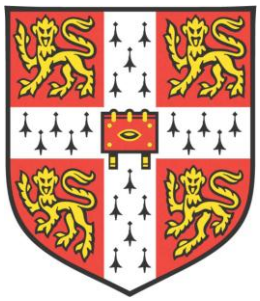
REGULATION OF THE CARBON-CONCENTRATING MECHANISM IN *CHLAMYDOMONAS REINHARDTII*



Madeline Mitchell

Gonville and Caius College

Department of Plant Sciences

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

September 2014

Dedicated to Ross

“It is good to have an end to journey toward;

but it is the journey that matters, in the end.”

Ursula K. Le Guin

DECLARATION

This dissertation is the result of my own work and includes no material that is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university or institution for any degree, diploma, or other qualification.

In accordance with the Degree Committee for Biology guidelines, this dissertation does not exceed 60,000 words.

Signed:

Date:

Madeline Mitchell

Cambridge

ACKNOWLEDGEMENTS

Firstly, my thanks go to my supervisor, Prof. Howard Griffiths for his boundless enthusiasm, support and encouragement. Secondly, thanks go to Dr Moritz Meyer for his day-to-day mentoring, humour, and for always making the time to help. I thank both him and Oliver Caspari for helpful discussions of the Chlamydomonas CCM and for generously allowing me to present some of their results in my dissertation. Thanks also to other members of the group for general discussions and advice and to Dr Jessica Royles, in particular.

I would also like to thank my second supervisor Prof. Julian Hibberd as well as members of his group, especially Dr Steven Burgess and Richard Smith-Unna, for their input. I have had the benefit of experimental support from Dr Jeremy Skepper (Cambridge Advanced Imaging Centre), Dr Mike Deery, Renata Feret and Dr Jo Rees (Cambridge Centre for Proteomics) as well as Dr Metodi Metodiev and Dr Gergana Metodieva (Proteomics group, University of Essex).

I would like to thank Dr Martin Jonikas and the members of his group, particularly Dr Leif Pallesen, Dr Luke Mackinder and Elizabeth Freeman, for welcoming me to the Carnegie Institution for Science (Stanford) during the summer of 2012. Thanks also to other members of the CAPP project, Prof. Alison Smith (John Innes Centre, Norwich) and Dr Alistair McCormick (University of Edinburgh), for constructive feedback.

I also wish to acknowledge the Herchel Smith Fund (University of Cambridge) for funding my PhD and the Cambridge Overseas Trust for appointing me as an Honorary Poynton Cambridge Australia Scholar. I am grateful to Prof. John Gray, the T. H. Middleton Fund (Graduate School of Life Sciences), Frank Smart Studentship (Department of Plant Sciences), Cambridge Philosophical Society, Society for Experimental Biology, Gonville and Caius College, Gatsby Foundation and organisers of Chlamy2012 for financial support for research and travel.

Thank you to my friends from Caius, Downing and the gymnastics club for making my time here so much fun, with special thanks to Daisy Hessenberger. Last but not least, I wish to thank my family, Jenny, David and Sarah Mitchell, for their unwavering love and support; and thanks always to Ross Dennis, for making me smile every day.

SUMMARY

Despite accounting for approximately half of global primary productivity, photosynthesis in aquatic environments is often limited by the availability of dissolved carbon dioxide ($CO_{2(aq)}$). To overcome the slow diffusion of $CO_{2(aq)}$ as well as kinetic limitations of the primary photosynthetic carboxylase, Rubisco, carbon concentrating mechanisms (CCMs) have evolved in many aquatic photosynthetic organisms to improve photosynthetic efficiency and growth in CO_2-limited environments.

In the model eukaryotic green alga *Chlamydomonas reinhardtii*, the CCM is induced under low CO_2 in the light and comprises: active inorganic carbon transport systems, carbonic anhydrases and the localisation of Rubisco to a central chloroplast microcompartment called the pyrenoid. In addition to changes in gene expression, acclimation to low CO_2 is accompanied by alterations to metabolism, physiology and even cellular ultrastructure. However, mechanisms governing the regulation and interaction of these molecular components to increase CCM activity remain poorly understood.

