs suggests that *Ta*TIL-1 and *At*TIL are also membrane-associated proteins. The sequence analysis also revealed that, like the *E. coli* Blc, *Ta*TIL-1 and *At*TIL differ from most lipocalins by the absence of intramolecular disulfide bonds. However, they are potentially N-glycosylated like human ApoD and Lazarillo. When the three SCRs of these five proteins are aligned, the start codons from *Ta*TIL-1 and *At*TIL are positioned at the cleavage sites of the N-terminal signal peptides of the three other proteins. This alignment suggests that *Ta*TIL-1 and *At*TIL do not possess an N-terminal signal peptide as is the case in Blc, ApoD and Lazarillo. The N-terminal portion of *Ta*TIL-1 is composed of hydrophilic residues followed by few hydrophobic residues. In *At*TIL, the hydrophobic section is even less accentuated. This profile does not fit the standard hydrophobic nature of the N-terminal signal peptide identified in ApoD, Blc and Lazarillo. Like Lazarillo, the *Ta*TIL-1 and *At*TIL proteins are longer than ApoD and Blc at their C-terminal end and possess a similar putative cleavage site. The hydrophobic C-terminal tail enables Lazarillo to receive a glycosylphosphatidylinositol (GPI) anchor.[12] This suggests that *Ta*TIL-1 and *At*TIL could also receive a GPI anchor. GPI anchoring is a post-translational addition of a lipid occurring in the endoplasmic reticulum lumen which links proteins to the external face of the plasma membrane. This type of modification has been reported in plants.[13] The fact that the N-glycosylation site is conserved between the wheat and *Arabidopsis* TIL orthologs supports the possibility that these proteins are processed in

the endoplasmic reticulum lumen. Another type of attachment to the membrane can also be suggested for *Ta*TIL-1 and *At*TIL. It has been proposed that human ApoD is associated with the external face of the membrane by a hydrophobic loop.[3,14-15] *Ta*TIL-1 and *At*TIL also possess a hydrophobic stretch of seven amino acids that is inserted into a loop between two b-strands. This hydrophobic stretch is in the loop between b-strands 5 and 6 instead of being in the loop between strands 7 and 8, as is the case in the human ApoD (Fig. 1 B1,C2). It is nevertheless possible that this stretch favours the attachment of TILs to the plasma membrane. The loop scaffold in *Ta*TIL-1 and *At*TIL is two amino acids longer than in the human ApoD and there is a proline at positions 32 and 29 respectively. These modifications suggest that the plant TILs have a different binding specificity. A recent proteomic analysis of highly purified plasma membranes from *Arabidopsis* showed that *At*TIL is associated with this membrane fraction.[17] This result confirms the prediction that *Ta*TIL-1 and *At*TIL are membrane-associated proteins. However, the nature of the association or attachment is still unknown.

Northern blot analysis revealed that the *TaTIL-1* transcripts accumulate to high levels upon exposure to low temperature and heat-shock treatments (10-fold) and to a lesser extent after a water stress (3.5-fold).[7] Abscisic acid, high salt and wounding treatments have no measurable effect. The *TaTIL-1* transcripts accumulate gradually to a maximum level after 36 days of cold acclimation. Upon deacclimation, the level of transcripts returns to the level seen in the control nonacclimated plants. The accumulation of *TaTIL-1* transcripts in wheat was found to be tissue-specific, as they were detected only in cold-acclimated leaves. The expression analyses revealed that the dicot ortholog *AtTIL* is also induced by low temperature (6-fold) and heat-shock treatments (9-fold). RNA blot hybridization studies also demonstrated that cold acclimation induces the accumulation of *TaTIL-1* transcripts in both less tolerant and cold hardy wheat. However, this increase is greater in the hardy winter cultivars. Low levels of expression are also found in oat and barley, two less cold tolerant species.

This difference in accumulation indicates that the *TaTIL-1* expression is correlated with the plant's capacity to develop freezing tolerance.

Analysis of the promoter regions of *AtTIL, TaTIL-1, OsTIL-1* and *OsTIL-2* revealed the presence of several low temperature response elements (LTREs), dehydration response elements (DREs) and heat shock elements (HSEs). *TaTIL-1* and *AtTIL* promoter sequences contain more LTREs than the *OsTIL-1* promoter sequence. On the other hand, the *OsTIL-1* promoter contains more HSEs than the *AtTIL* and *TaTIL-1* promoters. This situation is not unexpected since rice does not have the ability to cold acclimate but possesses a higher thermotolerance than wheat and *Arabidopsis*. The fact that TIL promoters possess several light-responsive elements supports the specific expression of the corresponding genes in green photosynthetic leaves.

Temperature stresses are known to induce membrane injuries.[17] The membrane-anchored lipocalins (Blc, ApoD, Lazarillo, and possibly *Ta*TIL-1 and *At*TIL) all appear to be expressed in response to conditions that cause membrane stresses, which suggests a biological role in membrane biogenesis and/or repair under severe stress conditions.[3] The plant *Ta*TIL-1 and *At*TIL proteins, like the human ApoD, may bind a wide variety of potential ligands of varying structures and functions. The mammalian ApoD is reported to bind arachidonic acid, bilirubin, steroid hormones (progesterone and pregnenolone) and cholesterol.[4] It is interesting to mention that plants also synthesize a wide variety of steroid hormones called brassinosteroids. A treatment with 24-epibrassinolide, a brassinosteroid, increases the tolerance of plants to heat and cold stresses.[18] The enhanced resistance to temperature stress is attributed to increased membrane stability and osmoregulation. It is known that sterol insertion in the plasma membrane increases its fluidity at low temperature and maintains the phospholipids order at high temperature.[19] *Ta*TIL-1 may be involved in the transport of these sterol molecules to the membrane in response to stress conditions.

Other Plant Lipocalins

Since plant lipocalins were last reviewed, the sequencing of the *Arabidopsis thaliana* and *Oryza sativa* (rice) genomes has been completed.[6,20-21] The newly identified *Ta*TIL-1 and *At*TIL proteins were used to search the proteins predicted from the DNA sequence information of these two genomes using the BLAST program. The search revealed that rice possesses two other lipocalin members, TIL-2 and CHL. Sequence analysis revealed the presence of two different genes in rice encoding TIL lipocalins: *OsTIL-1* and *OsTIL-2* on chromosomes 2 and 8, respectively, whereas *Arabidopsis thaliana* has only *AtTIL* on chromosome 5. The *Os*TIL-1 and *Os*TIL-2 proteins share 65% identity and 80% similarity. *Os*TIL-2 is a protein of 179 aa with a calculated molecular mass of 21 kDa (Table 1). The absence of a N-terminal target peptide suggests that the *Os*TIL-2 protein would, like *Os*TIL-1, accumulate in the cytosol. Further sequence analysis of the wheat and rice TIL-2 proteins indicated the presence of a conserved putative N-glycosylation site. In addition, a putative C-terminal cleavage site is predicted by several target peptide prediction programs: DGPI, PSORT,[10] and SignalP.[11] Considering this putative cleavage site, the calculated molecular mass of the mature *Os*TIL-2 protein is 19 kDa.

The second new member identified from *Arabidopsis* and rice was named CHL (for chloroplastic lipocalin). This protein was identified in *Arabidopsis* as a putative lipocalin (CAB41869).[6] An homology search revealed that AtCHL shares only 23% identity and 40% overall similarity with *At*TIL. However, a region of 16 amino acids corresponding to SCR1 shows a high similarity with TIL lipocalins. The encoded mature proteins in *Arabidopsis* and rice are respectively 314 aa and 322 aa long with calculated molecular masses of 35 and 36 kDa (Table 1). SignalP and ChloroP predict N-terminal chloroplastic targeting peptides with high scores in both proteins (Table 1).[11,22] However, the exact length of the chloroplast transit peptide and the location of the proteins within the chloroplast is still unknown. A pairwise sequence alignment predicts chloroplast transit peptide cleavage sites near the

beginning of SCR1 in both *At*CHL and *Os*CHL sequences. The mature CHL proteins would have a molecular mass of 26 kDa, which is approximately the usual lipocalin size (Table 1). CHL proteins also possess 8 conserved cysteine residues probably involved in the three-dimensional structure of the protein by forming disulfide bridges. Motif searches against the PROSITE database,[23] after exclusion of patterns with a high probability of occurrence, revealed that *Arabidopsis* and rice CHL proteins possess the SCR1 lipocalin signature (Table 1). This signature perfectly fits the SCR1 consensus used by the ScanProsite software and exhibits the two invariant amino acids G and W that are key features of SCR1.[8-9,24] As in most lipocalins, CHL SCR2 is found in the C-terminal half of the protein and bears the conserved TDY triplet (Table 1).[8-9] SCR3 is also found in the C-terminal portion of both proteins and the conserved R residue that characterizes this fingerprint is present (Table 1).[8-9]

Violaxanthin De-Epoxidases and Zeaxanthin Epoxidases

Violaxanthin de-epoxidases (VDEs) and zeaxanthin epoxidases (ZEPs) are the most puzzling members with regards to their classification as plant lipocalins. The size and the exon-intron architecture of the genes encoding these enzymes show no significant similarity to the genomic organization of bacterial and animal lipocalin genes and for these reasons, they were not considered as true lipocalins in most studies.[25-26] These enzymes are involved in photoprotection of the photosynthetic apparatus, and are first synthesized as precursor proteins that bear the transit peptide needed for translocation to the thylakoid space of chloroplasts.[6,27] They share the common substrate antheraxanthin and are believed to exhibit similar tertiary structure.[6] VDEs are predicted to be proteins with a central barrel structure flanked by a cysteine-rich N-terminal domain and a glutamate-rich C-terminal domain (Table 1).[28] ZEPs possess ADP-binding and FAD-binding domains and fit the description of a lipocalin based on SCR1 homology (Table 1). Functional analyses of the different domains of VDEs demonstrated that the deletion of any of the cysteine residues in the

N-terminal region resulted in a total loss of activity.[28] This is likely because cysteine residues allow the formation of disulfide bridges, which are important determinants of protein conformation. It thus appears that the conformation of the mature protein in the N-terminal portion of VDEs is essential to retain their activities. Deletion analysis of the C-terminal region demonstrated that 71 out of 98 aa could be removed without any loss of activity.[28] However, removal of another 12 aa resulted in a 90% loss of activity and an important reduction of the binding of VDEs to the thylakoid membrane.[28]

Given the feature of VDEs and ZEPs and the strict definition of lipocalins, it is difficult to unequivocally consider these two proteins as true lipocalins. They are at best lipocalin-like proteins that could have arisen from the fusion of an ancestral plant lipocalin to proteins with enzymatic functions.[26,29] Thus, VDEs and ZEPs may represent the first example of lipocalins evolution towards the acquisition of novel functions.

Evolutionary Origin of Plant Lipocalins and Lipocalin-Like Proteins

To help elucidate the evolutionary origin of plant lipocalins, we investigated the presence of lipocalins and lipocalin-like proteins in algae and cyanobacteria. Algae are considered primitive photosynthetic eukaryotes while cyanobacteria carry a complete set of oxygenic photosynthetic genes. The chloroplast is believed to have evolved from the endosymbiosis of a cyanobacterial ancestor with a eukaryotic host cell. An homology search performed with the *Ta*TIL-1 protein sequence revealed several ESTs from red algae. The search also revealed that cyanobacteria possess a lipocalin gene.

Phylogenetic analyses suggest that TIL lipocalin members were probably inherited from a bacterial gene present in the original host cell, the common ancestor of plants and animals.1 In some plant species, the TIL-2 lipocalin may have arisen from the duplication of the gene encoding the TIL-1 lipocalin. However, the

remaining plant lipocalin and lipocalin-like members CHLs, VDEs and ZEPs might have evolved from a series of duplication of the cyanobacterial ancestor gene after cyanobacteria endosymbiosis from which the chloroplast originated. VDE and ZEP sequences subsequently diverged and acquired new cellular function as xanthophylls cycle enzymes.

Conclusion

The identification and characterization of plant lipocalins and lipocalin-like proteins will help in designing experiments aimed at the understanding of their cellular function in plants and their role in modulating the responses to temperature and oxidative stresses. Using forward and reverse genetics in the model system *Arabidopsis* should provide the information needed to elucidate the function of each protein in the plant metabolism. In addition, microarray analyses will help in the identification of the target genes associated with over / under expression of the different proteins. The ease with which plants can be manipulated and the availability of mutants are tremendous tools that should enable us to understand the cellular function of lipocalins and lipocalin-like proteins in plants. This information could even help understand the cellular function of lipocalins in mammals.

Acknowledgements

This work was supported by a Natural Sciences and Engineering Research Council of Canada discovery grant, and by Genome Canada, Genome Québec, and Genome Prairie grants to F. Sarhan. We thank Dr. François Ouellet for helpful discussions and editorial help.

References

1. D. Sánchez, MD. Ganfornina, and G. Gutiérrez, et al. Exon intron structure and evolution of the lipocalin gene family. *Mol Biol Evol* 2003. 20: (5) 775-783.

2. B. Akerstrom, DR. Flower, and JP. Salier. Lipocalins: Unity in diversity. *Biochim Biophys Acta* 2000. 1482: (1-2) 1-8.

3. RE. Bishop. The bacterial lipocalins. *Biochim Biophys Acta* 2000. 1482: (1-2) 73-83.

4. É. Rassart, A. Bedirian, and S. Do Carmo, et al. Apolipoprotein D. *Biochim Biophys Acta* 2000. 1482: (1-2) 185-198.

5. D. Sánchez, MD. Ganfornina, and MJ. Bastiani. Lazarillo, a neuronal lipocalin in grasshoppers with a role in axon guidance. *Biochim Biophys Acta* 2000. 1482: (1-2) 102-109.

6. AD. Hieber, RC. Bugos, and HY. Yamamoto. Plant lipocalins: Violaxanthin de-epoxidase and zeaxanthin epoxidase. *Biochim Biophys Acta* 2000. 1482: (1-2) 84-91.

7. JB. Frenette Charron, G. Breton, and M. Badawi, et al. Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. *FEBS Lett* 2002. 517: (1-3) 129-132.

8. DR. Flower, AC. North, and TK. Attwood. Structure and sequence relationships in the lipocalins and related proteins. *Protein Sci* 1993. 2: (5) 753-761.

9. DR. Flower. The lipocalin protein family: Structure and function. *Biochem J* 1996. 318: (Pt1) 1-14.

10. K. Nakai and P. Horton. PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 1999. 24: (1) 34-36.

11. H. Nielsen, J. Engelbrecht, and S. Brunak, et al. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* 1997. 8: (5-6) 581-599.

12. MD. Ganfornina, D. Sánchez, and MJ. Bastiani. Lazarillo, a new GPI-linked surface lipocalin, is restricted to a subset of neurons in the grasshopper embryo. *Development* 1995. 121: (1) 123-134.

13. N. Morita, H. Nakazato, and H. Okuyama, et al. Evidence for a glycosylinositolphospholipid-anchored alkaline phosphatase in the aquatic plant Spirodela oligorrhiza. *Biochim Biophys Acta* 1996. 1290: (1) 53-62.

14. MC. Peitsch and MS. Boguski. Is apolipoprotein D a mammalian bilin-binding protein? *New Biol* 1990. 2: (2) 197-206.

15. RE. Bishop, SS. Penfold, and LS. Frost, et al. Stationary phase expression of a novel Escherichia coli outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins. *J Biol Chem* 1995. 270: (39) 23097-23103.

16. MC. Peitsch. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. *Biochem Soc Trans* 1996. 24: (1) 274-279.

17. Y. Kawamura and M. Uemura. Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation. *Plant J* 2003. 36: (2) 141-154.

18. SD. Clouse and JM. Sasse. Brassinosteroids: Essential regulators of plant growth and development. *Annu Rev Plant Physiol Plant Mol Biol* 1998. 49: 427-451.

19. RA. Demel and B. De Kruyff. The function of sterols in membranes. *Biochim Biophys Acta* 1976. 457: (2) 109-132.

20. The *Arabidopsis* Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana. Nature* 2000. 408: 796-815.

21. J. Yu, S. Hu, and J. Wang, et al. A draft sequence of the rice genome (Oryza sativa L. ssp. indica). *Science* 2002. 296: (5565) 79-92.

22. O. Emanuelsson, H. Nielsen, and G. von Heijne. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* 1999. 8: (5) 978-984.

23. CJ. Sigrist, L. Cerutti, and N. Hulo, et al. PROSITE: A documented database using patterns and profiles as motif descriptors. *Brief Bioinform* 2002. 3: (3) 265-274.

24. A. Gattiker, E. Gasteiger, and A. Bairoch. ScanProsite: A reference implementation of a PROSITE scanning tool. *Appl Bioinformatics* 2002. 1: (2) 107-108.

25. G. Gutiérrez, MD. Ganfornina, and D. Sánchez. Evolution of the lipocalin family as inferred from a protein sequence phylogeny. *Biochim Biophys Acta* 2000. 1482: (1-2) 35-45.

26. JP. Salier. Chromosomal location, exon/intron organization and evolution of lipocalin genes. *Biochim Biophys Acta* 2000. 1482: (1-2) 25-34.

27. RC. Bugos, AD. Hieber, and HY. Yamamoto. Xanthophyll cycle enzymes are members of the lipocalin family, the first identified from plants. *J Biol Chem* 1998. 273: (25) 15321-15324.

28. AD. Hieber, RC. Bugos, and AS. Verhoeven, et al. Overexpression of violaxanthin de-epoxidase: Properties of C-terminal deletions on activity and pH-dependent lipid binding. *Planta* 2000. 214: (3) 476-483.

29. MD. Ganfornina, G. Gutiérrez, and M. Bastiani, et al. A phylogenetic analysis of the lipocalin protein family. *Mol Biol Evol* 2000. 17: (1) 114-126.

Table 1. Structure features of known plant lipocalins.

Protein	Precursor/Mature Molecular Mass (kDa)	Subcellular Localisation	Cleavage site Position*	SCR1 GxWY	SCR2 TDY	SCR3 R	Conserved Cys Residues	Conserved N-glycosyl. Sites	Other Domains
At TIL-1	21 / 20	membrane	C-terminal	yes	D only	yes	0	1	no
Os TIL-1	22 / 20	membrane	C-terminal	yes	D only	yes	0	1	no
Ta TIL-1	22 / 20	membrane	C-terminal	yes	D only	yes	0	1	no
Os TIL-2	21 / 19	ND	C-terminal	yes	D only	yes	0	1	no
At CHL	39 / 26	chloroplast	N-terminal	yes	yes	yes	8	0	no
Os CHL	37 / 26	chloroplast	N-terminal	yes	yes	yes	8	0	no
At VDE	52 / 40	chloroplast	N-terminal	yes	DDY*	yes	14	1	yes**
Os VDE	50 / 40	chloroplast	N-terminal	yes	DDY*	yes	14	1	yes**
Ta VDE	52 / 40	chloroplast	N-terminal	yes	DDY*	yes	14	0	yes**
At ZEP	74 / 68	chloroplast	N-terminal	yes	D only	no	6	1	yes***
Os ZEP	68 / 63	chloroplast	N-terminal	yes	D only	no	5	1	yes***

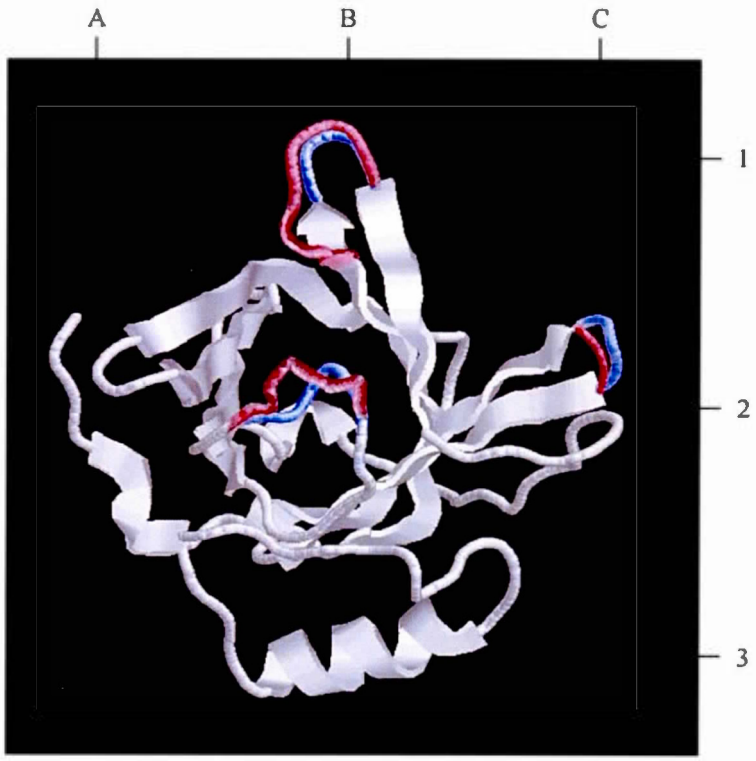
At = *Arabidopsis thaliana*; *Ta* = *Triticum aestivum* (wheat); *Os* = *Oriza sativa* (rice); Cys = Cysteine; ND = Not Determined

* – C-terminal, GPI anchor site; N-terminal, signal peptide.

** – N-terminal cyteine-rich region and C-terminal glutamic-acid-rich region

*** – N-terminal ADP-binding site and C-terminal FAD-binding site

Figure 1. Structural models of human ApoD and wheat *Ta*TIL-1.
Tertiary structure analyses were carried out using the Swiss-Model program.[16] The lower BLAST limit was set at 0.00001 and the human ApoD model (PDB ID: APD)[14] was used as template. The initial result was then resubmitted through the optimizing mode of the program. The final result was then visualized using the Swiss-Pdb Viewer and the model was adapted according to sequence comparison. Differences between the wheat and the human models were superimposed and colored. Grey sections are common to both models. The red (*Ta*TIL-1) and blue (ApoD) sections represent structural differences between the two proteins. Reprinted with permission from: Frenette Charron JB, Breton G et al. FEBS Lett 2002; 517(1-3):129-132.