The overall aim of this study was to investigate regulation of the CCM in wild-type and mutant Chlamydomonas strains at both the whole cell and molecular level. Firstly, investigation of CCM induction in synchronised cultures of wild-type cells identified changes in CCM activity that were uncoupled from accumulation of CCM-related mRNA and protein, contrasting with the coordinated response to low CO_2 observed in asynchronous control cultures. Pre-dawn induction of the CCM was coincident with preferential localisation of Rubisco and a thylakoid-lumenal carbonic anhydrase (CAH3) to the pyrenoid, highlighting the possible role of endogenous signals and post-translational modifications in modulating CCM activity.

Secondly, in order to probe the relationship between pyrenoid formation and CCM induction and activity, CCM expression was investigated in pyrenoid-negative mutants with substituted Rubisco small subunits (RBCS). Low CO_2-adapted pyrenoid-less RBCS mutants had impaired growth and low photosynthetic affinity for inorganic carbon (Ci). These pyrenoid-negative strains also showed a specific reduction in the accumulation of several CCM mRNAs, compared to pyrenoid-positive wild-type. Two-dimensional difference in gel electrophoresis (2D-DIGE)

was used to compare the soluble proteome of one low CO_2-adapted pyrenoid-less RBCS mutant compared to the pyrenoid-positive wild-type. This analysis identified only a few differentially expressed proteins, none of which were directly involved in CCM activity. Two primary metabolic enzymes were more abundant in the wild-type while eight proteins associated with protein synthesis and photosynthesis were more abundant in the pyrenoid-less mutant, suggesting that pyrenoid loss is accompanied by global metabolic, as well as CCM-specific, changes.

A shotgun proteomics approach (LC-MS/MS) was used to extend the analysis of the pyrenoid-less RBCS mutant proteome to the whole genome level. Approximately 10% of the total proteins detected using this method were identified as differentially expressed between pyrenoid-negative and pyrenoid-positive strains. Increased abundance of photosynthetic proteins was found in the pyrenoid-less RBCS mutant, confirming the results of 2D-DIGE. In contrast, increased accumulation of CCM and primary metabolic enzymes was detected in the pyrenoid-positive wild-type. Overall, detailed investigation of the phenotype of pyrenoid-negative RBCS mutants indicates that pyrenoid loss leads to impaired induction of the CCM as well as altered metabolism under low CO_2 conditions, perhaps as a result of decreased carbon fixation.

The results of these studies are explored in the context of the identification of additional CCM components and regulatory mechanisms as well as possible connections between Rubisco aggregation and CCM activity.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

2D-DIGE	two-dimensional difference in gel electrophoresis
ABC-type	ATP-binding cassette transporter, also known as the ABCC superfamily
ATP	adenosine-5'-triphosphate
Bicine	N,N-Bis(2-hydroxyethyl)glycine
BSA	bovine serum albumin
CAH	carbonic anhydrase
cAMP	3'5'-cyclic-adenosine-monophosphate
CBB	Calvin-Benson-Bassham cycle
CCM	carbon-concentrating mechanism/CO_2-concentrating mechanism
CCP	chloroplast carrier protein
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
Ci	inorganic carbon
CIA	inorganic carbon accumulation mutant
CO_2	carbon dioxide
DCMU	3-(3,4-Dichlorophenyl)-1,1-dimethylurea
DE	differentially expressed
HCO_3^-	bicarbonate ion
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HLA	high light-activated
$K_{0.5}$	concentration of external Ci required for half maximal photosynthesis
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LCI	low CO_2-inducible
NADP(H)	Nicotinamide adenine dinucleotide phosphate (reduced form)
OD	optical density
pI	isoelectric point
PSI	photosystem I
PSII	photosystem II
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
rbcL	Rubisco large subunit
RBCS	Rubisco small subunit
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
SDS	sodium dodecyl sulphate
PAGE	polyacrylamide gel electrophoresis
s.e.	standard error of the mean
TCA	Tricarboxylic acid cycle (also known as citric acid or Krebs cycle)
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
Triton X-100	polyethylene glycol p-(1,1,3,3-tetramethylbutyl)-phenyl ether
Tween20	Polyoxyethylene (20) sorbitan monolaurate
UTR	untranslated region

LIST OF APPENDICES

1 INTRODUCTION

1.1 Carbon-concentrating mechanisms are present in most unicellular aquatic photosynthetic organisms

Despite accounting for approximately half of global primary productivity (Field et al., 1998), photosynthesis in aquatic environments is often limited by the availability of dissolved carbon dioxide ($CO_{2(aq)}$). This is in part due to the slow diffusion of CO_2 in aqueous solution and to the large fluctuations in inorganic carbon (Ci) concentration and speciation that may occur in both freshwater and marine environments on an occasional, daily or seasonal basis (Maberly, 1996; Tokoro et al., 2014). In addition, the primary photosynthetic carboxylase, Rubisco, may also limit productivity. Rubisco is slow, catalysing a few reactions per second, has a relatively low affinity for CO_2 and an oxygenase function that competes with carboxylation (reviewed in Spreitzer and Salvucci, 2002). To overcome these limitations and thereby improve photosynthetic efficiency and growth in CO_2-limited environments, carbon-concentrating mechanisms (CCMs) have evolved in many aquatic photosynthetic organisms. Although there are several different means by which these CCMs operate, they all work to saturate Rubisco active sites with CO_2, resulting in the suppression of photorespiration and enhancement of carbon fixation.

CCMs are thought to have evolved during one of the periods of decreasing atmospheric CO_2 and increasing O_2 concentrations that have occurred several times since the evolution of oxygenic photosynthesis at least 2.4 Ga. However, with no direct evidence of CCM presence in the fossil record, it is difficult to

determine when these mechanisms first evolved (for a recent review see: Raven et al., 2012). Although it is assumed that prokaryotic and eukaryotic CCMs involved independently, that is, that the prokaryotic CCM evolved after the primary endosymbiotic event (Raven, 2010), CCMs are now widely distributed across both prokaryotic and eukaryotic algal lineages, including most microalgae and some macroalgae (Giordano et al., 2005; Raven et al., 2012).

CCMs in unicellular algae are generally induced when inorganic carbon is limited, although modulation of CCM expression may also occur in response to numerous other environmental factors such as light availability and energy and nutrient supply (Beardall and Giordano, 2002; Raven et al., 2012). The energy costs associated with inducing and maintaining an active CCM – for example, the protein synthesis, photosynthetic electron transport and production of reduced energy carriers required to actively transport Ci – may play a vital role in determining the regulation and expression of components (Raven et al., 2014).

Despite evolutionary and genetic differences, the most common CCMs in cyanobacteria and unicellular eukaryotic algae act as biophysical pumps that actively transport inorganic carbon (CO_2 or HCO_3^-) across membranes and release CO_2 in the vicinity of Rubisco. This Introduction will first provide a brief description of the cyanobacterial CCM because it is the best characterised algal CCM and has functional similarities that are relevant to understanding the eukaryotic system. The CCM in the model eukaryotic green alga *Chlamydomonas reinhardtii* Dangard (Chlamydomonas; Fig. 1.1) will then be discussed in more detail. This will include an overview of the identification, function and regulation of the individual CCM components investigated in this study.