A
B
C
1
2
3

CONCLUSION

Les travaux réalisés dans le cadre de ce projet de doctorat sont les premiers à identifier et à caractériser des membres de la famille des lipocalines chez les plantes. Les lipocalines étant une famille émergente de protéines impliquées dans de nombreuses fonctions cellulaires et dans de nombreuses maladies incurables chez l'être humain, il devenait impératif de caractériser cette famille de protéines chez les plantes.

Les résultats présentés dans le chapitre II identifient pour la première fois la présence d'une lipocaline chez les plantes. L'expression de cette dernière étant associée à la tolérance au gel chez le blé et aux tissus photosynthétiques, cela nous a permis de formuler une première hypothèse quant à la fonction des lipocalines chez les plantes. Cette hypothèse associait les lipocalines à la protection des tissus photosynthétiques lors de conditions de stress abiotiques tels que le gel.

Afin de raffiner cette hypothèse, nous avons entrepris dans le chapitre III de caractériser l'ensemble des lipocalines de plantes. Étant donné le peu d'information disponible sur les lipocalines végétales, nous avons utilisé une approche bioinformatique afin d'obtenir le maximum d'informations pertinentes sur les caractéristiques structurales et fonctionnelles de ces protéines. Ceci nous a révélé de nombreuses facettes encore inconnues des lipocalines de plantes et nous a permis d'identifier, de localiser et de catégoriser ces protéines. Cette approche bioinformatique nous a aussi permis de préciser notre hypothèse concernant la fonction de ces protéines. Les lipocalines de plante protègent la cellule contre les dommages dus aux stress abiotiques en jouant le rôle de détoxifiant afin d'enrayer les effets de l'oxydation causée par des molécules potentiellement dangereuses induites lors de stress de température et d'excès de lumière.

Les travaux présentés dans le quatrième chapitre visaient à vérifier cette hypothèse. Il a été démontré que des plantes qui n'accumulent pas *At*TIL sont très

sensibles aux baisses de température et au stress oxydatif et que ce phénotype peut être renversé lorsque l'accumulation de cette protéine est rétablie. Inversement, la surexpression de *At*TIL augmente la tolérance des plantes à ces deux stress. De plus, l'accumulation de *At*TIL retarde la floraison et le vieillissement de la plante. Tous ces résultats cadrent bien avec notre hypothèse de travail énoncée précédemment concernant la protection contre les dommages du stress oxydatif. Cependant, l'analyse de l'expression génétique globale a, quant à elle, été quelque peu surprenante. Cette surprise n'est pas nécessairement venue du fait que l'on a retrouvé, lors de cette analyse, des gènes régulant l'horloge circadienne ou étant impliqués dans la balance énergétique cellulaire, mais bien par ce que l'on n'a pas retrouvé. En effet, presque aucun gène associé à la tolérance aux stress ne semble être exprimé différentiellement chez les plantes n'exprimant pas *At*TIL. Ces résultats suggèrent donc que *At*TIL affecte une voie métabolique alternative qui module le niveau d'énergie cellulaire dans le but d'accroître la tolérance au stress oxydatif. Cette dernière affirmation est d'ailleurs appuyée par des études effectuées chez la Drosophile surexprimant une lipocaline apparentée à *At*TIL. À l'instar des plantes, les mouches surexprimant la lipocaline présentent une tolérance accrue au stress oxydatif et une longévité accrue tandis que les mouches n'exprimant pas cette lipocaline sont plus sensibles aux stress et vivent moins longtemps. Ces observations combinées à nos résultats suggèrent que les lipocalines possèdent une fonction conservée entre les espèces visant à accroître la tolérance aux stress.

L'ensemble des travaux présentés dans cette thèse de doctorat met en évidence le caractère multigénique du processus d'acclimatation au froid et de tolérance au gel. De plus, les lipocalines s'affichent comme des candidats de premier plan pour l'amélioration d'espèces économiquement importantes, tels le blé et autres plantes céréalières, dans le but de leur conférer une tolérance accrue aux stress abiotiques.

ANNEXE I (autre contribution)

Expression profiling and bioinformatics analyses of a novel stress-regulated multispanning transmembrane protein family from cereals and Arabidopsis

Ghislain Breton, Jean Danyluk, Jean-Benoit F. Charron, Fathey Sarhan

Plant Physiology (2003) 132: 64-74

Pour les travaux associés à cet article, j'ai participé à l'élaboration du design expérimental et aux analyses bioinformatiques. J'ai réalisé l'ensemble des analyses d'expression des ARN messagers. J'ai aussi participé à la rédaction du manuscrit en prenant sous mon aile les parties de textes traitant des analyses d'expression des ARN messagers dans les sections résultats et discussion. J'ai de plus rédigé l'ensemble des sections matériel et méthodes et légendes de figures.

Résumé

L'acclimatation au froid est un caractère multigénique permettant à certaines plantes de développer la tolérance au gel nécessaire à leur survie durant l'hiver. Afin de comprendre la nature génétique de ce caractère chez le blé, nous avons isolé plusieurs gènes induits par les basses températures. Certains de ces gènes codent pour de nouvelles protéines sans fonction connue. En intégrant les résultats de l'analyse des banques de données nucléotidiques, des prédictions bioinformatiques et des profils d'expression génétique, nous avons été en mesure de proposer des fonctions pour certaines des nouvelles protéines. Un des gènes identifiés fait partie d'une petite famille codant pour deux groupes distincts de protéines multitransmembranaires qui sont potentiellement localisées à la membrane plasmique (COR413-PM) ou à la membrane thylacoïdale (COR413-TM). Le premier groupe possède des sites potentiels de phosphorylation et d'ancrage de phosphatidylinositol glycosylé. Il est possible que cette famille de protéines soit spécifique aux plantes puisqu'aucun homologue n'a été trouvé dans les banques de données. L'analyse de la régulation des messagers de cette famille a démontré que l'expression de certains membres chez le blé et *Arabidopsis thaliana* était associée avec le développement de la tolérance au gel. De plus, l'expression de plusieurs est régulée par la sécheresse, la lumière et l'acide abscissique. L'analyse de la structure de plusieurs membres de la famille nous a mené à proposer que cette famille code possiblement pour des protéines similaires aux récepteurs couplés aux protéines G.

Mots clés : acclimatation au froid; *Arabidopsis thaliana*; cor413; lipocaline; membrane plasmique; récepteurs couplés aux protéines G; *Triticum aestivum* L

Abstract

Cold acclimation is a multigenic trait that allows hardy plants to develop efficient tolerance mechanisms needed for winter survival. To determine the genetic nature of these mechanisms, several cold-responsive genes of unknown function were identified from cold-acclimated wheat (*Triticum aestivum*). To identify the putative functions and structural features of these new genes, integrated genomic approaches of data mining, expression profiling, and bioinformatic predictions were used. The analyses revealed that one of these genes is a member of a small family that encodes two distinct groups of multispanning transmembrane proteins. The cold-regulated (COR)413-plasma membrane and COR413-thylakoid membrane groups are potentially targeted to the plasma membrane and thylakoid membrane, respectively. Further sequence analysis of the two groups from different plant species revealed the presence of a highly conserved phosphorylation site and a glycosylphosphatidylinositol-anchoring site at the C-terminal end. No homologous sequences were found in other organisms suggesting that this family is specific to the plant kingdom. Intraspecies and interspecies comparative gene expression profiling shows that the expression of this gene family is correlated with the development of freezing tolerance in cereals and *Arabidopsis*. In addition, several members of the family are regulated by water stress, light, and abscisic acid. Structure predictions and comparative genome analyses allow us to propose that the *cor413* genes encode putative G-protein-coupled receptors.

Introduction

To achieve their complete life cycle and reproduction in temperate regions, hardy plants like winter wheat (*Triticum aestivum*) have developed two major evolutionary adaptative mechanisms: vernalization and cold acclimation (CA). Overwintering plants sense the upcoming winter and delay flowering by postponing the transition from the vegetative to the cold-sensitive reproductive phase (Simpson et al., 1999). In addition, they develop the high degree of freezing tolerance (FT) needed for winter survival (Fowler et al., 1999). Following low temperature (LT) acclimation, some winter cereals can tolerate temperatures as low as −33°C. The regulatory mechanisms underlying these two processes and how they are interconnected are far from being fully understood. To gain further knowledge on the strategies that plants use for winter survival, the identification of cold-regulated (COR) genes is needed. A survey of the literature reveals that the expressions of a large number of genes are altered during the process of CA (Thomashow, 1999; Breton et al., 2000; Seki et al., 2002). These genes could be classified into four groups based on the presumed function of the encoded proteins. The first group comprises genes encoding structural proteins that may be involved in protecting the cell during LT stress. The second group represents those genes that regulate gene expression and signal transduction pathways, such as transcription factors, protein kinases, phosphatases, and the enzymes involved in phosphoinoside metabolism. The third group represents genes encoding enzymes involved in the biosynthesis of different osmoprotectants and membrane lipids and those of the antioxidative response. The fourth group contains cold-induced genes encoding proteins of unknown function.

To gain insight into the function of these novel proteins, a combination of expression profiling and bioinformatic analyses can be used to predict properties and features that may be important for their function. When a novel gene is found to be up-regulated by LT and its expression shows an association with the plants' capacity to develop FT, it is reasonable to assume that the encoded novel protein may play a

role in FT. By taking advantage of intraspecies variability in FT, a second level of association can be determined. Previous studies have shown that, compared with winter varieties, the less hardy spring wheat varieties cannot maintain the expression of COR genes (e.g. the WCS120 family) at a high level and that this differential expression is closely associated with their low degree of FT (Sarhan et al., 1997). A third level of association can be further established by taking advantage of the natural diversity of plant species. For example, species such as rice (*Oryza sativa*) and maize (*Zea mays*) are highly sensitive to LT above the freezing point, whereas species such as winter wheat and rye can tolerate temperatures as low as −33°C. This association can help differentiate between cold-responsive genes related to cold performance from those related to the acquisition of FT.

As a subsequent step, each novel LT-regulated protein sequence can be analyzed using available bioinformatic tools. These tools help in the identification of sorting signals, conserved posttranslational modifications, transmembrane helices, and secondary and tertiary structures. The most recent prediction software incorporate machine-learning algorithms in the form of a neural network and a hidden Markov model (Blom et al., 1999; Krogh et al., 2001). Comparative studies have shown that their prediction accuracy is often superior to older programs (Möller et al., 2001; Tusnàdy and Simon, 2001) and is bound to improve further when more newly characterized proteins are included in their training data sets. Knowledge gained from analyzing novel proteins with such tools can lead to the identification of important functional domains, an element needed to design future experiments to confirm the predicted function.

In the present study, the integrated approaches of expression profiling, structural analysis, and bioinformatic predictions were used to study a novel unknown gene family named *cor413*. This family encodes two distinct groups of proteins containing five putative transmembrane domains (TMD). COR413-plasma membrane (COR413-PM) proteins are potentially targeted to the plasma membrane and COR413-thylakoid membrane (COR413-TM) proteins to the thylakoid. The use of

intraspecies and interspecies comparative gene expression analysis shows that the regulation of this gene family is associated with the development of FT in cereals and *Arabidopsis*. A proposed structural and functional model for the COR413 protein family is discussed.

Results

Identification of TaCOR413-PM1 Homologs

Differential screening of a wheat cold-acclimated cDNA library was used to isolate LT-responsive clones. One of these clones, *Tacor413-pm1* (previously *Wcor413*), was selected for detailed molecular characterization (Danyluk, 1996). Sequence analysis revealed that the longest open reading frame (ORF) encodes a 210-amino acid (23 kD), highly hydrophobic protein with a predicted pI of 9.0 (Table I, *Ta*COR413-PM1). A search in the GenBank nonredundant sequence database using the BLAST program revealed that *Ta*COR413-PM1 is a novel protein with no characterized homologs. Data mining of the GenBank EST database with *Ta*COR413-PM1 revealed that plants possess several homologs of this protein. A combination of EST sequencing and in silico reconstitution allowed the generation of 27 new COR413-related protein sequences from plants. Using pair wise sequence alignments with the initial *Ta*COR413-PM1, these proteins were clustered into two distinct groups (Table I). The first group is named COR413-PM and contains members sharing more than 54% overall identity with *Ta*COR413-PM1 (Table I). The second group is named COR413-TM and contains members sharing less than 30% overall identity with *Ta*COR413-PM1 (Table I). However, a region of 40 amino acids shows a higher degree of homology among all members of both groups (Supplemental Figs. 1-3, square brackets; they can be viewed at www.plantphysiol.org).

Data mining of cereal EST databases and rice genomic sequence helped in the identification of two different COR413-PM members in wheat, maize, and barley (*Hordeum vulgare*), whereas only one was identified in rice (*O. sativa* subsp. *indica* cv 93-11; Yu et al., 2002). In addition, four COR413-PM proteins were identified in the *Arabidopsis* genome. On the other hand, only one member belonging to the COR413-TM group was identified in the four cereal species analyzed. In *Arabidopsis*, two COR413-TM were found in tandem repeat on chromosome 2 (The *Arabidopsis*

Genome Initiative, 2000). Furthermore, a search in the GenBank EST database revealed that other dicotyledonous plants such as tomato (*Lycopersicon esculentum*), soybean (*Glycine max*), ice plant (*Mesembryanthemum crystallinum*), poplar (*Populus* spp.), and cotton (*Gossypium hirsutum*) as well as the coniferales *Cryptomeria japonica* and *Pinus taeda* possess sequences encoding homologs of the COR413 groups (see Supplemental Table II; supplemental tables can be viewed at www.plantphysiol.org).

Other embryophytes such as the marchantiales *Marchantia polymorpha* and the moss *Physcomitrella patens* also have COR413 homologs. The deduced moss COR413 proteins share slightly higher identity with the COR413-PM group, suggesting that they are related to this group (Table I). Because of their lower degree of homology, they were classified separately in this study as moss COR413 (Table I). This lower homology may result from the evolutionary distance between moss and other plants listed in Table I. Because no entries encoding COR413 homologs were found in the green algae *Chlamydomonas reinhardtii* sequence database, it is possible that COR413 would be present only in multicellular Viridiplantae. COR413 homologous sequences were neither found in other eukaryotes nor in prokaryote databases, suggesting that this family is specific to the plant kingdom.

The cor413 Genes Encode Membrane Proteins Potentially Targeted to the Plasma and Thylakoid Membranes

The analyses of both COR413-PM and -TM sequences revealed that they are rich in hydrophobic amino acids, suggesting that they may be membrane proteins (Supplemental Figs. 1 and 2). The sequence alignments of both groups show many regions of high identity (shaded in black in Supplemental Figs. 1 and 2). In addition, many residues normally considered important for protein structure or activity such as Cys residues and Pro residues are conserved within COR413-PM or COR413-TM proteins (Supplemental Figs. 1-3, asterisks). Five of the Pro residues are even

conserved between members of both groups (yellow-shaded asterisks). On the other hand, the sequence alignments also revealed that approximately the first 50 amino acids of COR413-PM and the first 80 amino acids of COR413-TM are poorly conserved. This observation prompted us to analyze these regions for subcellular targeting signals.

Analysis using the PSORT program revealed that there is no consensus targeting or retention signal present in COR413-PM sequences (Nakai and Kanehisa, 1992; Supplemental Table III). Although the program suggested different cellular localizations for each member, the average score was slightly higher for the plasma membrane localization. Because many proteins targeted to the plasma membrane possess a cleavable signal peptide, COR413-PM sequences were analyzed with SignalP (Nielsen and Krogh, 1998). The analysis of SignalP-HMM results revealed that six proteins have a high probability to possess a non-cleavable signal anchor for endoplasmic reticulum (ER) translocation (Supplemental Table III). These results are consistent with those obtained with PSORT and suggest that COR413-PM proteins are targeted to the plasma membrane. Moss COR413 shows the same features as the COR413-PM group, suggesting that they are also targeted to the plasma membrane (Supplemental Table IV). Using several secondary structure prediction programs available on the Network Protein Sequence Analysis server, we found that the N-terminal region of COR413-PM proteins contains a possible hinge-like structure consisting of two segments of 20 to 25 residues predicted to form α-helices that are separated by a Gly-rich region (Supplemental Fig. 1).

The use of the targeting signal programs PSORT, iPSORT, and TargetP for COR413-TM sequence analyses revealed that they are all likely to be targeted to the thylakoid membrane (Supplemental Table V; Nakai and Kanehisa, 1992; Emanuelsson et al., 2000; Bannai et al., 2002). Chloroplast targeting signals are generally highly basic and rich in Ser and Thr (Agarraberes and Dice, 2001). The N-terminal sequence of all COR413-TM members shows these two properties (Supplemental Fig. 2).

COR413 Proteins Contain Five TMD

As expected from the overall amino acid composition, the Kyte and Doolittle hydrophobicity plot of *Ta*COR413-PM1 shows a highly hydrophobic pattern with six clear spikes (Supplemental Fig. 1, S1-S6; Kyte and Doolittle, 1982). Superposition of the Kyte and Doolittle plot of the 15 available COR413-PM sequences showed that the overall hydrophobicity is well conserved among the different members (Fig. 1A). To analyze the number of TMD and the possible topology of COR413-PM proteins, the newly developed and accurate membrane topology prediction program TMHMM was used (Krogh et al., 2001; Möller et al., 2001). The final prediction generated by the program for each COR413-PM members is listed in Supplemental Table III, and the compilation of all TMD predictions is presented in Figure 1B. These analyses allowed us to propose two structural models (Fig. 2). In the first, COR413-PM proteins would have five TMD with the N-terminal end outside and the C-terminal end inside (Fig. 2, model 1). This model is supported by the following observations: (a) The final prediction of 11 of 15 proteins have this topology; and (b) the compilation of the N-probability graphs for TMD shows that the five-TMD topology is favored. In this compilation, spike 4 was chosen as the third TMD in nine of the 11 proteins. The data analysis of the inside/outside probability graphs generated by TMHMM and the calculation of the median probability revealed that the N-terminal topology is favored (64%). Although model 1 is the software's preferred topology, it does not take into account the following points: (a) All 15 proteins contain six hydrophobic spikes suggesting six TMD; and (b) four proteins of 15 are predicted to have six TMD with TMHMM. Therefore, an alternative model could be proposed where group COR413-PM members would have six TMD with both the N-terminal and C-terminal ends inside (Fig. 2, model 2). The main difference with the first model is the inverted topology in the first one-half of the protein. COR413-TM sequences were analyzed before and after removal of the putative N-terminal chloroplastic targeting signal. The comparison of the 10 Kyte and Doolittle profiles clearly shows

that COR413-TM proteins possess six hydrophobic spikes (S1-S6 in Supplemental Fig. 2 and Fig. 1A). Despite the clear hydrophobic pattern, TMHMM had difficulties generating clear topology predictions (Supplemental Table V). However, the compilation of the 10 TMHMM graphs suggests that group II may also have a five-TMD structure (Fig. 1B). Because no clear consensus can be deduced for the TMHMM inside-outside topology, it is impossible to predict which loops are exposed on the lumenal and stromal side of the thylakoids. The extracellular loop 1 of COR413-PM model 1 (Fig. 1, S3) falls in the region that is conserved between COR413-PM and -TM, and corresponds to the second predicted loop and the third TMD of COR413-TM (Supplemental Figs. 1-3; identified in blue in Fig. 2).

COR413-PM Proteins Contain Conserved Putative Phosphorylation and Glycosylated Phosphatidylinositol (GPI)-Anchoring Sites

Motif searches against the PROSITE, Pfam, and Smart databases, after exclusion of patterns with a high probability of occurrence, did not detect known motifs. However, the neural network-based NetPhos phosphorylation site prediction software generated several interesting findings (Blom et al., 1999). Even though the *Ta*COR413-PM1 sequence contains eight Ser residues, 10 Thr residues, and four Tyr residues, only one Thr residue is predicted to be a phosphorylation site (Supplemental Fig. 1 in yellow). Analysis of the other COR413-PM members with the NetPhos software always identified a putative phosphorylation site at this position (Supplemental Fig. 1). Interestingly, in both of our models, this phosphorylation site is located on the internal side of the membrane where it may be the target for intracellular kinases (see Fig. 2). For the chloroplastic COR413-TM proteins, NetPhos predicted a phosphorylation site in the same region (between TMD3 and TMD4) for eight of the 10 proteins. The other two proteins are those from *Arabidopsis*, which raises the possibility that the prediction of the phosphorylation site for the chloroplastic proteins may be incorrect (Supplemental Fig. 2).

The DGPI program predicted a GPI-anchoring site at the C-terminal end of all COR413-PM family members (D. Buloz and J. Kronegg, unpublished data). The conserved features are a highly hydrophobic C-terminal end and a consensus cleavage site needed for the addition of the GPI anchor (Supplemental Fig. 1). This second posttranslational modification fits well with our structural models because GPI anchors are modifications located on the external side of the membrane (Fig. 2). This modification will result in the cleavage of the second extracellular loop as schematized in Figure 2. The chloroplastic COR413-TM proteins' C-terminal tail is also very hydrophobic but the potential cleavage sites are less conserved.

cor413 Genes Are Regulated by Environmental Stresses

Northern-blot analyses indicated that *Tacor413-pm1* and *Tacor*413-*tm1* transcripts are strongly up-regulated by LT in leaf tissues (Fig. 3 A). In contrast, the *Tacor413-pm2* transcript was down-regulated. The LT kinetics study in winter wheat cv Norstar leaves shows that the *Tacor413-pm1* and *tm1* transcripts accumulate rapidly within 24 h and remain at high levels throughout the acclimation period (Fig. 3A). In comparison, the transcripts accumulation in the less freezing-tolerant spring wheat cv Glenlea peaks at 24 h and then declines (Fig. 3A). When the plants are deacclimated at 24°C for 5 d, *Tacor413-pm1* and *tm1* transcripts decline to the nonacclimated control levels in both cultivars. The intra- and interspecies comparative expression analyses are shown in Figures 3B and 4. *Tacor413-pm1* and *tm1* mRNA levels are higher in winter wheat cultivars compared with the less FT spring wheat cultivars (Glenlea and Concorde). These results suggest that the accumulation of *Tacor413-pm1* and *tm1* transcripts is associated with the capacity of the plants to develop FT. This figure also shows that *Tacor413-pm2* level is slightly down-regulated by long term LT treatments because transcript levels are higher in nonacclimated wheat leaves than in the 36-d-acclimated ones. The use of the wheat *Tacor413-pm1* and *tm1* full-length probes revealed that LT-sensitive oat and LT-

tolerant barley and rye also possess cold-inducible homologs of the *cor413* family (Fig. 4A). The wheat probes did not detect any signal in rice, but the use of rice-specific probes showed that the transcript level of *Oscor413-pm1* is detectable but not LT-regulated under the four temperature regimes used. In contrast, *Oscor413-tm1* transcripts are practically undetectable (Fig. 4B). Results obtained with the maize *Zmcor413-pm1* and *-tm1* probes using similar treatments have shown that both transcripts are undetectable (data not shown). In *Arabidopsis*, the *Atcor413-pm1* and *Atcor413-tm1* transcripts accumulate in response to the LT treatments, but *Atcor413-pm2* transcripts are undetectable (Fig. 4C).

To determine whether the wheat and *Arabidopsis* cor413 gene families are specifically regulated by LT, plants were subjected to different stress treatments (Figs. 4C and 5A). RNA gel-blot analysis indicated that water stress induces the accumulation of *Atcor413-pm1* and *-tm1* as well as *Tacor413-pm1* and *-tm1* transcripts to a level comparable to 1 d of LT exposure. Exogenous application of the stress-associated growth regulator abscisic acid (100 μM) also induced the accumulation of the four transcripts. Taken together, these results suggest that the *At*COR413-PM1 and TM1 proteins could be dicotyledonous orthologs of the wheat COR413-PM1 and TM1 proteins.

Tissue Specificity and Light Regulation of cor413 Genes

The expression data in Figure 5B shows that, under LT conditions, *Tacor413-pm1* is expressed more abundantly in leaves and roots, whereas the chloroplastic protein-encoding *Tacor413-tm1* accumulated only in the photosynthetic tissues. To further investigate the association of the *Tacor413-tm1* expression profile with photosynthetic tissue, we analyzed the regulation of the *cor413* family members under different light conditions. The results in Figure 5C show that the LT-induced expression of *Tacor413-pm1* is not light dependent and is not associated with the chloroplast differentiation stage. In contrast, *Tacor413-tm1* LT accumulation is

dependent on the chloroplast differentiation stage because *Tacor413-tm1* accumulation is higher in light-grown plants than in etiolated plants.

To take advantage of the large body of information generated from the different plant EST projects, we analyzed systematically each GenBank *cor413*-related entry for information regarding tissue specificity. The result of our survey is presented in Supplemental Table II. In addition to the leaf and roots tissues, c*or413-pm* and *cor413-tm* transcripts were found in wheat pre-anthesis spike, maize glume, and rice panicle. In dicotyledonous plants, they are found in *Arabidopsis* flower buds, cotton post-anthesis fiber bolls, potato (*Solanum tuberosum*) sprouting eyes, alfalfa (*Medicago sativa*) root tips, soybean immature flowers, and tomato flower buds and maturing fruits. *Cor413* members were also found in *P. taeda* bark tissue, *P. patens* protonemata, and *M. polymorpha* immature sex organs. This survey reveals that *cor413-pm* and *-tm* expression is not restricted to the plant vegetative stage but also occurs in the final phase of the reproductive stage.

Discussion

A combination of comparative expression profiling and bioinformatic tools was used to identify and to characterize a novel family of plant multispanning transmembrane proteins. Data mining of various nucleotide databases and protein sequence alignments revealed that the higher plant COR413 family can be clustered into two distinct groups. Sequence analyses revealed four important conserved features on COR413-PM and two on COR413-TM. These features are related to the cellular localization, the structure, and the presence of a phosphorylation site and of a GPI-anchoring site.

The bioinformatic approach used allowed us to propose that COR413-PM proteins are targeted to the plasma membrane and that COR413-TM proteins are targeted to the thylakoid membrane. The predicted localization of COR413-TM proteins is further corroborated by the fact that their corresponding transcripts are more abundant in photosynthetic tissues and are regulated by the chloroplast differentiation stage. The existence of some EST entries from non-photosynthetic tissues such as alfalfa developing flowers, *Arabidopsis* flower buds, potato sprouting eyes, and barley etiolated tissues suggests that the proteins may also be associated with other plasts.

The hidden Markov-based TMHMM software predicted that both COR413-PM and -TM proteins are likely to possess five transmembrane helices and that the N-terminal end of group COR413-PM may be located on the extracellular side of the plasma membrane. Although this is the most probable topology for the moment, TMHMM had a difficulty reaching a consensus topology, especially with COR413-TM proteins. This may be due to the fact that TMHMM, which uses a machine-learning algorithm, was tested using a data set containing very few plant membrane proteins (Krogh et al., 2001). This hypothesis suggests that the prediction of plant membrane protein structures will certainly become more accurate with time when more plant data becomes available for data set generation. This particular program

was used for two main reasons: first, it is considered the most accurate prediction software (Möller et al., 2001), and second, because it is based on a hidden Markov algorithm. A recent review on membrane protein topogenesis concluded that the current knowledge makes it impossible to establish consensus rules because too many different processes seem to influence simultaneously the insertion of the protein into the membrane (Goder and Spiess, 2001). On the basis of this conclusion, we believe that machine learning algorithms such as the hidden Markov algorithm can better take into account the subtle differences in amino acids that cannot be deduced by any other method.

The use of the neural network-based predictor NetPhos 2.0 (Blom et al., 1999) suggested that COR413-PM members are likely to possess a different phosphorylation site on the second intracellular loop. It is worth mentioning that the software predicted a phosphorylation site in the loop between TMD3 and TMD4 of the wheat *Ta*COR413-PM1 and -PM2 proteins, and this loop is a region that is highly divergent between the two proteins. This suggests that the two proteins may be regulated by different kinases (Supplemental Fig. 1).

The last consensus prediction was obtained from the anchoring site predictor DGPI. This program found the presence of a cleavage site and a favorable environment for the addition of a GPI anchor (proper hydrophobic tail length and hydrophilic region length) at the second extracellular loop of COR413-PM proteins. The only feature that DGPI did not detect on COR413-PM sequences is the presence of an N-terminal cleavable signal peptide for translocation to the ER, and neither was this feature detected by the accurate SignalP signal sorting predictor (Nielsen and Krogh, 1998). However, SignalP did predict the presence of a non-cleavable signal anchor for ER translocation in several COR413-PM. It is thus possible that in our case, a signal anchor may replace the signal peptide. Although no multispanning transmembrane proteins are currently known to be GPI-anchored (Borner et al., 2002), the results obtained from our bioinformatic analyses do not at this point rule out the possibility that COR413-PM proteins could be GPI-anchored. The GPI-

modified proteins are usually identified by their presence in the soluble fraction after GPI cleavage by specific lipases. Thus, multispanning transmembrane proteins will always remain attached to the membrane fraction and will not be identified as GPI-containing proteins. Thus a special experimental procedure needs to be developed to confirm our hypothesis.

Gene Expression Studies

To understand the function of the COR413 family, expression patterns were determined during several environmental stresses. The expression of one member of group COR413-PM and one from group COR413-TM was closely associated with the acquisition of FT in several plant species such as wheat, rye, and *Arabidopsis*. This observation is in agreement with the recent microarray analysis that identified *Atcor413-pm1* as an LT-inducible gene in *Arabidopsis* (Seki et al., 2001; clone FL3-5A3). On the other hand, group *Cor413-pm* and *-tm* transcripts were not induced in the LT-sensitive species rice and maize. Together, these results suggest that the *cor413* expression is not associated with a general metabolic response to LT. Furthermore, the wheat and *Arabidopsis cor413-pm1* and *-tm1* genes were also induced by water stress and abscisic acid. Interestingly, a *cor413* homolog was found in an EST survey of ABA-treated protonemata cells of the moss *P. patens* (Machuka et al., 1999). Furthermore, recent results have shown that a *P. patens* homolog of *cor413* is induced by ABA and slightly by LT, and these increases were associated with the development of FT of the protonemata cells (Nagao et al., 2001).

Putative COR413 Function

It is known that the plasma membrane is the primary site of freezing injury (Steponkus, 1984). To date, only highly soluble amphipatic proteins have been proposed to act as membrane-stabilizing proteins (Artus et al., 1996; Danyluk et al.,

1996). As an integral membrane protein, COR413-PM could play a structural role by stabilizing the plasma membrane lipid bilayer. If the proposed function is exact, the existence of a thylakoid COR413 may suggest that this membrane also needs structural reinforcement.

The second proposed hypothesis is that the COR413 protein family is associated with environmental stress signaling. This hypothesis is based on the comparison between our structural model (Fig. 2, model 1) and that of the mammalian Rhodopsin-like G-protein-coupled receptor (GPCR) family (Bockaert and Pin, 1999). GPCR is the largest family of receptors in animals, and sequence alignment studies have helped classify them into five large clusters. The largest cluster is named the Rhodopsin-like class A GPCR and contains at least 1,000 different members (Horn et al., 1998). All of these proteins share little sequence identity, but one triplet motif (E/D-R-Y) is highly conserved and has been the subject of numerous mutational studies (Scheer et al., 1996; Alewijnse et al., 2000; Chung et al., 2002, and refs. therein). It is located on the internal side of the membrane at the border of the third TMD and second intracellular loop of GPCR (Fig. 2). The aspartic/Glu residues contribute to maintain the receptor in its quiescent state (Chung et al., 2002). A similar motif (D-R/K-T) was found in the most conserved region of COR413-PM and moss COR413 (Supplemental Figs. 1 and 3), and it is also located on the internal side of the membrane at the border of TMD1 and the first intracellular loop (Fig. 2). Although there is compelling biochemical and molecular evidence for the existence of GPCR-based signaling in plants (the three components of the heterotrimeric G-protein are identified), no receptor has been clearly shown to act as a GPCR (Millner, 2001). Two plant proteins are actually considered GPCR. The first is the MLO protein family that is related to the animal GPCR family because it also possesses seven TMD (Devoto et al., 1999). The second GCR1, was isolated by its sequence homology with the *Dictyostelium* spp. cAMP GPCR (Josefsson and Rask, 1997; Plakidou-Dymock et al., 1998). These cAMP receptors are not clustered with the Rhodopsin-like family, but GCR1 possesses a motif similar to the D-R-Y triplet

(H-R-T). On the basis of transgenic studies Colucci et al. (2002) recently suggested that GCR1 may be the gibberellic acid receptor.

Three other features support the assumption that group COR413-PM proteins are related to the Rhodopsin-like GPCR family. GPCR are regulated by kinases, and the phosphorylation sites are often located on the last intracellular loop (Pitcher et al., 1998). It is on this loop that the conserved putative phosphorylation site was predicted for the COR413-PM proteins. The second feature is the presence of a lipid-anchoring site. It is known that saturated acyl chains are sometimes added on GPCR to link the C-terminal tail to the inner leaflet of the plasma membrane (Fig. 2; Bouvier et al., 1995). Therefore, a hydrophobic molecule is added to an already highly hydrophobic protein, as is predicted for COR413-PM members with the addition of a GPI anchor. In animal cells, GPI anchors are used to target the modified proteins into special cholesterol and sphingolipid-rich membrane domains named lipid rafts (Brown and London, 2000). These membrane domains were shown to be the site of intense signaling events. Similar domains were recently identified in the plant plasma membrane (Peskan et al., 2000). The third feature linking COR413-PM to GPCR is the presence of seven highly conserved Pro residues (Supplemental Fig. 1). In transmembrane proteins such as GPCR, some highly conserved Pro residues located inside the TMD are known to be important for correct folding and function (Sansom and Weinstein, 2000). The rigid body motion of the two portions of a Pro-kinked TMD is proposed as a key dynamic component in the rearrangement of GPCR structure upon activation by ligand binding. Interestingly, five of the seven Pro residues in COR413-PM sequences are conserved in COR413-TM sequences, suggesting that they may play the same role in both subgroups (Supplemental Figs. 1-3).

On the basis of these analyses, one may ask what is the specific ligand for COR413-PM? Several molecules can act as GPCR ligands in animal cells, and molecules sharing chemical characteristics with some of these ligands do exist in plants (Supplemental Table VI; Wink, 1997). Knowing that the larger extracellular

loop of COR413-PM members is the region with the highest homology with the chloroplastic COR413-TM, it is possible that both proteins bind the same ligand. The exact biochemical properties and function of this new protein family during LT acclimation remains to be determined. Nevertheless, the combination of data mining, bioinformatic analyses, and expression profiling presented here will help us in the design of experimental procedures aimed at answering those questions.

Materials and methods

Plant Material and Growth Conditions

In this study, we used two spring wheat genotypes (*Triticum aestivum* L. cv Glenlea, LT_{50} of −8°C; and cv Concorde, LT_{50} of -8°C); four winter wheat genotypes (*T. aestivum* L. cv Monopole, LT_{50} of -15°C; cv Absolvent, LT_{50} of -16°C; cv Fredrick, LT_{50} of -16°C; and cv Norstar, LT_{50} of -19); winter rye (*Secale cereale* L. cv Musketeer, LT_{50} of -21°C); oat (*Avena sativa* L. cv Laurent, LT_{50} of -6°C); barley (*Hordeum vulgare* L. cv Winchester, LT_{50} of −7°C); rice (*Oryza sativa* subsp. *indica* cv IR36, LT_{50} of 4°C); maize (*Zea mays*, LT_{50} of 4°C); and *Arabidopsis* ecotype Columbia (LT_{50} of -9°C). Growth of plants and stress treatments were as previously described (Frenette Charron et al., 2002).

Cloning and Data Mining

The *Ta*cor413-pm1 clone (previously p*Wcor413*) was isolated by differential screening of a Lambda Zap II library constructed from poly(A^+) RNA isolated from 1-d cold-acclimated winter wheat (cv Norstar; Houde et al., 1992). The *Ta*cor413-pm1 clone was purified and excised as a pBluescript vector following the library supplier's protocol (Stratagene, La Jolla, CA).

Database searches to identify *Ta*cor413-pm1 homologs were performed using the Canadian Bioinformatics Resource (Halifax, Nova Scotia, Canada; http://www.cbr.nrc.ca) and National Center for Biotechnology Information (Bethesda, MD; http://www.ncbi.nlm.nih.gov/BLAST) Web implementation of BLAST (Altschul et al., 1990) against the GenBank nonredundant sequence database and GenBank EST database (Benson et al., 2002). In the first round of data mining, COR413 homologs were identified by using the *Ta*COR413-pm1 protein sequence as query with TBLASTN against the GenBank EST database. In the second round, the

identified EST from each different plant species containing the longest 5' or 3' end were used as query to search the same database. Overlapping ESTs were assembled, and a consensus cDNA was deduced when two or more identical sequence could be aligned. To obtain the largest number of complete COR413 sequences, available clones were ordered, sequenced, and submitted to GenBank, and the others were deduced from the available genomic and EST sequences. The ORF of the assembled gene was identified using ORFinder on the NCBI Web site (T. Tatusov and R. Tatusov, unpublished data; http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The longest ORF was always chosen (except for *Cj*COR413 wher e the third ATG was chosen). The complete in silico assembled nucleotide sequence and its encoded protein were then used to screen back the EST database. This other round of data mining was useful for the identification of very near homologs with subtle amino acid differences. Only the complete COR413 homologs were used for subsequent structural and functional domain prediction analyses. Survey of all the homologs identified can be found in Supplemental Table VII. *At*COR413-pm3 and *At*COR413-pm4 were not used in the bioinformatic analysis. Although *At*COR413-pm3 and pm4 seem to be related to group COR413-PM, their sequences are slightly different, and no other similar plant proteins were found.

The degree of sequence identity in Table I was determined using ALIGN (Pearson, 1990) on the Biology Workbench workstation (http://workbench.sdsc.edu/). Group I and II sequences were aligned and analyzed by using ClustalW (Thompson et al., 1994) on the Biology Workbench (http://workbench.sdsc.edu/) and Network Protein Sequence Analysis servers (Combet et al., 2000; http://pbil.ibcp.fr/). Shading of amino acids was performed with the BOXSHADE program at the BOXSHADE Web server at the University of Lausanne (Switzerland; http://ulrec3.unil.ch/software/boxshade/boxshade.html).

Structural Analyses

For detection of specific targeting sequences, we used PSORT, iPSORT (Nakai and Kanehisa, 1992; Bannai et al., 2002; http://psort.nibb.ac.jp/), and TargetP v1.01 (Emanuelsson et al., 2000; http://www.cbs.dtu.dk). For detection of signal peptides, SignalP v2.0 was used (Nielsen and Krogh, 1998; http://www.cbs.dtu.dk). Before performing structural prediction, a Kyte and Doolittle hydropathic plot was generated by using the Protscale program (http://ca.expasy.org/cgi-bin/protscale.pl) with the Kyte and Doolittle option and a window of nine amino acids (Kyte and Doolittle, 1982). The superposition of COR413 hydropathic plot was generated by transferring the raw data to Microsoft Excel (Microsoft, Redmond, WA). The graph was constructed by aligning the data table to the last C-terminal amino acid therefore compensating for the various protein N-terminal lengths. For TMD prediction, TMHMM (http://www.cbs.dtu.dk) was used (Krogh et al., 2001). The TMHMM data tables were processed with Microsoft Excel as for hydropathic plot. α-Helical regions were identified with secondary structure prediction programs integrated in the Web implementation of ClustalW at the Network Protein Sequence Analysis Web site (Combet et al., 2000; http://pbil.ibcp.fr).

Other Prediction Servers

For functional domain identification, we first used ScanPROSITE on the Expasy Web server for PROSITE motif database screening (http://ca.expasy.org/tools/scanprosite/) and NCBI RPS-BLAST for Pfam and Smart conserved domain databases screening (*Altschul et al., 1997*; http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). In a subsequent search, we used most of the software available on the Expasy server (http://ca.expasy.org/). Two types of the software gave interesting results. NetPhos was used for consensus phosphorylation site detection (http://www.cbs. dtu.dk; Blom et al., 1999), and DGPI

was used for GPI-anchoring site detection (http://129.194.186.123/GPI-anchor/index_en.html; D. Buloz and J. Kronegg, unpublished data).

Expression Studies

Cereal RNA extraction and RNA gel-blot analysis were performed as already described (Houde et al., 1992). Total RNA from *Arabidopsis* and etiolated wheat were extracted using Tri Reagent (Molecular Research Center, Cincinnati) according to the manufacturer protocol. To prevent cross-hybridization between *Tacor413-pm1* and *-pm2* probe in northern analysis, specific probes for the 3'-non-coding region of each cDNA were used. These probes were generated by PCR with the following primers: *Tacor413- pm1*, 5'-ttcatctacccggtctgggccgtc and 5'-ccaggaaacaaactaagacgtgacacc; and *Tacor413-pm2*, 5'-agtctgggtcctggtgctc and 5'-tcataccagaactacaacaaatcg (*Tacor413-pm2* and *Tacor413-tm1* clones were kindly provided by Dr Anderson [U.S. Department of Agriculture-Agricultural Research Service-Plant Gene Expression Conter, Albany, CA). Northern blots of rice and maize samples were performed using the complete *Oscor413-pm1* or *Oscor413-tm1* clones (kindly provided by Dr. Sasaki as part of the Japanese Rice Genome Research Program of the National Institute of Agrobiological Sciences and the Institute of the Society of Techno-Innovation in Agriculture, Forestry and Fisheries; Yamamoto and Sasaki, 1997) and *Zmcor413-pm1* or *Zmcor413-tm1* clones (kindly provided by Dr. Singh [Agriculture and Agri-Food Canada]). The *Arabidopsis Atcor413-pm1* and *pm2* probes also showed cross-hybridization. Therefore, probes specific to the 5' non-coding regions were generated by PCR and used in the hydridizations. The first primer hybridized with the vector cloning site left border (5'-atagagctcactagtccggaattcccgggtcga) and the second hybridized specifically to the *Atcor413-pm1* or *-pm2* sequences (*Atcor413-pm1*, 5'-gtatatggcggcgattgaagcaacc; and *Atcor413-pm2*, 5'-tggcagcgaaagaagcgaggaatttga). For *Atcor413*-tm1 the complete cDNA was used as probe. Dr. Newman (Department of Energy-Plant Research

Laboratory, East Lansing, MI) kindly provided the three *Arabidopsis* clones from *Arabidopsis* Biological Resource Center (Ohio State University, Columbus) distribution services. Northern analyses for each sample were performed at least three times from two biological replicates. For other COR413 sequence analysis, Dr. Ujino-Ihara from the Forestry and Forest Products Research Institute kindly provided the *C. japonica* clone, Dr. Bashiardes as part of the Physcomitrella EST Program at the University of Leeds (UK) and Washington University (St. Louis) kindly provided the *P. patens* clone, and Dr. Anderson from Clemson University Genomic Institute kindly provided the barley clone. All distributed clones are identified in Supplemental Table VII.