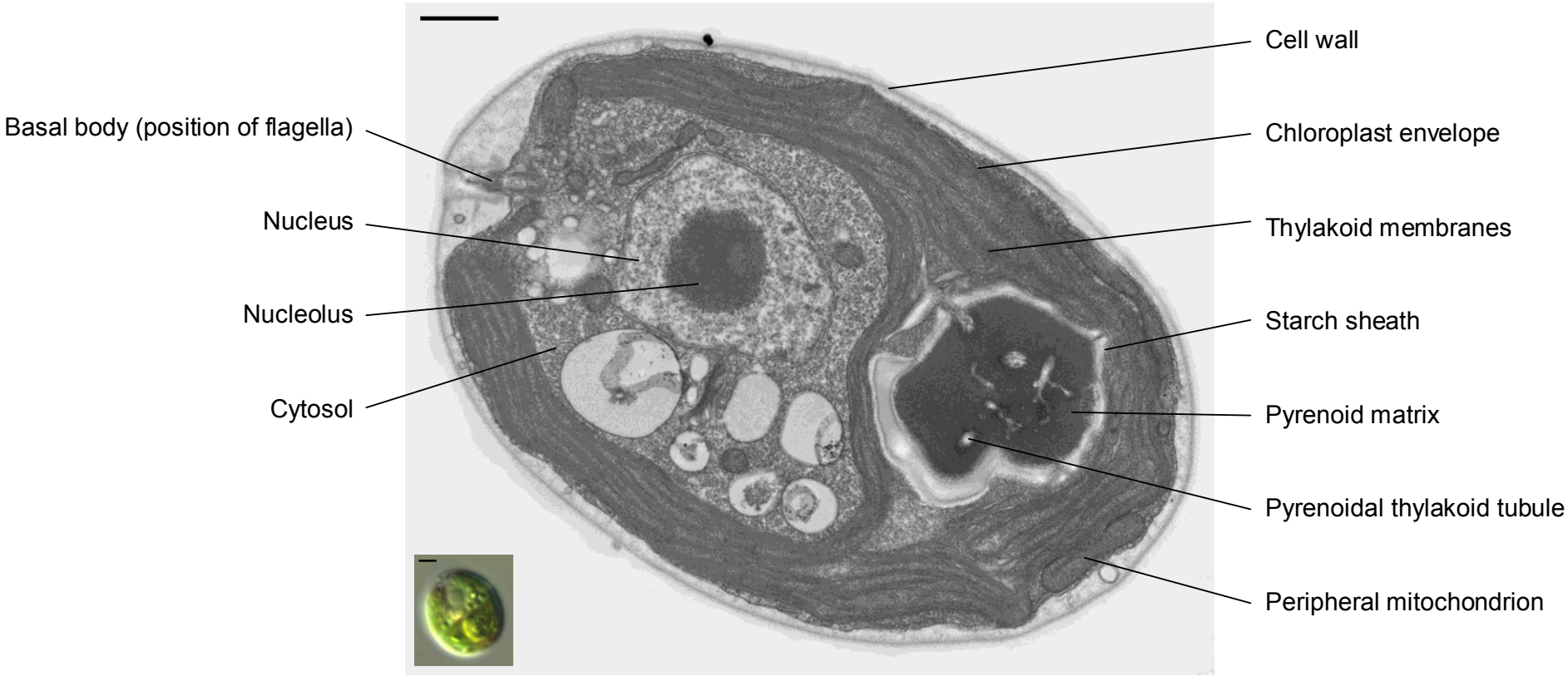


Figure 1.1 Transmission electron micrograph of wild-type *Chlamydomonas reinhardtii* (strain 2137 mt+) grown under CCM-induced conditions (low CO_2 in the light). Image courtesy of Dr Moritz Meyer. Inset: differential interference contrast light micrograph of wild-type Chlamydomonas cell. Scale bars = 1 μm.

1.2 The cyanobacterial/prokaryotic CCM is well characterised

Of all algal CCMs, the cyanobacterial system (for example, *Synechocystis* and *Synechococcus* species) is the best understood, with genetic and biochemical characterisation of many components (Price et al., 2008; Price et al., 2012; Rae et al., 2013). Cyanobacteria possess multiple Ci transport systems including both high and low affinity transporters and those that are inducible or constitutively expressed (Fig. 1.2; Price, 2011). Uptake of HCO_3^- across the plasma membrane can occur as symport of HCO_3^- and Na^+ or may be energised by the hydrolysis of ATP. CO_2 pumps on the thylakoid membrane are in fact specialised complexes that drive the conversion of CO_2 to HCO_3^- using NADPH as an electron donor. These pumps capture CO_2 entering the cell as well as recycling CO_2 before it leaks from the cell. Together with these CO_2 pumps, accumulation of the ionic form of Ci (HCO_3^-) and the lack of a cytosolic carbonic anhydrase minimise CO_2 leakage from the cell (Price and Badger, 1989; see also, Price et al., 2012).