Literature cited

Agarraberes FA, Dice JF (2001) Protein translocation across membranes. Biochim Biophys Acta 1513: 1-24

Alewijnse AE, Timmerman H, Jacobs EH, Smit MJ, Roovers E, Cotecchia S, Leurs R (2000) The effect of mutations in the DRY motif on the constitutive activity and structural instability of the histamine H(2) receptor. Mol Pharmacol 57: 890-898

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410

Altschul SF, Madden TL, Schäfer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3381-3402

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* Nature 408: 796-815

Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF (1996) Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. Proc Natl Acad Sci USA 93: 13404-13409

Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18: 298-305

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL (2002) Genbank. Nucleic Acids Res 30: 17-20

Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J Mol Biol 294: 1351-1362

Bockaert J, Pin JP (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J 18: 1723-1729

Borner GH, Sherrier DJ, Stevens TJ, Arkin IT, Dupree P (2002) Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis* A genomic analysis. Plant Physiol 129: 486-499

Bouvier M, Moffett S, Loisel TP, Mouillac B, Hebert T, Chidiac P (1995) Palmitoylation of G-protein-coupled receptors: a dynamic modification with functional consequences. Biochem Soc Trans 23: 116-120

Breton G, Danyluk J, Ouellet F, Sarhan F (2000) Biotechnological applications of plant freezing associated proteins. Biotechnol Annu Rev 6: 59-101

Brown DA, London E (2000) Structure and function of sphingolipid- and cholesterol-rich membrane rafts. J Biol Chem 275: 17221-17224

Chung DA, Wade SM, Fowler CB, Woods DD, Abada PB, Mosberg HI, Neubig RR (2002) Mutagenesis and peptide analysis of the DRY motif in the alpha2A adrenergic receptor: evidence for alternate mechanisms in G protein-coupled receptors. Biochem Biophys Res Commun 293: 1233-1241

Colucci G, Apone F, Alyeshmerni N, Chalmers D, Chrispeels MJ (2002) GCR1, the putative *Arabidopsis* G protein-coupled receptor gene is cell cycle-regulated, and its overexpression abolishes seed dormancy and shortens time to flowering. Proc Natl Acad Sci USA 99: 4736-4741

Combet C, Blanchet C, Geourjon C, Deléage G (2000) NPS@: Network Protein Sequence Analysis. Trends Biochem Sci 25: 147-150

Danyluk J (1996) Identification et caractérisation moléculaire de gènes induits au cours de l'acclimatation au froid chez le blé (*Triticum aestivum*). PhD thesis. Université de Montréal, Montréal

Danyluk J, Perron A, Houde M, Limin A, Fowler B, Benhamou N, Sarhan F (1996) Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. Plant Cell 10: 623-638

Devoto A, Piffanelli P, Nilsson I, Wallin E, Panstruga R, von Heijne G, Schulze-Lefert P (1999) Topology, subcellular localization, and sequence diversity of the Mlo family in plants. J Biol Chem 274: 34993-35004

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300: 1005-1016

Fowler DB, Limin AE, Ritchie JT (1999) Low-temperature tolerance in cereals: model and genetic interpretation. Crop Sci 39: 626-633

Frenette Charron JB, Breton G, Badawi M, Sarhan F (2002) Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. FEBS Lett 517: 129-132

Goder V, Spiess M (2001) Topogenesis of membrane proteins: determinants and dynamics. FEBS Lett 504: 87-93

Horn F, Weare J, Beukers MW, Horsch S, Bairoch A, Chen W, Edvardsen O, Campagne F, Vriend G (1998) GPCRDB: an information system for G protein-coupled receptors. Nucleic Acids Res 26: 275-279

Houde M, Danyluk J, Laliberté J-F, Rassart E, Dhindsa RS, Sarhan F (1992) Cloning, characterization and expression of a cDNA encoding a 50-kilodalton protein specifically induced by cold acclimation in wheat. Plant Physiol 99: 1381-1387

Josefsson LG, Rask L (1997) Cloning of a putative G-protein-coupled receptor from *Arabidopsis thaliana*. Eur J Biochem 249: 415-420

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132

Machuka J, Bashiardes S, Ruben E, Spooner K, Cuming A, Knight C, Cove D (1999) Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. Plant Cell Physiol 40: 378-387

Millner PA (2001) Heterotrimeric G-protein in plant cell signaling. New Phytol 151: 165-174

Möller S, Croning MD, Apweiler R (2001) Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 17: 646-653

Nagao M, Minami A, Takezawa D, Arakawa K, Fujikawa S (2001) ABA-induced freezing tolerance in *Physcomitrella patens* and gene expression (abstract no. 354[F455]). Plant Cell Physiol 42: s121

Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14: 897-911

Nielsen H, Krogh A (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. *In* J Glasgow, T Littlejohn, F Major, R Lathrop, D Sankoff, C Sensen, eds, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp 122-130

Pearson WR (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 183: 63-98

Peskan T, Westermann M, Oelmuller R (2000) Identification of low-density Triton X-100-insoluble plasma membrane microdomains in higher plants. Eur J Biochem 267: 6989-6995

Pitcher JA, Freedman NJ, Lefkowitz RJ (1998) G protein-coupled receptor kinases. Annu Rev Biochem 67: 653-692

Plakidou-Dymock S, Dymock D, Hooley R (1998) A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. Curr Biol 8: 315-324

Sansom MS, Weinstein H (2000) Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices. Trends Pharmacol Sci 21: 445-451

Sarhan F, Ouellet F, Vazquez-Tello A (1997) The wheat wcs120 gene family: a useful model to understand the molecular genetics of freezing tolerance in cereals. Physiol Plant 101: 439-445

Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) Constitutively active mutants of the alpha 1-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. EMBO J 15: 3566-3578

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13: 61-72

Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T, et al (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31: 279-292

Simpson GG, Gendall AR, Dean C (1999) When to switch to flowering. Annu Rev Cell Dev Biol 15: 519-550

Steponkus PL (1984) Role of the plasma membrane in freezing injury and cold acclimation. Annu Rev Plant Physiol 35: 543-584

Thomashow MF (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu Rev Plant Physiol Plant Mol Biol 50: 571-599

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680

Tusnàdy GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. Bioinformatics 17: 849-850

Wink M (1997) Special nitrogen metabolism. *In* PM Dey, JB Harbourne, eds, Plant Biochemistry. Academic Press, San Diego, pp 439-486

Yamamoto K, Sasaki T (1997) Large-scale EST sequencing in rice. Plant Mol Biol 35: 135-144

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). Science 296: 79-92

Table I. *Characteristics of cor413-pm, -tm, and moss cor413 genes*

Genes from wheat (Ta), barley (Hv), rice (Os), maize (Zm), sorghum (Sb), Arabidopsis (At), alfalfa (Mt), soybean (Gm), tomato (Le), potato (St), *O. japonica* (Oj), and *P. patens* (Pp) were translated, and the length in amino acids and pI of the longest ORF is presented. For genes sequenced in this study, accession numbers are provided. For sequences constructed from data available in the EST database, one accession number is provided and identified in italic. More information can be found in Supplemental Table VII on the source EST for in silico sequencing and on the Arabidopsis Genome Initiative number for Arabidopsis proteins.

Type	Name	GenBank Accession No.	Amino Acid Residues	pI	Identity with *Ta*COR413-PM1	Identity with *Ta*COR413-TM1
COR413-PM	*Tacor413-pm1*	AAB18207	210	9.0	100	29
	Tacor413-pm2	AA123724	208	9.7	78	27
	Hvcor413-pm1	*BI421687*	210	8.6	95	28
	Hvcor413-pm2	*BF628071*	208	9.3	79	27
	Oscor413-pm1	AI2B3006	210	9.4	71	26
	Zmcor413-pm1	AY181208	212	9.4	72	27
	Sbcor413-pm1	*BI075784*	213	9.1	72	28
	Atcor413-pm1	AI283004	197	9.1	55	26
	Atcor413-pm2	AI2B3005	203	9.4	55	25
	Mtcor413-pm1	*BI003463*	198	8.6	55	26
	Mtcor413-pm2	*BG647116*	199	8.6	55	25
	Mtcor413-pm3	*BG456396*	195	5.2	54	26
	Gmcor413-pm1	*BI211677*	198	9.1	59	27
	Lecor413-pm1	*AW039062*	198	9.2	56	23
	Lecor413-pm2	*BG642925*	202	8.6	61	24
COR413-TM	*Tacor413-tm1*	AY181206	221	10.2	29	100
	Hvcor413-tm1	AI465840	215	10.4	30	88
	Oscor413-tm1	AY181210	222	10.7	27	69
	Zmcor413-tm1	AY181209	226	10.5	27	64
	Atcor413-tm1	AAK76616	226	10.5	25	44
	Atcor413-tm2	AA187293	225	10.6	24	45
	Mtcor413-tm1	*BI639516*	234	10.0	23	41
	Lecor413-tm1	*AW034114*	221	9.8	26	53
	Stcor413-tm1	*BG047649*	220	9.8	26	53
	Ojcor413-tm1	AY181207	241	10.3	27	39
Moss COR413	*Ppcor413-1*	AAL16410	207	5.7	44	26
	Ppcor413-2	***BI200341***	**204**	**7.0**	**44**	**28**
	Ppcor413-3	***BI169989***	**205**	**7.1**	**45**	**10**

Figure 1. Hydropathy and transmembrane predictions. A, Compilation of Kyte and Doolittle profiles of all group COR413-PM and -TM members. S1 to S6, Spikes 1 to 6. B, Compilation of profiles generated by TMHMM 2.0 for all group COR413-PM and -TM members. TMD1 to TMD5, Transmembrane helices 1 to 5. For group COR413-TM, the two profiles were generated without the N-terminal chloroplastic targeting signal (cut after the conserved Cys residue identified in Supplemental Fig. 2).

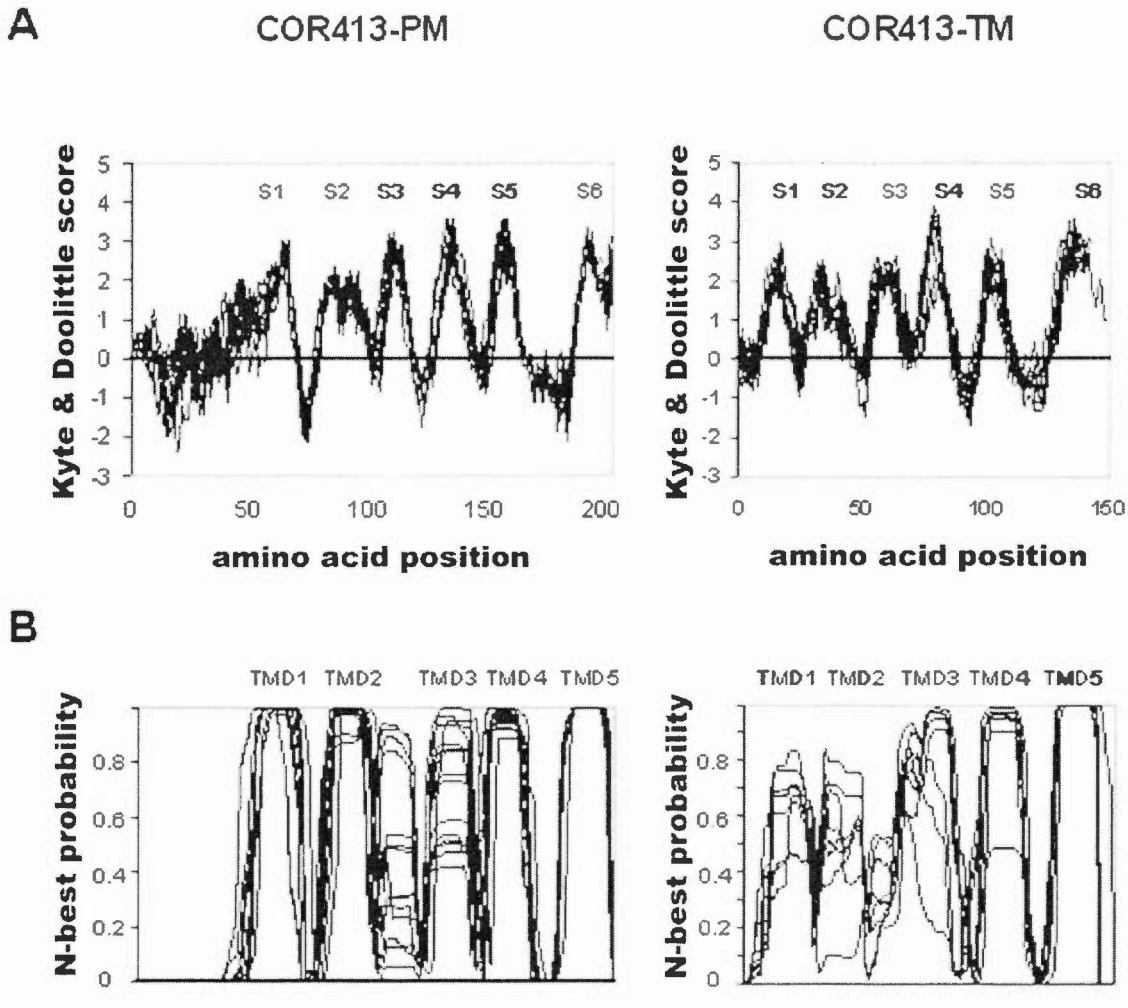
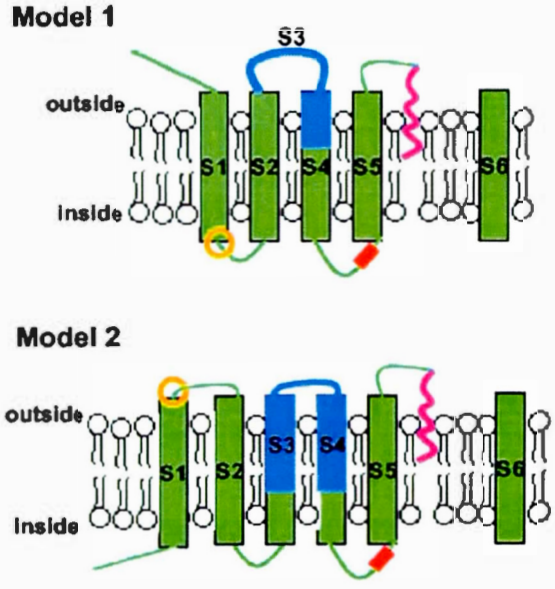
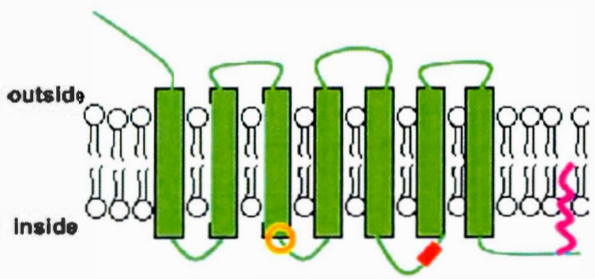

A
COR413-PM
COR413-TM
Kyte & Doolittle score
S1 S2 S3 S4 S5 S6
amino acid position
B
N-best probability
TMD1 TMD2 TMD3 TMD4 TMD5

Figure 2. Proposed models for COR413-PM proteins and comparison with the GPCR Rhodopsin-like family. Green boxes and green lines, TMD and interconnecting loops, respectively.Boxes and lines shaded in blue, Region containing the highest similarity with COR413-TM proteins. S1 to S6 correspond to Kyte and Doolittle spikes from Figure 1. In model I and II, S6 is separated from the rest of the proteins due to the addition of a GPI anchor. Red box, Phosphorylation sites. Pink line, Lipidic anchor. Orange circle, Position of the highly conserved DRT (COR413-PM) or DRY (GPCR) triplet motif.



COR413-PM



G-protein coupled receptor
Rhodopsin-like family

Figure 3. Accumulation of Tacor413-pm and -tm mRNAs during CA in spring and winter wheat. A, Accumulation of Tacor413-pm and -tm mRNAs during CA in spring wheat cv Glenlea and winter wheat cv Norstar. NA7, NA12, nonacclimated control plants grown for 7 and 12 d; CA1, CA6, and CA36, 7-d-old plants were cold-acclimated plants for 1, 6, and 36 d; DA5, cold-acclimated plants (36 d) were deacclimated for 5 d. B, Accumulation of Tacor413-pm and -tm mRNAs during CA in spring and winter wheat cultivars. Total RNA (7.5 μg) from shoots of two spring wheat genotypes (cv Glenlea [Glen], LT50 [lethal temperature that kills 50% of the seedlings] of 8°C; and cv Concorde [Con], LT50 of 8°C), four winter wheat genotypes (cv Monopole [Mon], LT50 of 15°C; cv Absolvent [Abs], LT50 of16°C; cv Fredrick [Fred], LT50 of 16°C; and cv Norstar [Nor], LT50 of 19°C. NA, Nonacclimated plants grown for 13 d; CA36, 7-d-old plants were cold acclimated for 36 d. The 28S ribosomal band stained with ethidium bromide is included to show RNA loads.

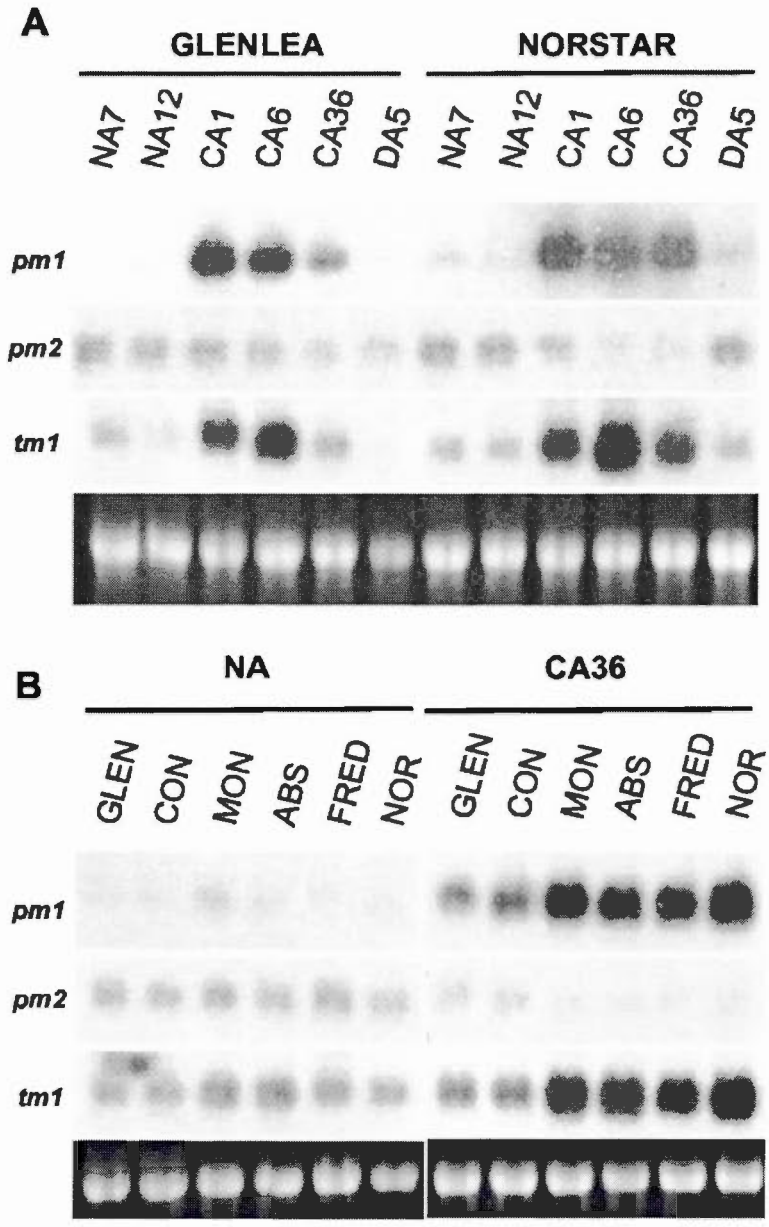

A
GLENLEA
NORSTAR
NA7
NA12
CA1
CA6
CA36
DA5
NA7
NA12
CA1
CA6
CA36
DA5
pm1
pm2
tm1
B
NA
CA36
GLEN
CON
MON
ABS
FRED
NOR
GLEN
CON
MON
ABS
FRED
NOR
pm1
pm2
tm1

Figure 4. Accumulation of *Cor413-pm* and *-tm* mRNAs during CA in cereals and *Arabidopsis*. A, Differential accumulation in various species. In this study, total RNA (7.5 μg) from oat (*Avena sativa* L. cv Laurent, LT_{50} of -6°C), barley (cv Winchester, LT_{50} of -7°C), and winter rye (*Secale cereale* L. cv Musketeer, LT_{50} of -21°C) were used. NA, Nonacclimated plants grown for 13 d; CA36, 7-d-old plants were cold acclimated for 36 d. B, Accumulation in rice (*O. sativa* subsp. *indica* cv IR36). Plants grown for 24 h under the corresponding day/night temperatures in degrees Celsius. C, Accumulation in *Arabidopsis*. NA, Nonacclimated plants grown for 40 d under short photoperiod; CA4h, CA1, CA4, CA7, CA14, and CA21, cold-acclimated plants for 4 h and 1, 4, 7, 14, and 21 d; CTRL, dehydration control plants were removed from pots and placed in water; WS, water-stressed plants were water-stressed by removing them from pots and allowing them to dry for the indicated periods of time; ABA, plants treated with 0.1 mM ABA (Sigma-Aldrich, St. Louis) for 2 and 18 h. The 28S ribosomal band stained with ethidium bromide is included to show RNA loads.

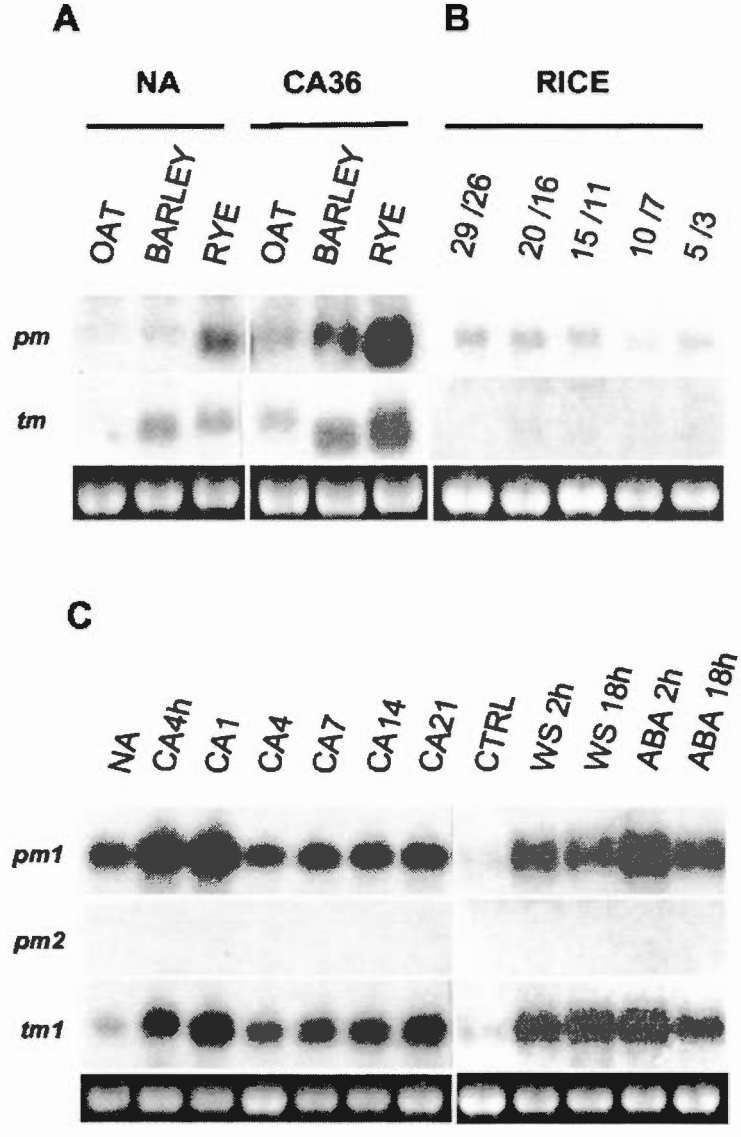

A
B
NA
CA36
RICE
OAT
BARLEY
RYE
OAT
BARLEY
RYE
29 /26
20 /16
15 /11
10 /7
5 /3
pm
tm
C
NA
CA4h
CA1
CA4
CA7
CA14
CA21
CTRL
WS 2h
WS 18h
ABA 2h
ABA 18h
pm1
pm2
tm1

Figure 5. Accumulation of *Tacor413-pm* and *-tm* mRNAs under stress conditions and tissue specificity in winter wheat cv Norstar. A, Accumulation of *Tacor413-pm* and *-tm* mRNAs under different stress conditions. NA, Nonacclimated plants grown for 7 d; CA1, plants cold acclimated for 1 d; HS1 and 3 h, plants exposed to 40°C for 1 and 3 h (heat shock); ABA18h, plants treated with 0.1 mM ABA (Sigma-Aldrich) for 18 h; NaCl, plants treated with 300 mM NaCl for 18 h; DHN CTRL, dehydration control plants grown for 7 d; DHN30% and 70%, water stressed plants with a relative water content of 30% and 70%. B, Tissue specificity in winter wheat cv Norstar. Leaf, crown, and roots of nonacclimated plants (NA) and 6-d cold-acclimated (CA6) plants. C, *Tacor413-tm1* expression is dependent on the chloroplast differentiation stage. NA, Nonacclimated plants grown for 7 d in the presence of a light cycle (8 h of light:16 h of dark; green) or in the dark (etiolated); NA+L, 7-d-old etiolated plants after one light cycle; CA, 7-d-old green or etiolated plants cold acclimated for 24 h in the presence of light (L) or in the dark (D). The 28S ribosomal band stained with ethidium bromide is included to show RNA loads (7.5 μg).

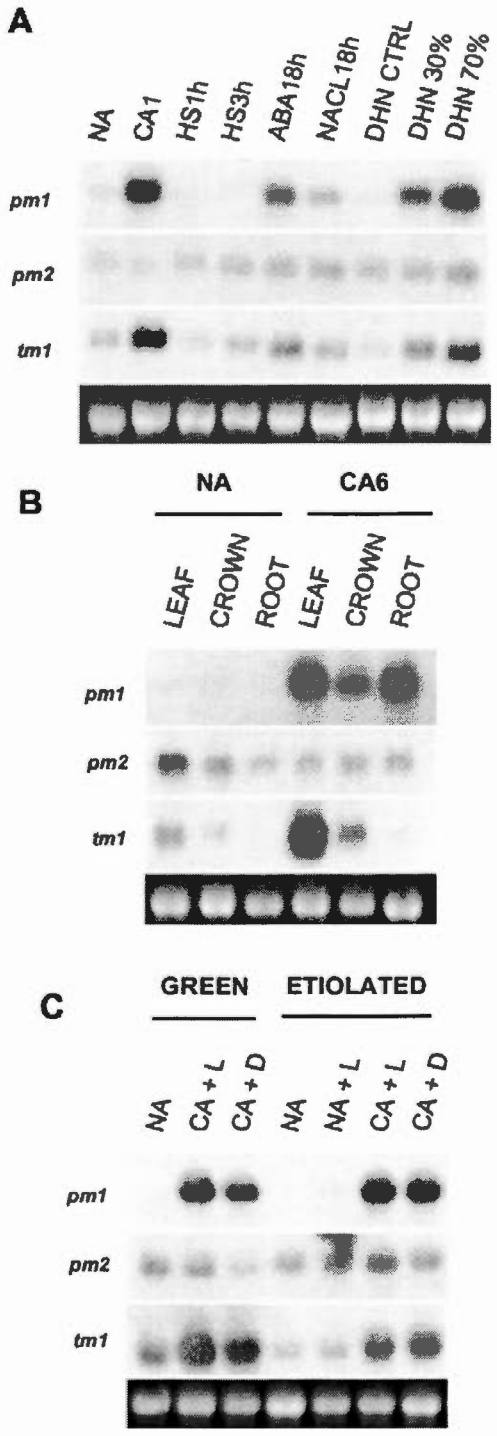

A
NA
CA1
HS1h
HS3h
ABA18h
NACL18h
DHN CTRL
DHN 30%
DHN 70%
pm1
pm2
tm1
B
NA
CA6
LEAF
CROWN
ROOT
LEAF
CROWN
ROOT
pm1
pm2
tm1
C
GREEN
ETIOLATED
NA
CA + L
CA + D
NA
NA + L
CA + L
CA + D
pm1
pm2
tm1

Table II (Supplemental data). COR413 related entries in Genbank EST database.

The tissue from which RNA was extracted is listed when known. Special treatments (hormone, temperature, pathogens) are not listed since it is not a sign of specific expression.

Species	No of Orthologs	Genbank acc. number	Tissue
P. patens	21 cor413-1/14834	AAL16410	protonemata (7d)
	20 cor413-2	BJ200341	n.m.[a]
	20 cor413-3	BJ169989	n.m.
	1 cor413-4	BJ169579	n.m.
M. polymorpha	1 cor413-1/ 1507	C96082	immature sex organ (female)
C. japonica	2 tm1 / 2495	AY181207	inner bark
P. taeda	8 pm1 / 41696	BG040065	pollen cone (immature)(1); xylem (6), xylem (secondary)(1)
	5 tm1	BF517115	shoot tip (2), xylem side wood (2); xylem compression wood (1)
G. max	46 pm1 / 233557 with 11 different splice site variants	BE211677	germinating shoots (2-3d)(3), hypocotyl and plumule (3d)(2) cotyledons (3-7d) (4) roots (7-8d)(2), etiolated hypocotyl (9-10d)(3); cotyledons (11d)(1) degenerating cotyledons (2w)(2); leaf and shoot tip (2w)(1) leaves (2-4w)(4); stem tissue (4w)(3) immature flower (2); mature flower (1); immature seed coats (1) mature seed pods (2); roots (3) whole seedlings (1-3w)(11) ; n.m. (1)
	1 tm1	BF423820	floral meristematic
G. clandestina	2 pm1 / 944	BG838527	leaves and stem
M. truncatula	37pm1/140818	BF003463	cotyledons + first leaves (3); roots (3-4d)(5) developing leaves (4); root cell culture (4), leaves (4); nodules (4); arbuscular mycorrhiza (10); n.m (3)
	7 pm2	BG647116	root (3); cell culture (roots)(1); arbuscular mycorrhiza (2), developping leaves (1)
	9 pm3	BG456396	arbuscular mycorrhiza (2), root tips (2); n.m. (5)
	8 tm1	BF639516	cotyledons and primary leaves (1); root (2) developing flower(1);leaves(1),n.m. (3)
L. japonicus	4 pm1 / 31600	AI967693	whole roots (2) young plants (2w)(2)
P. tremulus X P. tremuloides	6 pm1 / 20137	AI161923	cambial region -1.5 m actively growing tree (6)
	1 tm1	AI162978	cambial region
G. arboreum	1 pm1 / 24032	BG440267	7-10d post-anthesis fiber bolls
	1 pm2	BG444068	7-10d post-anthesis fiber bolls
G. hirsutum	2 pm1 12088	AI727568	fiber (6d)(2)
T. salsuginea	2 pm1/681	BE758599	n.m.

	1 tm1	BI698744	aerial part
A. thaliana	49 pm1/ 219653	AF283004	leaves + roots (2-3w)(1); seedling (2-6w)(6); flower buds (3); roots (6); young flower (1); green silique (2); hypocotyl (3d)(1); suspension (1); leaves (12w)(6); n.m. (22)
	11 pm2	AF283005	etiolated (5d)(1); seedlings (2-6w)(2); flower buds (1); green silique (1); n.m. (6)
	1 pm3	AY063884	n.m.
	0 pm4 (chr4)	Z99707	no hit found
	5 tm1	AAK76616	flower buds (1), n.m. (4)
	1 tm2	AAL87293	rosette (4-7w)
L. esculentum	22 pm1 /155304	AW039062	fruit mature green (3-5d pre-ripening)(1); breaker fruits (pericarp (4); fruit red ripe (pericarp)(4); callus (25-40d)(5); leaf 4w (2); suspension (1); ovary carpel 5d pre-anthesis to 5d post-anthesis (1); root pre-anthesis (1); n.m. (3)
	3 pm2	BG642925	flower buds and open flowers (1); shoot meristem (4-6w)(1); n.m. (1)
	6 tm1	AW034114	suspension (1); callus (1); mature green fruit (1), radicle (5d)(1); n.m. (2)
S. tuberosum	4 tm1 / 62501	BG593744	leaves and petiole (6-8w); sprouting eyes (12-14w post harvest)
M. crystallinum	1 pm1 / 16925	BE033418	roots
	1 pm2	BE037100	apical meristem and leaf primordia (6w)
	1 pm3	BG269058	n.m.
	1 tm1	AW265829	leaves (6w)
O. sativa	12 pm1 / 206242	AF283006	panicle flowering stage (4); immature leaf (2); shoot etiolated (8d)(2); shoot (1); root (1); young root (1); immature seed (1)
	2 tm1	AY181210	immature seed (5d after pollination); panicle at ripening stage
S. bicolor	21 pm1 / 87226	BI075784	immature panicle (pre-anthesis)(6); etiolated (5d)(5); seedling (10-14d)(1) plants (5w) (9)
	3 tm1	BG047649	ovary (mix of ovaries of varying immature stages) (8w)
S. propinquum	3 pm1 / 21388	BG101906	rhizome
Z. mays	13 pm1 / 191177	AY181208	4th leaf (1); ear tissue (3); tassel primordium (3); glume (1); anther + pollen (1); n.m. (4)
	6 pm2	AW438128	n.m.
	3 tm1	AW787650	mixed adult tissues (3)
H. vulgare	40 pm1 / 126972	BE421687	leaves (1); leaves (14d) (1); shoot (8); first leaves (22); roots (2d) (3); n.m. (5)
	9 pm2	BF628071	pre-anthesis spike (white to yellow anther)(1); caryopsis 3-15 DAP (1); etiolated roots (1); testa/pericarp from developing kernel (1); leaves (1); etiolated leaves (2); n.m. (2)
	6 tm1	AF465840	shoot (5d)(1); testa-pericarp from developing kernel (1); root (5d)

			etiolated tissue (1) developing caryopsis 3-15 DAP (2) leaves (1)
T. aestivum	9 pm1 / 72258	AAB18207	leaves (14d)(3); pre-anthesis spike (1), seedlings (5d)(2); crown (5w)(2)
			n m. (1)
	6 pm2	AAL23724	leaf (14d)(2), pre-anthesis spike (1); spike (1); n m (2)
	2 tm1	AY181206	young spikelet flag leaf
T. turgidum	4 pm1 / 1643	BE428950	roots (3d) (4)
	2 pm2	BE428224	root
T.monococcum	2 pm1 / 5655	BG607123	early reproductive apex - double ridge stage to terminal spikelet stage
			(7w) (2)

[a] n.m.: not mentioned.

Table III (Supplemental data). Targeting signals and structure predictions of the COR413-PM protein family.

Name	Targeting localization and score[a]	Signal type and score[b]	Topology prediction[c]
*Ta*COR413-PM1	c.t.m. (0.752), p.m (0.600)	signal anchor (0.957)	6 TMD N-in C-in
*Ta*COR413- PM2	c.t.m. (0.866); p.m (0.600)	signal anchor (0.003)	6TMD N-in C-in
*Hv*COR413- PM1	c.t.m. (0.844), p.m (0.600)	signal anchor (0.955)	6TMD N-in C-in
*Hv*COR413- PM2	c.t.m. (0.849), p.m. (0.600)	signal anchor (0.004)	6TMD N-in C-in
*Os*COR413- PM1	p.m. (0.600), c.t.m. (0.409)	signal anchor (0.007)	5TMD N-out C-in
*Zm*COR413- PM1	p.m. (0.600) g.b. (0.400)	signal anchor (0.007)	5TMD N-out C-in
*Sb*COR413- PM1	p.m. (0.600); c.t.m. (0.413)	signal anchor (0.010)	5TMD N-out C-in
*At*COR413- PM1	p.m. (0.600), c.t.m. (0.487)	signal anchor (0.059)	5TMD N-out C-in
*At*COR413- PM2	p.m. (0.600); c.t.m. (0.441)	signal anchor (0.009)	5TMD N-out C-in
*Mt*COR413- PM1	p.m. (0.600), c.t.m. (0.402)	signal anchor (0.681)	5TMD N-out C-in
*Mt*COR413- PM2	p.m. (0.600); c.t.m. (0.506)	signal anchor (0.857)	5TMD N-out C-in
*Mt*COR413- PM3	p.m. (0.600), c.t.m. (0.417)	signal anchor (0.631)	5TMD N-out C-in
*Gm*COR413- PM1	c.t.m. (0.619), p.m. (0.600)	signal anchor (0.991)	5TMD N-out C-in
*Le*COR413- PM1	p.m. (0.600), c.t.m. (0.447)	signal anchor (0.067)	5TMD N-out C-in
*Le*COR413-PM2	p.m. (0.600), c.t.m. (0.441)	signal anchor (0.044)	5TMD N-out, C-in
Mean	p.m. (0.600), c.t.m. (0.551)	signal anchor (0.350)	5TMD N-out 11/15
		signal peptide (0.06)	C-in 15/15

[a] PSORT was used to predict the cellular localization p.m.: plasma membrane, c.t.m.: chloroplast thylakoid membrane, g.b.: golgi bodies

[b] SignalP was used to predict the presence of a signal peptide or anchor

[c] TMHMM was used for topology prediction N-in: N-terminal inside, N-out: N-terminal outside, C-in: C-terminal inside, C-out: C-terminal outside

Table IV (Supplemental data). Targeting signals and structure predictions of the Moss COR413 protein family.

Name	Targeting localization and score[a]	Signal type and score[b]	Topology prediction[c]
*Pp*COR413-1	p.m (0.600), c.t.m (0.434)	signal anchor (0.149)	5TMD N-out, C-in
*Pp*COR413-2	p.m (0.600), c.t.m (0.504)	signal anchor (0.935)	5TMD N-out; C-in
*Pp*COR413-3	p.m. (0.600); m.i.m (0.511)	signal anchor; (0.291)	4TMD N-out; C-out
Mean	p.m (0.600)	signal anchor (0.460)	5TMD N-out, C-in
		signal peptide (0.130)	

[a] PSORT was used to predict the cellular localization p.m plasma membrane c.t.m. chloroplast thylakoid membrane, m.i.m. mitochondrial inner membrane

[b] SignalP was used to predict the presence of a signal peptide or anchor

[c] TMHMM was used for topology prediction N-in N-terminal inside N-out N-terminal outside, C-in C-terminal inside C-out C-terminal outside

Table V. (Supplemental data). Targeting signals and structure predictions of the COR413-TM protein family.

Name	Targeting localization and score			Topology prediction[b]
	PSORT[a]	iPSORT	TargetP	
*At*COR413-TM1	c.r. (0.685); p.m. (0.640)	chloroplast	chloro (0.932)	2TMD N-in; 5TMD N-in
*Hv*COR413-TM1	p.m. (0.600); c.t.m. (0.496)	chloroplast	chloro (0.774)	2TMD N-in, 4TMD N-out
*Os*COR413-TM1	p.m. (0.600); c.t.m. (0.519)	mitochondrion	chloro (0.878)	4TMD N-out, 5TMD N-out
*Sm*COR413-TM1	m.i.m. (0.860); c.t.m. (0.762)	mitochondrion	chloro (0.766)	4TMD N-in; 4TMD N-in
*At*COR413-TM1	c.t.m. (0.950); p.m. (0.600)	chloroplast	chloro (0.840)	4TMD N-in, 4TMD N-in
*At*COR413-TM2	m.i.m. (0.826); c.t.m. (0.730)	chloroplast	chloro (0.843)	4TMD N-in, 4TMD N-in
*Os*COR413-TM1	c.t.m. (0.972); m.i.m. (0.900)	chloroplast	chloro (0.956)	3TMD N-out; 4TMD N-in
*Zm*COR413-TM1	c.t.m. (0.938); p.m. (0.600)	chloroplast	chloro (0.615)	3TMD N-out; 4TMD N-in
*Sm*COR413-TM1	c.t.m. (0.950); p.m. (0.600)	chloroplast	chloro (0.849)	3TMD N-out, 5TMD N-out
*Tt*COR413-TM1	c.t.m. (0.936); m.i.m. (0.875)	chloroplast	chloro (0.833)	2TMD N-in; 1TMD N-in
mean	c.t.m. (0.720); p.m. (0.600)	chloroplast	chloro (0.830)	-

[a] p.m.: plasma membrane; c.t.m.: chloroplast thylakoid membrane; m.i.m.: mitochondrial inner membrane; e.r.: endoplasmic reticulum

[b] TMHMM was used for topology prediction N-in: N-terminal inside, N-out: N-terminal outside, C-in: C-terminal inside, C-out: C-terminal outside.

Table VI (Supplemental data). List of plant molecules showing structural similarities to animal GPCR ligands

PCR Ligand	Plant Molecules
Serotonin	-Tryptamine
	-N-methyltryptamine
	-N,N-dimethyltryptamine
	-5-Methoxy-N,N-dimethylserotonin
	-Gramine
	-Auxins
Dopamine	Tyramine
Adrenaline	N-methyl tyramine
Noradrenaline	Hordenine
Histamine	Histamine
Retinal	Abscisic acid
Acetylcholine	Acetylcholine

Table VII (Supplemental data). Fasta files of COR413 sequences and predicted open reading frames.

Group I COR413

>*Ta*COR413-PM1 clone wcor413 sequenced and submitted to Genbank (AAB18207).

>*Ta*COR413-PM2 clone WHE0316_D08_D08 sequenced and submitted to Genbank (AAL23724).

>*Hv*COR413-PM1 deduced from ESTs BE421687 and AL502741.

>*Hv*COR413-PM2 deduced from ESTs BG343566 and BF628071.

>*Os*COR413-PM1. clone S1249 sequenced and submitted to Genbank (AF283006).

>*Zm*COR413-PM1. clone Zm03_05a03 sequenced and submitted to Genbank (AY181208).

>*Sb*COR413-PM1 deduced from ESTs BI075784 and AW680600.

>*At*COR413-PM1 clone l15O23 sequenced and submitted to Genbank (AF283004) At2g15970.

>*At*COR413-PM2 clone 90M2 sequenced and submitted to Genbank (AF283005) At3g50830.

>*Mt*COR413-PM1 deduced from EST BF003463.

>*Mt*COR413-PM2 deduced from EST BG647116.

>*Mt*COR413-PM3 deduced from ESTs BG456396 and AL384664.

>*Gm*COR413-PM1 deduced from ESTs BE211677 and AW309837.

>*Le*COR413-PM1 deduced from EST AW039062.

>*Le*COR413-PM2 deduced from ESTs BG642925, AW041686 and BG629922.

Group I not used for analysis

>*At*COR413-PM4 from genomic sequence (Z99707) At4g37220.

>*At*COR413-PM3 from Genbank AY063884 and corrected with the genomic sequence AC004482 At2g23680.

Group II COR413

>*Ta*COR413-TM1 clone WHE0316_D08_D08 sequenced and submitted to Genbank (AY181206).

>*Hv*COR413-TM1. clone HVSMEb0007K15f sequenced and submitted to Genbank (AF465840).

>*Os*COR413-TM1 clone E10543_6Z sequenced and submitted to Genbank (AY181210).

>ZmCOR413-TM1 clone Zm03_01b06 sequenced and submitted to Genbank (AY181209).

>AtCOR413-TM1 clone 97C22 sequenced and shown identical to AAK76616, At1g29395.

>AtCOR413-TM2 from RIKEN AAL87293 and completed with NP_564327 At1g29390.

>StCOR413-TM1 deduced from ESTs BM109760, BM109832 and BG 596886.

>LeCOR413-TM1 deduced from ESTs AW034114, AW443083 and AW041036.

>MtCOR413-TM1 deduced from ESTs BF639516 and BG588918.

>CjCOR413-TM1. clones CC1364 and CC0164 sequenced and submitted to Genbank (AY181207).

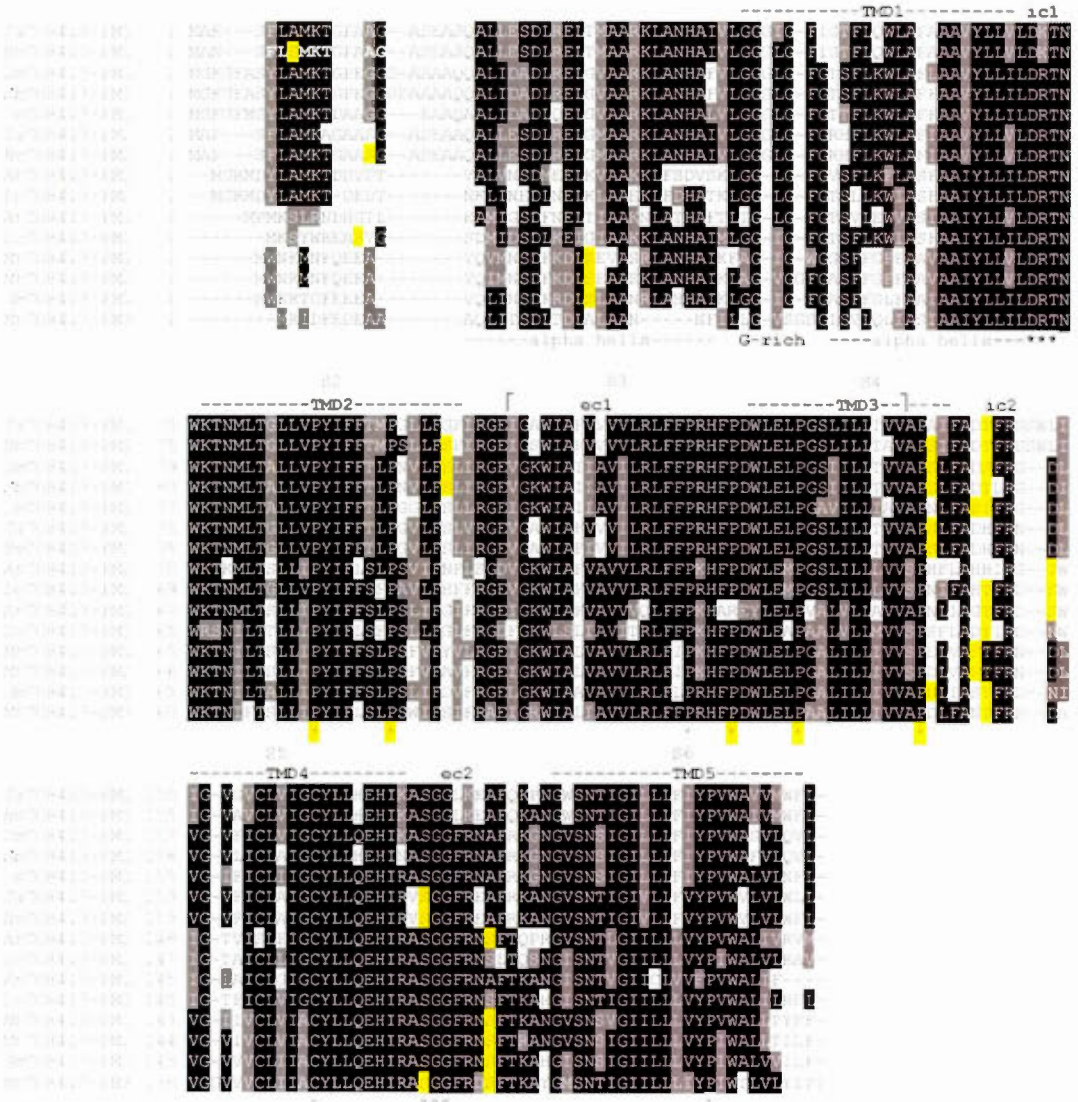
<u>Moss COR413</u>

>PpCOR413-1 clone PPU030403 sequenced and submitted to Genbank (AAL16410).

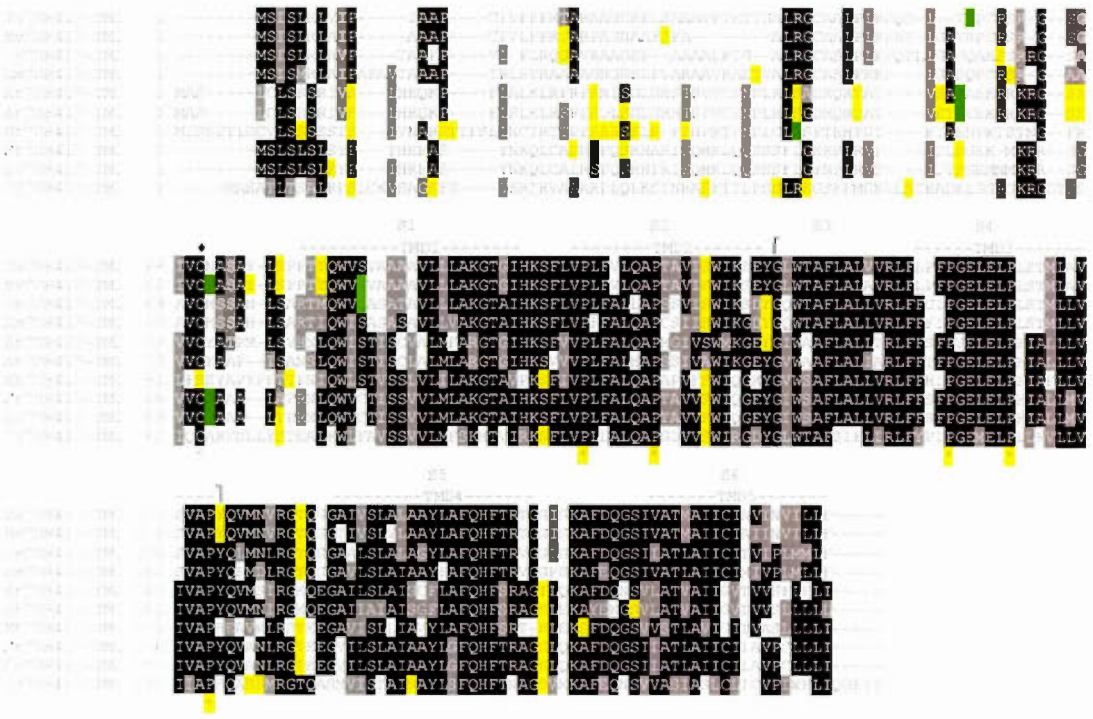
>PpCOR413-2 deduced from ESTs BJ200341, BJ167268 and BI487528.

>PpCOR413-3 deduced from ESTs BJ169989 and BJ203803.

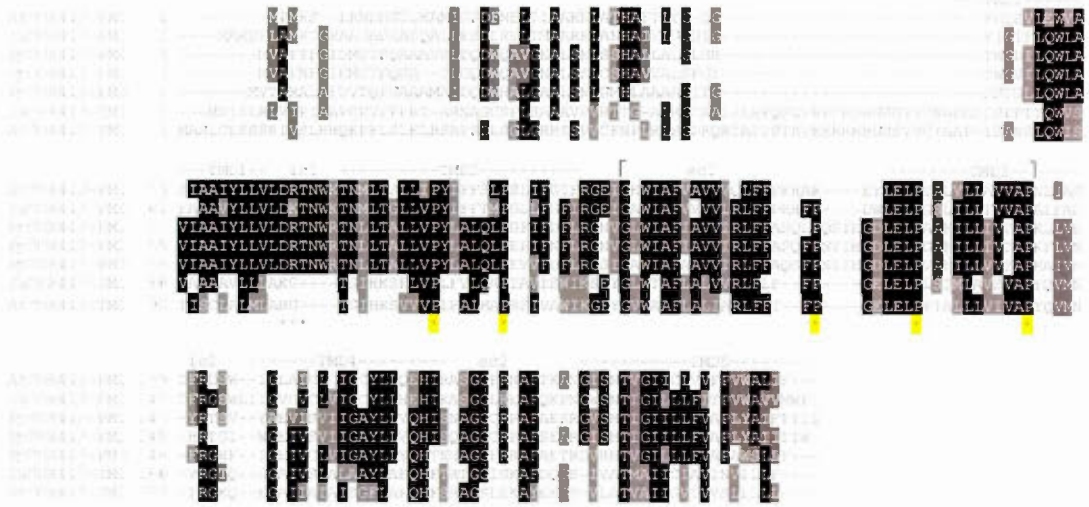
Supplemental Figure 1. Aligment and structural features of COR413-PM proteins. COR413-PM members from Table I were aligned using CLUSTALW. Identical and similar amino acids are shaded in black and gray, respectively. S1 to S6: Spikes identified in the Kyte and Doolittle profile (Fig.4). TMD1 to TMD5: Transmembrane domains identified by TMHMM 2.0.ic1 and ic2: intracellular loop 1 and 2. ec1 and ec2: extracellular loop 1 and 2. Yellow shaded letters: Putative phosphorylation sites identified by NetPhos 2.0. ###: indicates putative cleavage site for GPI anchor identified by DGPI. *: indicates proline and cysteine residues with potential structural function. Yellow shaded asterics: proline residues conserved between COR413-PM and –TM. ***: indicates putative GPCR domain. Brackets indicate the region of highest homology between COR413-PM and –TM.


TMD1
ic1
G-rich
TMD2
ec1
TMD3
ic2
TMD4
ec2
TMD5

Supplemental Figure 2. Aligment and structural features of COR413-TM proteins.COR413-PM members from Table I were aligned using CLUSTALW. Identical and similar amino acids are shaded in black and gray, respectively. : indicates the first amino acid of the mature protein that was used for Kyte and Dolittle and TMHMM analyses. S1 to S6: Spikes identified in the Kyte and Doolittle profile (Fig.4). TMD1 to TMD5: Transmembrane domains identified by TMHMM. Yellow shaded letters: Putative phosphorylation sites identified by NetPhos 2.0. Green shaded letters: Putative chloroplastic cleavage site identified by TargetP. *: indicates proline and cysteine residues with potential structural function. Yellow shaded asterics: proline residues conserved between COR413-PM and –TM. Brackets indicate the region of highest homology between COR413-PM and –TM.



Supplemental Figure 3. Aligment and structural features of COR413-PM, COR413-TM and Moss COR413 proteins. Selected COR413 members from wheat, Arabidopsis and Moss were aligned using CLUSTALW. Identical and similar amino acids are shaded in black and gray, respectively. TMD1 to TMD5: Transmembrane domains identified by TMHMM. *: indicates proline residues with potential structural function. Brackets indicate the region of highest homology between group I and group II COR413. ic1 and ic2: intracellular loop 1 and 2. ec1 and ec2: extracellular loop 1 and 2. ***: indicates putative GPCR domain.



BIBLIOGRAPHIE GÉNÉRALE

Agarraberes, F.A. et Dice, J.F. (2001) Protein translocation across membranes. Biochim. Biophys. Acta 1513: 1-24.

Alabadí, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Más, P. et Kay, S.A. (2001) Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. Science 293: 880-883.

Alabadí, D., Yanovsky, M.J., Más, P., Harmer, S.L. et Kay, S.A. (2002) Critical role for CCA1 and LHY in maintaining circadian rhythmicity in *Arabidopsis*. Curr. Biol. 12: 757-761.

Alewijnse, A.E., Timmerman, H., Jacobs, E.H., Smit, M.J., Roovers, E., Cotecchia, S. et Leurs, R. (2000) The effect of mutations in the DRY motif on the constitutive activity and structural instability of the histamine H(2) receptor. Mol. Pharmacol. 57: 890-898.

Allhorn, M., Lundqvist, K., Schmidtchen, A. et Åkerström, B. (2003) Heme-scavenging role of alpha1-microglobulin in chronic ulcers. J Invest Dermatol. 121: 640-646.

Allhorn, M., Klapyta, A. et Åkerström B. (2005) Redox properties of the lipocalin alpha1-microglobulin: reduction of cytochrome c, hemoglobin, and free iron. Free Radic Biol Med. 38: 557-567

Allard, F., Houde, M., Kröl, M., Ivanov, A., Huner, N.P.A. et Sarhan, F. (1998) Betaine improves freezing tolerance in wheat. Plant Cell Physiol. 39: 1194-1202.

Allen, R.D., Webb, R.P. et Schake, S.A. (1997) Use of transgenic plants to study antioxidant defenses. Free Radic. Biol. Med. 23: 473–479.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C.C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D.E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W.L., Berry, C.C. et Ecker, J.R. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301: 653-657.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. et Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Altschul, S.F., Madden, T.L., Schäfer, A.A., Zhang, J., Zhang, Z., Miller, W. et Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3381-3402.

Åkerström, B., Flower, D.R. et Salier, J.P. (2000) Lipocalins: unity in diversity. Biochim. Biophys. Acta 1482: 1-8.

Andreasson, E., Jenkins, T., Brodersen, P., Thorgrimsen, S., Petersen, N.H., Zhu, S., Qiu, J.L., Micheelsen, P., Rocher, A., Petersen, M., Newman, M.A., Bjørn Nielsen, H., Hirt, H., Somssich, I., Mattsson, O. et Mundy, J. (2005) The MAP kinase substrate MKS1 is a regulator of plant defense responses. EMBO J. 24: 2579-2589.

Apel, K. et Hirt, H. (2004) Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55: 373–399.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.

Artus, N.N., Uemura, M., Steponkus, P.L., Gilmour, S.J., Lin, C. et Thomashow, M.F. (1996) Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. Proc. Natl. Acad. Sci. USA 93: 13404-13409.

Asada, K. (1999) The water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. Ann. Rev. Plant Physiol. Plant Mol. Biol. 50: 601–639.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K. et Miyano, S. (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18: 298-305.

Bargiello, T.A. et Young, M.W. (1984) Molecular genetics of a biological clock in *Drosophila*. Proc. Natl. Acad. Sci. USA. 81: 2142-2146.

Bartels, D. et Sunkar, R. (2005) Drought and salt tolerance in plants. CRC Crit. Rev. Plant Sci. 24: 23–58.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Rapp, B.A. et Wheeler, D.L. (2002) Genbank. Nucleic Acids Res. 30: 17-20.

Bishop, R.E., Penfold, S.S., Frost, L.S., Holtje, J.V. et Weiner, J.H. (1995) Stationary phase expression of a novel *Escherichia coli* outer membrane lipoprotein and its relationship with mammalian apolipoprotein D. Implications for the origin of lipocalins. J. Biol. Chem. 270: 23097-23103.

Bishop, R.E. (2000) The bacterial lipocalins. Biochim. Biophys. Acta 1482: 73-83.

Blanco-Vaca, F., Via, D.P., Yang, C.Y., Massey, J.B. et Pownall, H.J. (1992) Characterization of disulfide-linked heterodimers containing apolipoprotein D in human plasma lipoproteins. J. Lipid Res. 12: 1785-1796.

Blom, N., Gammeltoft, S. et Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294: 1351-1362 .

Bockaert, J. et Pin, J.P. (1999) Molecular tinkering of G protein-coupled receptors: an evolutionary success. EMBO J. 18: 1723-1729.

Bohnert, H. et Shen, B. (1999) Transformation and compatible solutes. Sci. Hortic. 78: 237-260.

Bold, H.C. et Wynne, M.J. (1985) Introduction to the algae. 2^{e} édition, Prentice Hall, Inc., Englewood Cliffs, NJ. 702 pp.

Borner, G.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T. et Dupree, P. (2002) Prediction of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis* A genomic analysis. Plant Physiol, 129: 486-499.

Bouché, N., Fait, A., Bouchez, D., Moller, S.G. et Fromm, H. (2003) Mitochondrial succinic-semialdehyde dehydrogenase of the gamma-aminobutyrate shunt is required to restrict levels of reactive oxygen intermediates in plants. Proc. Natl. Acad. Sci. USA. 100: 6843-6848.

Bouvier, M., Moffett, S., Loisel, T.P., Mouillac, B., Hebert, T. et Chidiac, P. (1995) Palmitoylation of G-protein-coupled receptors: a dynamic modification with functional consequences. Biochem. Soc. Trans. 23: 116-120.

Bowler, C., Slooten, L., Vandenbranden, S., De Rycke, R., Botterman, J., Sybesma, C., Van Montagu, M. et Inzé, D. (1991) Manganese superoxide dismutase can reduce cellular damage mediated by oxygen radicals in transgenic plants. EMBO J. 10: 1723–1732.

Bown, A.W. et Shelp, B.J. (1997) The Metabolism and Functions of [gamma]-Aminobutyric Acid. Plant Physiol. 115: 1-5.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R. et Gorlach, J. (2001) Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. Plant Cell 13: 1499-1510.

Breton G, Danyluk J, Ouellet F, Sarhan F. (2000). Biotechnological applications of plant freezing associated proteins. Biotechnol. Annu. Rev. 6: 57-99.

Breton, G., Danyluk, J., Charron, J.B.F. et Sarhan, F. (2003) Expression profiling and bioinformatic analyses of a novel stress-regulated multispanning transmembrane protein family from cereals and *Arabidopsis*. Plant Physiol. 132: 64-74.

Broadbent, P., Creissen, G.P., Kular, B., Wellburn, A.R. et Mullineaux, P.M. (1995) Oxidative stress responses in transgenic tobacco containing altered levels of glutathione reductase activity. Plant J. 8: 247-255.

Brown, D.A. et London, E. (2000) Structure and function of sphingolipid- and cholesterol-rich membrane rafts. J. Biol. Chem. 275: 17221-17224.

Bugos, R.C., Hieber, A.D. et Yamamoto, H.Y. (1998) Xanthophyll cycle enzymes are members of the lipocalin family, the first identified from plants. J. Biol. Chem. 273: 15321-15324.

Burke, B.J., Redondo, C., Redl, B. et Findlay, J.B.C. (2006) Lipocalin Receptors: Into the Spotlight. *Dans* Lipocalins. *Éditeurs*: B. Åkerström, N. Borregaard, D.R. Flower, et J.P. Salier. Landes Bioscience. Georgetown, TX. pp.167-176.

Burke, M.J., Gusta, L.V., Quamme, H.A., Weiser, C.J. et Li, P.H. (1976) Freezing and injury in plants. Ann. Rev. Plant Physiol. 27: 507-528.

Cadieux, C., Sarhan, F. et Perras, M. (1988) Osmotic adjustement and photosynthetic electron transport response to cold hardening in winter and spring wheat. Plant Physiol. Biochem. 26: 313-322.

Castonguay, Y. et Guckert, A. (1996) Adaptation of forage legumes to cold climates. Grassland and Land use systems. EGF Meetings contributions 532: 911-917.

Chakraborti, T., Das, S., Mondal, M., Roychoudhury, S. et Chakraborti, S. (1999) Oxidant, mitochondria and calcium: an overview. Cell. Signal. 11: 77–85.

Charron, J.B., Breton, G., Danyluk, J., Muzac, I., Ibrahim, R.K. et Sarhan, F. (2002) Molecular and biochemical characterization of a cold-regulated phosphoethanolamine Nmethyltransferase from wheat. Plant Physiol. 129: 363-373.

Charron, J.B., Ouellet, F., Pelletier, M., Danyluk, J., Chauve, C. et Sarhan, F. (2005) Identification, expression, and evolutionary analyses of plant lipocalins. Plant Physiol. 139: 2017-2028.

Charron, J.B.F. et Sarhan, F. (2006) Plant lipocalins *Dans* Lipocalins. *Éditeurs*: B. Åkerström, N. Borregaard, D.R. Flower, et J.P. Salier. Landes Bioscience. Georgetown, TX. Pp.41-48.

Chung, D.A., Wade, S.M., Fowler, C.B., Woods, D.D., Abada, P.B., Mosberg, H.I. et Neubig, R.R. (2002) Mutagenesis and peptide analysis of the DRY motif in the alpha2A adrenergic receptor: evidence for alternate mechanisms in G protein-coupled receptors. Biochem. Biophys. Res. Commun. 293: 1233-1241.

Clough, S.J. et Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. 16: 735-743.

Clouse, S.D. et Sasse, J.M. (1998) Brassinosteroids: Essential regulators of plant growth and development. Annu. Rev. Plant Physiol. Plant Mol. 49: 427-451.

Colucci, G., Apone, F., Alyeshmerni, N., Chalmers, D. et Chrispeels, M.J. (2002) GCR1, the putative *Arabidopsis* G protein-coupled receptor gene is cell cycle-regulated, and its overexpression abolishes seed dormancy and shortens time to flowering. Proc. Natl. Acad. Sci. USA 99: 4736-4741.

Combet, C., Blanchet, C., Geourjon, C., Deléage, G. (2000) NPS@: Network Protein Sequence Analysis. Trends Biochem. Sci. 25: 147-150.

Crowe, J.H., Hoekstra, F.A. et Crowe, L.M. (1992) Anhydrobiosis. Ann. Rev. Plant Physiol. 54: 579-599.

Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. et Barton, G.J. (1998) JPred: a consensus secondary structure prediction server. Bioinformatics 14, 892-893.

Danyluk, J. (1996) Identification et caractérisation moléculaire de gènes induits au cours de l'acclimatation au froid chez le blé (*Triticum aestivum*). Thèse de doctorat. Université de Montréal, Montréal

Danyluk, J. et Sarhan, F. (1990) Differential mRNA transcription during the induction of freezing tolerance in spring and winter wheat. Plant Cell Physiol. 31: 609-619.

Danyluk, J., Rassart, E. et Sarhan, F. (1991) Gene expression during cold and heat shock in wheat. Biochem. Cell Biol. 69: 383-391.

Danyluk, J., Houde, M., Rassart, E. et Sarhan, F. (1994) Differential expression of a gene encoding an acidic dehydrin in chilling sensitive and freezing tolerant gramineae species. FEBS Lett. 344: 20-24.

Danyluk, J., Carpentier, E. et Sarhan, F. (1996) Identification and characterization of a low temperature regulated gene encoding an actin-binding protein from wheat. FEBS Lett. 389: 324-327.

Danyluk, J., Perron, A., Houde, M., Limin, A., Fowler, B., Benhamou, N. et Sarhan, F. (1996) Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. Plant Cell 10: 623-638.

Delwiche, C.F., Kuhsel, M. et Palmer, J.D. (1995) Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. Mol. Phylogenet. Evol. 4: 110-128.

Demel, R.A. and De Kruijff, B. (1976) The function of sterols in membranes.Biochim. Biophys. Acta 457: 109-132.

Demel, R.A., Dorrepaal, E., Ebskamp, M.J., Smeekens, J.C. et de Kruijff, B. (1998) Fructans interact strongly with model membranes. Biochim. Biophys. Acta. 1375: 36-42.

Devoto, A., Piffanelli, P., Nilsson, I., Wallin, E., Panstruga, R., von Heijne, G. et Schulze-Lefert, P, (1999) Topology, subcellular localization, and sequence diversity of the Mlo family in plants. J. Biol. Chem. 274: 34993-35004.

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

Doyle, M.R., Davis, S.J., Bastow, R.M., McWatters, H.G., Kozma-Bognar, L., Nagy, F., Millar, A.J. et Amasino, R.M. (2002) The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. Nature 419: 74-77.

Dunlap, J.C., Loros, J. et DeCoursey, P.J. (2004) Chronobiology: Biological Timekeeping. Sinauer, Sunderland, UK. 382 p.

Edwards, K.D., Lynn, J.R., Gyula, P., Nagy, F. et Millar AJ. (2005) Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. Genetics 170: 387-400.

Elstner, E.F. (1987) Metabolism of activated oxygen species. *Dans* Biochemistry of plants. *Éditeur*: D.D. Davies. Academic Press, London, UK. pp.253–315.

Emanuelsson, O., Nielsen, H. et von Heijne, G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. 8: 978-984.

Emanuelsson, O., Nielsen, H., Brunak, S. et von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300: 1005-1016.

Farré, E.M., Harmer, S.L., Harmon, F.G., Yanovsky, M.J. et Kay, S.A. (2005) Overlapping and distinct roles of PRR7 and PRR9 in the *Arabidopsis* circadian clock Curr. Biol. 15: 47-54.

Felsenstein, J. (1993) PHYLIP, Phylogeny Inference Package, Version 3.6. Distributed by the Author. Department of Genetics, University of Washington, Seattle.

Fleet, G. (1990) Yeasts in dairy products. J. Appl. Bacteriol. 68: 199-211.

Flower, D.R., North, A.C. et Attwood, T.K. (1993) Structure and sequence relationships in the lipocalins and related proteins. Protein Sci. 2: 753-761.

Flower, D.R. (1995) Multiple molecular recognition properties of the lipocalin protein family. J. Mol. Recognit. 8: 185-195.

Flower, D.R. (1996) The lipocalin protein family: structure and function. Biochem. J. 318: 1-14.

Flower, D.R., North, A.C. et Sansom, C.E. (2000) The lipocalin protein family: structural and sequence overview. Biochim. Biophys. Acta 1482: 9-24.

Flower, D.R. (2000) Beyond the superfamily: the lipocalin receptors. Biochim. Biophys. Acta 1482: 327-336.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G. et Putterill, J. (1999) GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. EMBO J. 18: 4679-4688.

Fowler, D.B., Limin, A.E. et Ritchie, J.T. (1999) Low-temperature tolerance in cereals: model and genetic interpretation. Crop Sci. 39: 626-633.

Fowler, S. et Thomashow, M.F. (2002) *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690.

Fowler, S.G., Cook, D. et Thomashow, M.F. (2005) Low temperature induction of *Arabidopsis* CBF1, 2, and 3 is gated by the circadian clock. Plant Physiol. 137: 961-968.

Foyer, C.H. et Noctor, G. (2005) Redox homeostasis and antioxidant signaling: A metabolic interface between stress perception and physiological responses. Plant Cell 17: 1866–1875.

Frenette-Charron, J.B., Breton, G., Badawi, M. et Sarhan, F. (2002) Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. FEBS Lett. 517: 129-132.

Ganfornina, M.D., Sanchez, D. et Bastiani, M.J. (1995) Lazarillo, a new GPI-linked surface lipocalin, is restricted to a subset of neurons in the grasshopper embryo. Development 121: 123-134.

Ganfornina, M.D., Gutiérrez, G., Bastiani, M. et Sánchez, D. (2000) A phylogenetic analysis of the lipocalin protein family. Mol. Biol. Evol. 17: 114-126.

Gattiker, A., Gasteiger, E. et Bairoch, A. (2002) ScanProsite: A reference implementation of a PROSITE scanning tool. Appl Bioinformatics. 1: 107-108.

Gille, L. et Nohl, H. 2001. The ubiquinol/bc_1 redox couple regulates mitochondrial oxygen radical formation. Arch. Biochem. Biophys. 388: 34-38.

Goder, V. et Spiess, M. (2001) Topogenesis of membrane proteins: determinants and dynamics. FEBS Lett. 504: 87-93.

Goodman, D.S. (1984) Overview of current knowledge of metabolism of vitamin A and carotenoids. J. Natl. Cancer Inst. 73: 1375-1379.

Gould, P.D., Locke, J.C., Larue, C., Southern, M.M., Davis, S.J., Hanano. S., Moyle, R., Milich, R., Putterill, J., Millar, A.J. et Hall, A. (2006) The molecular basis of temperature compensation in the *Arabidopsis* circadian clock. Plant Cell 18: 1177-1187.

Guindon, S. et Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52: 696-704.

Gutiérrez, G., Ganfornina, M.D. et Sánchez, D. (2000) Evolution of the lipocalin family as inferred from a protein sequence phylogeny. Biochim. Biophys. Acta 1482: 35-45.

Guy, C.L., Niemi, K.J. et Brambl, R. (1985) Altered gene expression during cold acclimation of spinach. Proc. Natl. Acad. Sci. USA. 82: 3673-3677.

Guy, C.L. (1990) Cold acclimation and freezing stress tolerance: role of protein metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 187-223.

Hajdukiewicz, P., Svab, Z. et Maliga, P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol. Biol. 25: 989–994.

Hamblen, M.J., White, N.E., Emery, P.T., Kaiser, K. et Hall, J.C. (1998) Molecular and behavioral analysis of four period mutants in *Drosophila melanogaster* encompassing extreme short, novel long, and unorthodox arrhythmic types. Genetics 149: 165-178.

Hannah, M.A., Heyer, A.G. et Hincha, D.K. (2005) A global survey of gene regulation during cold acclimation in Arabidopsis thaliana. PLoS Genet. 1:e26.

Hare, P.D., Cress, W.A. et van Staden, J. (1999). Proline synthesis and degradation: a model system for elucidating stress-related signal transduction. J. Exp. Bot. 50: 413-434.

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

Harmer, S.L. et Kay, S.A. (2005) Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*. Plant Cell 17: 1926-1940.

Havaux, M. et Kloppstech, K. (2001) The protective functions of carotenoid and flavonoid pigments against excess visible radiation at chilling temperature investigated in *Arabidopsis npq* and *tt* mutants. Planta 213: 953-966.

Hazen, S.P., Schultz, T.F., Pruneda-Paz, J.L., Borevitz, J.O., Ecker, J.R. et Kay, S.A. (2005) LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc. Natl. Acad. Sci. USA. 102: 10387-10392.

Hieber, A.D., Bugos, R.C et Yamamoto, H.Y. (2000) Plant lipocalins: violaxanthin de-epoxidase and zeaxanthin epoxidase. Biochim. Biophys. Acta 1482: 84-91.

Hieber, A.D., Bugos, R.C., Verhoeven, A.S. et Yamamoto, H.Y. (2002) Overexpression of violaxanthin de-epoxidase: properties of C-terminal deletions on activity and pH-dependent lipid binding. Planta 214: 476-483.

Higgins, J., Hodges, N.A., Olliff, C.J. et Phillips, A.J. (1987) A comparative investigation of glycinebetaine and dimethylsulphoxide as liposome cryoprotectants. J. Pharm. Pharmacol. 39: 577-582.

Higo, K., Ugawa, Y., Iwamoto, M. et Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res. 27: 297-300.

Holmberg, N. et Bülow, L. (1998) Improving stress tolerance in plants by gene transfer. Trends Plant Sci. 3: 61-66.

Horn, F., Weare, J., Beukers, M.W., Horsch, S., Bairoch, A., Chen, W., Edvardsen, O., Campagne, F. et Vriend, G. (1998) GPCRDB: an information system for G protein-coupled receptors. Nucleic Acids Res. 26: 275-279.

Houde, M., Dhindsa, R.S. et Sarhan, F. (1992) A molecular marker to select for freezing tolerance in Gramineae. Mol. Gen. Genet. 234: 43-48.

Houde, M., Danyluk, J., Laliberte, J.F., Rassart, E., Dhindsa, R.S. et Sarhan, F. (1992) Cloning, Characterization, and Expression of a cDNA Encoding a 50-Kilodalton Protein Specifically Induced by Cold Acclimation in Wheat. Plant Physiol. 99: 1381-1387.

Houde, M., Belcaid, M., Ouellet, F., Danyluk, J., Monroy, A.F., Dryanova, A., Gulick, P., Bergeron, A., Laroche, A., Links, M.G., MacCarthy, L., Crosby, W.L. et Sarhan, F. (2006) Wheat EST resources for functional genomics of abiotic stress. BMC Genomics 7: 149.

Huner, N.P.A., Oquist, G. et Sarhan, F. (1998) Energy balance and acclimation to light and cold. Trends Plant Sci. 3: 224-230.

Ikeda, K., Nakayashiki, H., Takagi, M., Tosa, Y. et Mayama, S. (2001) Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. Mol. Gen. Genet. 266: 318-325.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. et Speed, T.P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249-264.

Jones, D.T., Taylor, W.R. et Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences. CABIOS 8: 275-282.

Josefsson, L.G. et Rask, L. (1997) Cloning of a putative G-protein-coupled receptor from *Arabidopsis thaliana*. Eur. J. Biochem. 249: 415-420.

Journot-Catalino, N., Somssich, I.E., Roby, D. et Kroj, T. (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in Arabidopsis thaliana. Plant Cell 18: 3289-3302.

Kaczorowski, K.A. et Quail, P.H. (2003) *Arabidopsis* PSEUDO-RESPONSE REGULATOR7 is a signaling intermediate in phytochrome-regulated seedling deetiolation and phasing of the circadian clock. Plant Cell 15: 2654-2665.

Kawamura, Y. et Uemura, M. (2003) Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation. Plant J. 36: 141-154.

Kim, N.S., Park, N.I., Kim, S.H., Kim, S.T., Han, S.S. et Kang, K.Y. (2000) Isolation of TC/AG repeat microsatellite sequences for fingerprinting rice blast fungus and their possible horizontal transfer to plant species. Mol. Cells 10: 127-134.

Kishor, P.B.K, Hong, Z., Miao, G.H., Hu, C.A. et Verma, D.P.S. (1995) Overexpression of DI-pyrroline-5-carboxylate synthetase increases proline production and confers osmotolerance in transgenic plants. Plant Physiol. 108: 1387-1394.

Ko, R., Smith, L.T. et Smith, G.M. (1994) Glycine betaine confers enhanced osmotolerance and cryotolerance on Listeria monocytogenes. J.Bacteriol. 176: 426-431.

Krogh, A., Larsson, B., von Heijne, G. et Sonnhammer, E.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305: 567-580.

Kumar, A. et Minocha, S.C. (1998) Transgenic Manipulation of Polyamine Metabolism. Amsterdam: Harwood.Ladyman JAR, Hitz WD, Hanson AD. 1980. Translocation and metabolism of glycine betaine by barley plants in relation to water stress. Planta 150: 191-196.

Kyte, J. et Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132.

Larsson, J., Allhorn, M. et Kerstrom, B. (2004) The lipocalin alpha(1)-microglobulin binds heme in different species. Arch Biochem Biophys. 432: 196-204.

Leegood, R.C. et Furbank, R.T. (1986) Stimulation of photosynthesis by 2% oxygen at low-temperatures is restored by phosphate. Planta 168: 84-93.

Limin, A.E., Danyluk, J., Chauvin, L.P., Fowler, D.B. et Sarhan, F. (1997) Chromosome mapping of low-temperature induced Wcs120 family genes and regulation of cold-tolerance expression in wheat. Mol. Gen. Genet. 253: 720-727.

Liu, Y., Garceau, N.Y., Loros, J.J. et Dunlap, J.C. (1997) Thermally regulated translational control of FRQ mediates aspects of temperature responses in the neurospora circadian clock. Cell 89: 477-486.

Liu, Z., Chang G.Q. et Leibowitz, S.F. (2001) Apolipoprotein D interacts with the long-form leptin receptor: a hypothalamic function in the control of energy homeostasis. FASEB J. 15: 1329-1331.

Livak, K.J. et Schmittgen, T.D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. Methods 25:402–408.

Locke, J.C.W., Southern, M.M., Kozma-Bognar, L., Hibberd, V., Brown, P.E., Turner, M.S. et Millar, A.J. (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. Mol. Syst. Biol. 1: 2005.0013.

Lloyd, A.W., Ollif, C.J. et Rutt, K.J. (1994) A study of modified betaines as cryoprotective additives. J. Pharm. Pharmacol. 46: 704-707.

Lynch, D.V. et Steponkus, P.L. (1987) Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (Secale cereale L. cv Puma). Plant Physiol. 83: 761-767.

Lyons, J.M. (1973) Chilling injury in plants. Ann. Rev. Plant Physiol. 24: 445-466.

Machuka, J., Bashiardes, S., Ruben, E., Spooner, K., Cuming, A., Knight, C. et Cove, D. (1999) Sequence analysis of expressed sequence tags from an ABA-treated cDNA library identifies stress response genes in the moss *Physcomitrella patens*. Plant Cell Physiol. 40: 378-387.

Maggio, A., Miyazaki, S., Veronese, P., Fujita, T., Ibeas, J.I., Damsz, B., Narasimhan, M.L., Hasegawa, P.M., Joly, R.J. et Bressan, R.A. (2002) Does proline accumulation play an active role in stress-induced growth reduction? Plant J. 31: 699-712.

Majercak, J., Sidote, D., Hardin, P.E. et Edery, I.. (1999) How a circadian clock adapts to seasonal decreases in temperature and day length. Neuron 24: 219-230.

Malaba, L., Smeland, S., Senoo, H., Norum, K.R., Berg, T., Blomhoff, R. et Kindberg, G.M. (1995) Retinol-binding protein and asialo-orosomucoid are taken up by different pathways in liver cells. J. Biol. Chem. 270: 15686-15692.

Malpeli, G., Folli, C. et Berni, R. (1996) Retinoid binding to retinol-binding protein and the interference with the interaction with transthyretin. Biochim. Biophys. Acta 1294: 48-54.

Makino, S., Matsushika, A., Kojima, M., Oda, Y. et Mizuno, T. (2001) Light response of the circadian waves of the APRR1/TOC1 quintet: when does the quintet start singing rhythmically in *Arabidopsis*? Plant Cell Physiol. 42: 334-339.

Marin, E., Nussaume, L., Quesada, A., Gonneau, M., Sotta, B., Hugueney, P., Frey, A. et Marion-Poll, A. (1996) Molecular identification of zeaxanthin epoxidase of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. EMBO J. 15: 2331-2342.

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

Matsushika, A., Imamura, A., Yamashino, T. et Mizuno, T. (2002) Aberrant expression of the light-inducible and circadian-regulated APRR9 gene belonging to

the circadian-associated APRR1/TOC1 quintet results in the phenotype of early flowering in *Arabidopsis thaliana*. Plant Cell Physiol. 43: 833-843.

McClung, C.R., Fox, B.A. et Dunlap, J.C. (1989) The Neurospora clock gene frequency shares a sequence element with the Drosophila clock gene period. Nature 339: 558-562.

McClung, C.R. (2001) Circadian rhythms in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52:139-162.

McClung, C.R. (2006) Plant circadian rhythms. Plant Cell 18: 792-803.

McCue, K.F. et Hanson, A.D. (1992) Salt-inducible betaine aldehyde dehydrogenase from sugar beet: cDNA cloning and expression. Plant Mol. Biol. 18: 1-11.

McKersie, B.D., Chen, Y., de Beus, M., Bowley, S.R., Bowler, C., Inze, D., D'Halluin, K. et Botterman, J. (1993) Superoxide dismutase enhances tolerance of freezing stress in transgenic alfalfa (Medicago sativa L.). Plant Physiol. 103: 1155–1163.

McKersie, BD., Bowley, SR., Harjanto, E. et Le Prince, O. (1996) Water-deficit tolerance and field performance of transgenic alfalfa overexpressing superoxide dismutase. Plant Physiol. 111: 1177–1181.

McKersie, B.D., Bowley, S.R., et Jones, K.S. (1999) Winter survival of transgenic alfalfa overexpressing superoxide dismutase. Plant Physiol. 119: 839–848.

McKersie, B.D., Murnaghan, J., Jones, K.S. et Bowley, S.R. (2000) Iron-superoxide dismutase expression in transgenic alfalfa increases winter survival without a detectable increase in photosynthetic oxidative stress tolerance. Plant Physiol. 122: 1427–1438.

McNeil, S., Nuccio, M. et Hanson, A. (1999) Betaines and related osmoprotectants. Targets for metabolic engineering of stress resistance. Plant Physiol. 120: 945-949.

Michael, T.P., Salome, P.A., Yu, H.J., Spencer, T.R., Sharp, E.L., McPeek, M.A., Alonso, J.M., Ecker, J.R. et McClung, C.R. (2003) Enhanced fitness conferred by naturally occurring variation in the circadian clock. Science 302: 1049-1053.

Millar, A.J. (2004) Input signals to the plant circadian clock. J. Exp. Bot. 55: 277-283.

Millner, P.A. (2001) Heterotrimeric G-protein in plant cell signaling. New Phytol. 151: 165-174.

Mittler, R. (2002) Oxidative stress, antioxidants and stress tolerance. Trends Plant Sci. 7: 405–410.

Mittler, R., Vanderauwere, S., Gollery, M., et van Breusegem, F. (2004) Reactive oxygen gene network of plants. Trends Plant Sci. 9: 490–498.

Mizuno, T. et Nakamichi, N. (2005) Pseudo-Response Regulators (PRRs) or True Oscillator Components (TOCs). Plant Cell Physiol. 46: 677-685.

Möller, I.M. (2001) Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52: 561–91.

Möller, S., Croning, M.D. et Apweiler, R. (2001) Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 17: 646-653.

Monaco, H.L. et Zanotti, G. (1992) Three-dimensional structure and active site of three hydrophobic molecule-binding proteins with significant amino acid sequence similarity. Biopolymers 32: 457-465.

Morita, N., Nakazato, H., Okuyama, H., Kim, Y. et Thompson, G.A. Jr. (1996) Evidence for a glycosylinositolphospholipid-anchored alkaline phosphatase in the aquatic plant *Spirodela oligorrhiza*. Biochim. Biophys. Acta 1290: 53-62.

Murakami, M., Yamashino, T. et Mizuno, T. (2004) Characterization of circadian-associated APRR3 pseudo-response regulator belonging to the APRR1/TOC1 quintet in *Arabidopsis thaliana*. Plant Cell Physiol. 45: 645-650.

Nagao, M., Minami, A., Takezawa, D., Arakawa, K. et Fujikawa, S. (2001) ABA-induced freezing tolerance in *Physcomitrella patens* and gene expression (abstract no. 354[F455]). Plant Cell Physiol. 42: s121.

Nakai, K. et Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14: 897-911.

Nakai, K. et Horton, P. (1999) PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. 24: 34-36.

Nakamichi, N., Kita, M., Ito, S., Yamashino, T. et Mizuno, T. (2005) PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of *Arabidopsis thaliana*. Plant Cell Physiol. 46: 686-698.

Nakamura, T., Yokota, S., Muramoto, Y., Tsutsui, K., Oguri, Y., Fukui, K. et Takabe, T. (1997) Expression of a betaine aldehyde dehydrogenase gene in rice, a glycine betaine nonaccumulator, and possible localization of its protein in peroxisomes. Plant J. 11: 1115-1120.

Nanjo, T., Kobayashi, M., Yoshiba, Y., Sanada, Y., Wada, K., Tsukaya, H., Kakubari, Y.,Yamaguchi-Shinozaki, K. et Shinozaki, K. (1999) Biological functions of proline in morphogenesis and osmotolerance revealed in antisense transgenic *Arabidopsis thaliana*. Plant J. 18: 185-193.

Ndong, C., Danyluk, J., Huner, N.P.A. et Sarhan, F. (2001) Survey of gene expression in winter rye during changes in growth temperature, irradiance or excitation pressure. Plant Mol. Biol. 45: 691-703.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. et Somerville, C. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. Plant Physiol. 106: 1241-1255.

Nielsen, H., Engelbrecht, J., Brunak, S. et von Heijne, G. (1997) A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int J Neural Syst. 8: 581-599.

Nielsen, H. et Krogh, A. (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. *Dans* J. Glasgow., T. Littlejohn., F. Major., R. Lathrop., D. Sankoff., C. Sensen. *éditeurs*, Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp 122-130.

Niki, T. et Sakai, A. (1983) Effect of cycloheximide on the freezing tolerance and ultrastructure of cortical parenchyma cells from mulberry twigs. Can. J. Bot. 61: 2205–2221.

Nishida, I. et Murata, N. (1996) Chilling sensitivity in plants and cyanobacteria: The crucial contribution of membrane lipids. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 541-568.

Niyogi, K.K., Grossman, A.R. et Bjorkman, O. (1998) *Arabidopsis* mutants define a central role for the xanthophyll cycle in the regulation of photosynthetic energy conversion. Plant Cell 10: 1121-1134.

North, A.C. (1989) Three-dimensional arrangement of conserved amino acid residues in a superfamily of specific ligand-binding proteins. Int. J. Biol. Macromol. 11: 56-58.

Nuccio, M.L., Rhodes, D., McNeil, S.D. et Hanson, A.D. (1999) Metabolic engineering of plants for osmotic stress resistance. Curr. Opin. Plant Biol. 2: 128-134.

Olien, C.R. et Clark, J.L. (1993) Changes in soluble carbohydrate composition of barley, wheat, and rye during winter. Agr.J. 85: 21-29.

Page, R.D. (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput. Appl. Biosci. 12: 357-358.

Pearson, W.R. (1990) Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol. 183: 63-98.

Peitsch, M.C. et Boguski, M.S. (1990) Is apolipoprotein D a mammalian bilin-binding protein? New Biol. 2: 197-206.

Peitsch, M.C. (1996) ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. Biochem. Soc. Trans. 24: 274-279.

Perl-Treves, R. et Galun, E. (1991) The tomato Cu, Zn superoxide dismutase genes are developmentally regulated and respond to light and stress. Plant Mol. Biol. 17: 745-760.

Perras, M. et Sarhan, F. (1984) Energy state of spring and winter wheat during cold hardening. Soluble sugars and adenine nucleotides. Physiol. Plant. 60: 129-132.

Pervaiz, S. et Brew, K. (1985) Homology of beta-lactoglobulin, serum retinol-binding protein, and protein HC. Science 228: 335-337.

Pervaiz, S. et Brew, K. (1987) Homology and structure-function correlations between alpha 1-acid glycoprotein and serum retinol-binding protein and its relatives. FASEB J. 1: 209-214.

Peskan, T., Westermann, M. et Oelmuller, R. (2000) Identification of low-density Triton X-100-insoluble plasma membrane microdomains in higher plants. Eur. J. Biochem. 267: 6989-6995.

Pitcher, J.A., Freedman, N.J. et Lefkowitz, R.J. (1998) G protein-coupled receptor kinases. Annu. Rev. Biochem. 67: 653-692.

Plakidou-Dymock, S., Dymock, D. et Hooley, R. (1998) A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. Curr. Biol. 8: 315-324.

Prasad, T.K., Anderson, M.D., Martin, B.A. et Stewart, C.R. (1994) Evidence for chilling-induced oxidative stress in maize seedlings and a regulatory role for hydrogen peroxide. Plant Cell 6: 65-74.

Price, J.L. (1997) Insights into the molecular mechanisms of temperature compensation from the Drosophila PERIOD and TIMELESS mutants. Chronobiol. Int. 14: 455-468.

Raison, J.K. (1973) The influence of temperature-induced phase changes on kinetics of respiratory and other membrane-associated enzymes. J. Bioenerg 4: 258-309.

Rassart, E., Bedirian, A., Do Carmo, S., Guinard, O., Sirois, J., Terrisse, L. et Milne, R. (2000) Apolipoprotein D. Biochim. Biophys. Acta 1482: 185-198.

Restrepo, M.A., Freed, D.D. et Carrington, J.C. (1990) Nuclear transport of plant potyviral proteins. Plant Cell 2: 987-998.

Rivlin, R.S. (1996) Riboflavin. *Dans* Present knowledge in nutrition. *Éditeurs* E.E. Ziegler, et L. J. Filer. ILSI Press. Washington, DC. pp. 167-173.

Rosen, K.M. et Villa-Komaroff, L. (1990) An alternative method for the visualization of RNA in formaldehyde agarose gels. Focus 12: 23-24.

Rothenfluh, A., Abodeely, M., Price, J.L. et Young, M.W. (2000) Isolation and analysis of six timeless alleles that cause short- or long-period circadian rhythms in Drosophila. Genetics 156: 665-675.

Roughan, P.G. (1985) Phosphatidylglycerol and chilling sensitivity in plants. Plant Physiol. 77: 740-746.

Saito, A., Pietromonaco, S., Loo, A.K. et Farquhar, M.G. (1994) Complete cloning and sequencing of rat gp330/"megalin," a distinctive member of the low density lipoprotein receptor gene family. Proc. Natl. Acad. Sci. USA. 91: 9725-9729.

Salier, J.P. (2000) Chromosomal location, exon/intron organization and evolution of lipocalin genes. Biochim. Biophys. Acta 1482: 25-34.

Salomé, P.A. et McClung, C.R. (2005) What makes *Arabidopsis* tick: Light and temperature entrainment of the circadian clock. Plant Cell Environ. 28: 21-38.

Sánchez, D., Ganfornina, M.D. et Bastiani, M.J. (2000) Lazarillo, a neuronal lipocalin in grasshoppers with a role in axon guidance. Biochim Biophys Acta 1482: 102-109.

Sánchez, D., Ganfornina, M.D., Gutiérrez, G. et Marín, A. (2003) Exon–intron structure and evolution of the lipocalin gene family. Mol. Biol. Evol. 20: 775–783.

Sánchez, D., Ganfornina, M.D., Gutierrez, G., Gauthier-Jauneau, A.C., Risler, J.L. et Salier, J.P. (2006a) Lipocalins genes and their evolutionary history. *Dans* Lipocalins. *Éditeurs*: B. Åkerström, N. Borregaard, D.R. Flower, et J.P. Salier. Landes Bioscience.Georgetown, TX. pp.5-16.

Sánchez, D., Lopez-Arias, B., Torroja, L., Canal, I., Wang, X., Bastiani. M.J. et Ganfornina, M.D. (2006b) Loss of glial lazarillo, a homolog of apolipoprotein D, reduces lifespan and stress resistance in Drosophila. Curr. Biol. 16: 680-686.

Sansom, M.S. et Weinstein, H. (2000) Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices. Trends Pharmacol. Sci. 21: 445-451.

Sarhan, F. et D'Aoust, M.J. (1975) RNA synthesis in spring and winter wheat during cold acclimation. Physiol. Plant. 35: 62-65.

Sarhan, F. et Chevrier, N. (1985) Regulation of RNA synthesis by DNA-dependent RNA polymerases and RNases during cold acclimation in winter and spring wheat. Plant Physiol. 78: 250-255.

Sarhan, F. et Perras, M. (1987) Accumulation of a high molecular weight protein during cold hardening of wheat (Triticum aestivum L.). Plant Cell Physiol. 28: 1173-1179.

Sarhan, F., Ouellet, F. et Vazquez-Tello, A. (1997) The wheat wcs120 gene family: a useful model to understand the molecular genetics of freezing tolerance in cereals. Physiol. Plant. 101: 439-445.

Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. et Tabata, S. J. (2000) Structural analysis of *Arabidopsis thaliana* chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty P1 and TAC clones.DNA Res. 7: 31-63.

Sato, E., Nakamichi, N., Yamashino, T. et Mizuno, T. (2002) Aberrant expression of the *Arabidopsis* circadian-regulated APRR5 gene belonging to the APRR1/TOC1 quintet results in early flowering and hypersensitiveness to light in early photomorphogenesis. Plant Cell Physiol. 43: 1374-1385.

Sawyer, L. et Richardson, J.S. (1991) Using appropriate nomenclature. Trends Biochem. Sci. 16: 11.

Schaffer, M.A. et Fischer, R.L. (1988) Analysis of mRNA that accumulate in response to low temperature identifies a thiol protease gene in tomato. Plant Physiol. 87: 431-436.

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

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

Scheer, A., Fanelli, F., Costa, T., De Benedetti, P.G. et Cotecchia, S. (1996) Constitutively active mutants of the alpha 1-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. EMBO J. 15: 3566-3578.

Schleiff, E., Tien, R., Salomon, M. et Soll, J. (2001) Lipid composition of outer leaflet of chloroplast outer envelope determines topology of OEP7. Mol. Biol. Cell 12: 4090-4102.

Schubert, D. et LaCorbiere, M. (1985) Isolation of an adhesion-mediating protein from chick neural retina adherons. J. Cell Biol. 101: 1071-1077.

Sehgal, A., Price, J.L., Man, B. et Young, M.W. (1994) Loss of circadian behavioral rhythms and per RNA oscillations in the Drosophila mutant timeless. Science 263: 1603-1606.

Seiler, H. et Busse, M. (1990) The yeasts of cheese brines. Int. J. Food Microbiol. 11: 289-303.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. et Shinozaki, K. (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13: 61-72.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A. Sakurai, T. Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y. et Shinozaki, K. (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31: 279-292.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y. et Shinozaki, K. (2002) Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J. 31: 279-292.

Sen Gupta, A., Heinen, J.L., Holaday, AS., Burke, J.J. et Allen, R.D. (1993a). Increased resistance to oxidative stress in transgenic plants that overexpress chloroplastic Cu/Zn superoxide dismutase. Proc. Natl. Acad. Sci. USA. 90: 1629–1633.

Sen Gupta, A., Webb, R.P., Holaday, A.S. et Allen, R.D. (1993b). Over-expression of superoxide dismutase protects plants from oxidative stress: Induction of ascorbate peroxidase in superoxide dismutase-overexpressing plants. Plant Physiol. 103: 1067–1073.

Senoo, H., Stang, E., Nilsson, A., Kindberg, G.M., Berg, T., Roos, N., Norum, K.R. et Blomhoff, R. (1990) Internalization of retinol-binding protein in parenchymal and stellate cells of rat liver. J. Lipid Res. 31: 1229-1239.

Sharkey, T.D., Stitt, M., Heineke, D., Gerhardt, R., Raschke, K. et Heldt, H.W. (1986) Limitation of photosynthesis by carbon metabolism. II. O2-insensitive CO2 uptake results from limitations of triose phosphate utilization. Plant Physiol. 81: 1123-1129.

Shelp, B.J., Bown, A.W. et McLean, M.D. (1999) Metabolism and functions of gamma-aminobutyric acid. Trends Plant Sci. 4: 446-452.

Shieh, M.W., Wessler, S.R. et Raikhel, N.V. (1993) Nuclear targeting of the maize R protein requires two nuclear localization sequences. Plant Physiol. 101: 353-361.

Sigrist, C.J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A. et Bucher, P. (2002) PROSITE: A documented database using patterns and profiles as motif descriptors. Brief Bioinform. 3: 265-274.

Siminovitch, D., Rheaume, B., Pomeroy, K. et Lepage, M. (1968) Phospholipid, protein, and nucleic acid increases in protoplasm and membrane structures associated with development of extreme freezing resistance in black locust tree cells. Cryobiology 5: 202-205.

Simpson, G.G., Gendall, A.R. et Dean, C. (1999) When to switch to flowering. Annu. Rev. Cell. Dev. Biol. 15: 519-550.

Singh, J. et Laroche, A. (1988) Freezing tolerance in plants: a biochemical overview. Biochemistry and Cell Biol. 66: 650-657.

Sivaprasadarao, A., Boudjelal, M. et Findlay, J.B. (1993) Lipocalin structure and function. Biochem. Soc. Trans. 21: 619-622.

Slooten, L., Capiau, K., Van Camp, W., Van Montagu, M., Subesma, C. et Inzé, D. (1995) Factors affecting the enhancements of oxidative stress tolerance in transgenic tobacco overexpressing manganese superoxide dismutase in the chloroplasts. Plant Physiol. 107: 737–750.

Snedden, W.A. et Fromm, H. (1999) Regulation of the - aminobutyrate-synthesizing enzyme, glutamate decarboxylase, by calcium-calmodulin: a mechanism for rapid activation in response to stress. *Dans* Plant Responses to Environmental Stresses: From Phytohormones to Genome Reorganization, *Éditeurs*: H.R. Lerner. Marcel Dekker, New York, NY. 549-574.

South, G.R. et Whittick, A. (1987) Introduction to Phycology. Blackwell Scientific Publications, Oxfordshire, UK. pp. 341.

Steponkus, P.L. (1984) Role of the plasma membrane in freezing injury and cold acclimation. Annu. Rev. Plant Physiol. 35: 543-584.

Steponkus, P.L., Uemura, M., Balsamo, R.A., Arvinte, T. et Lynch, D.V. (1988) Transformation of the cryobehavior of rye protoplasts by modification of the plasma membrane lipid composition. Proc Natl Acad Sci USA. 85: 9026-9030.

Stitt, M., Grosse, H. et Woo, K.C. (1988) Interactions between sucrose synthesis and CO2 fixation. II. Alterations of fructose 2,6-bisphosphate during photosynthetic oscillations. J. Plant Physiol. 133:138-143.

Stitt, M. et Hurry, V. (2002) A plant for all seasons: alterations in photosynthetic carbon metabolism during cold acclimation in *Arabidopsis*. Curr. Opin. Plant Biol. 5: 199-206.

Suzuki, K., Lareyre, J.J., Sánchez, D., Gutiérrez, G., Araki, Y., Matusik, R.J. et Orgebin-Crist, M.C. (2004) Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2. Gene 339: 49-59.

Taiz, L. et Zeigler, E. (1991) Plant physiology. Benjamin/Cummings Pub Co. Redwood City, CA. 559 p.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K. et Shinozaki, K. (2002) Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. Plant J. 29: 417-426.

Thomashow, M.F. (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50: 571-599.

Thompson, A.J., Jackson, A.C., Parker, R.A., Morpeth, D.R., Burbidge, A. et Taylor, I.B. (2000) Abscisic acid biosynthesis in tomato: regulation of zeaxanthin epoxidase and 9-cis-epoxycarotenoid dioxygenase mRNAs by light/dark cycles, water stress and abscisic acid. Plant Mol. Biol. 42: 833-845.

Thompson, J.D., Higgins, D.G. et Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. et Higgins, D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876-4882.

Tramontano, W.A. et Jouve, D. (1997) Trigonelline accumulation in salt-stressed legumes and the role of other osmoregulators as cell cycle control agents. Phytochemistry 44: 1037-1040.

Tsang, E.W. Bowler, C., Herouart, D., Van Camp, W., Villarroel, R., Genetello, C., Van Montagu, M. et Inze, D. (1991) Differential regulation of superoxide dismutase in plants exposed to environmental stress. Plant Cell 3: 783-792.

Tusnàdy, G.E., Simon, I. (2001) The HMMTOP transmembrane topology prediction server. Bioinformatics 17: 849-850.

Uemura, M. et Yoshida, S. (1984) Involvement of plasma membrane alterations in cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). Plant Physiol. 75: 818-826.

Uemura, M., Joseph, R.A. et Steponkus, P.L. (1995) Cold acclimation of *Arabidopsis thaliana* (Effect on plasma membrane lipid composition and freeze-induced lesions). Plant Physiol. 109: 15-30.

Vágújfalvi, A., Kerepesi, I., Galiba, G., Tischner, T. et Sutka, J. (1999) Frost hardiness depending on carbohydrate changes during cold acclimation in wheat. Plant Sci. 144: 85-92.

Van Breusegem, F., Slooten, L., Stassart, J.M., Moens, T., Botterman, J., Van Montagu, M. et Inze, D. (1999) Overproduction of Arabidopsis thaliana FeSOD confers oxidative stress tolerance to transgenic maize. Plant Cell Physiol. 40: 515–523.

Van Camp, W., Capiau, K., Van Montagu, M., Inze, D. et Stoolen, L. (1996) Enhancement of oxidative stress tolerance in transgenic tobacco plants overproducing Fe-superoxide dismutase in chloroplasts. Plant Physiol. 112: 1703–1714.

van de Peer, Y. et de Wachter, R. (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput. Appl. Biosci. 10: 569-570.

Vazquez-Tello, A., Ouellet, F. et Sarhan, F. (1998) Low temperature-stimulated phosphorylation regulates the binding of nuclear factors to the promoter of *wcs120*, a wheat cold-specific gene. Mol. Gen. Genet. 257: 157-166.

Vereyken, I.J., Chupin, V., Demel, R.A., Smeekens, S.C. et De Kruijff, B. (2001) Fructans insert between the headgroups of phospholipids. Biochim. Biophys. Acta 1510: 307-320.

von Arnim, A.G., Deng, X.W. et Stacey, M.G. (1998). Cloning vectors for the expression of green fluorescent protein fusion proteins in transgenic plants. Gene 221: 35-43.

Walker, D.W., Muffat, J., Rundel, C. et Benzer, S. (2006) Overexpression of a Drosophila homolog of apolipoprotein D leads to increased stress resistance and extended lifespan. Curr. Biol. 16: 674-679.

Wang, D., Amornsiripanitch, N. et Dong X. (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS.* Pathog. 2: e123.

Weiser, C.J. (1970) Cold resistance and injury in woody plants. Science 169: 1269-1278.

Welti, R., Li, W., Li, M., Sang, Y., Biesiada, H., Zhou, H.E., Rajashekar, C.B., Williams, T.D. et Wang, X. (2002) Profiling membrane lipids in plant stress responses. Role of phospholipase D alpha in freezing-induced lipid changes in *Arabidopsis*. J Biol Chem. 277: 31994-32002.

Weretilnyk, E.A. et Hanson, A.D. (1990) Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. Proc. Natl. Acad. Sci. USA. 87: 2745-2749.

Wink, M. (1997) Special nitrogen metabolism. *Dans* P.M. Dey, J.B. Harbourne, *éditeurs*. Plant Biochemistry. Academic Press, San Diego, pp. 439-486.

Yamamoto, K. et Sasaki, T. (1997) Large-scale EST sequencing in rice. Plant Mol. Biol. 35: 135-144.

Yamamoto, Y., Yoshizawa, T., Kamio, S., Aoki, O., Kawamata, Y., Masushige, S. et Kato, S. (1997) Interactions of transthyretin (TTR) and retinol-binding protein (RBP) in the uptake of retinol by primary rat hepatocytes. Exp. Cell Res. 234: 373-378.

Yancey, P.H., Clark, M.E., Hand, S.C., Bowlus, R.D. et Somero, G.N. (1982) Living with water stress: evolution of osmolyte systems. Science 217: 1214-1222.

Yancey, P. (1994). Compatible and counteracting solutes. *Dans* Cellular and molecular physiology of cell volume regulation. *Éditeur* K. Strange. CRC Press. Boca Raton, FL. pp 81-109.

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Huang, X., Li, W., Li, J., Liu, Z., Li, L., Liu, J., Qi, Q., Liu, J., Li, L., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Zhang, J., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Ren, X., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W., Li, P., Chen, W., Wang, X., Zhang, Y., Hu, J., Wang, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo, W., Li, G., Liu, S., Tao, M., Wang, J., Zhu, L., Yuan. L. et Yang, H. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). Science. 296: 79-92.

Zehring, W.A., Wheeler, D.A., Reddy, P., Konopka, R.J., Kyriacou, C.P., Rosbash, M. et Hall, J.C. (1984) P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic Drosophila melanogaster. Cell 39: 369-376.

Zeigler, R.S., Tohme, J., Nelson, J., Levy, M. et Correa, F. (1994) Linking blast population analysis to resistance breeding: A proposed strategy for durable resistance. *Dans* Rice blast disease. *Éditeurs:* R.S. Ziegler, S.A. Leong, et P.S. Teng. CAB International, Wallingford, UK. 267-292.

Zhang, L., Dunn, M.A., Pearce, R.S. et Hughes, M.A. (1993) Analysis of organ specificity of low-temperature-responsive gene family in rye (*Secale cereale L.*). J. Exp. Bot. 44: 1787-1793.

Zhao, Y., Aspinall, D. et Paleg, L.G. (1992) Protection of membrane integrity in Medicago sativa L. by glycinebetaine against the effects of freezing. J. plant physiol. 140: 541-543.

Zhou, B.L., Arakawa, K., Fujikawa, S. et Yoshida, S. (1994) Cold-induced alterations in plasma membrane proteins that are specifically related to the development of freezing tolerance in cold-hardy winter wheat. Plant Cell Physiol. 35: 175-182.