Ci is further accumulated in specialised proteinaceous microcompartments called carboxysomes. Carboxysomes are defined as either α- or β- depending on the type of associated proteins and, despite evolutionary and structural differences, they appear to perform similar physiological functions (Whitehead et al., 2014). An ordered carboxysome shell surrounds a paracrystalline arrangement of Rubisco and a carbonic anhydrase is associated with the inside of the shell (Espie and Kimber, 2011; Rae et al., 2013). The carbonic anhydrase feeds CO_2 to Rubisco and the protein shell is thought to limit back diffusion, thereby elevating CO_2 concentrations around Rubisco active sites. Depending on the species and growth conditions, cyanobacterial cells generally contain 5-15 carboxysomes, and mutants that cannot make normal carboxysomes are only able to grow at high CO_2 concentrations (Price et al., 2012).

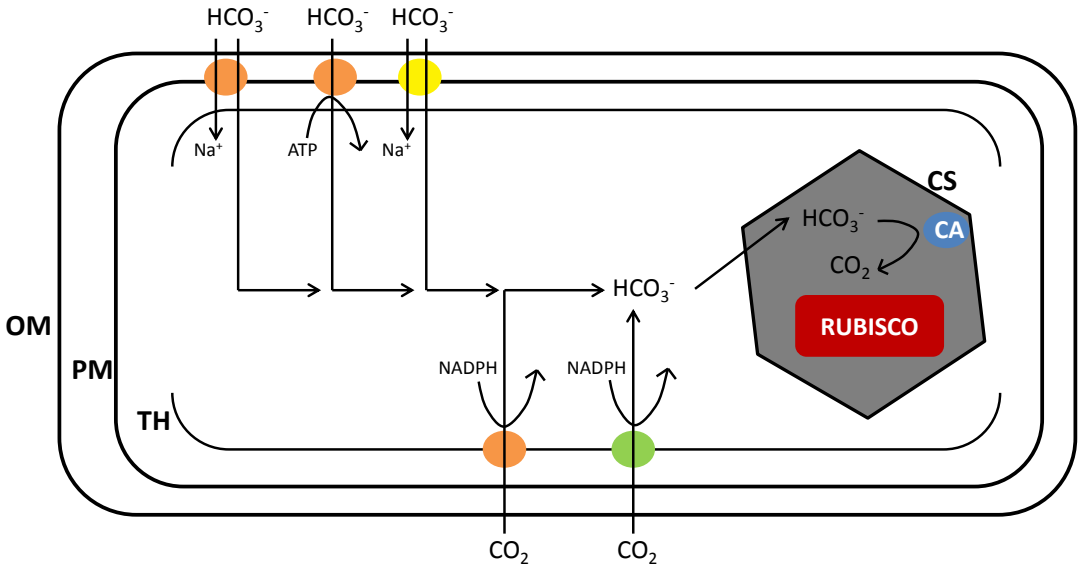


Figure 1.2 Simplified model of the CCM in β-cyanobacteria. Components include: inducible (orange), mostly inducible (yellow) and constitutive (green) inorganic carbon transport and recapture proteins at the plasma membrane (PM) and thylakoid membrane (TH); and a carbonic anhydrase (CA; blue) associated with the inner side of the carboxysome shell (CS). Rubisco (red) is packaged within the carboxysome; OM, outer membrane. Redrawn after Price et al. (2012).

1.3 Chlamydomonas is the model eukaryotic green alga with a CCM

Amongst the eukaryotic algae, the Chlamydomonas CCM is the best understood system (for reviews see: Moroney and Ynalvez, 2007; Wang et al., 2011; Jungnick et al., 2014). In Chlamydomonas, three elements are important for CCM activity (Fig. 1.3). First, Ci transporters localised to the plasma membrane, chloroplast envelope and, putatively, the thylakoid membrane, are required for Ci uptake and accumulation (Wang et al., 2011). Second, carbonic anhydrases operate in parallel within each cellular compartment (periplasm, cytosol, stroma, thylakoid lumen) and facilitate the interconversion of CO_2 and HCO_3^- (Moroney et al., 2011). Third, the localisation of Rubisco to a chloroplast microcompartment called the pyrenoid is required for CCM function (Ma et al., 2011; Meyer et al., 2012) and is thought to facilitate delivery of CO_2 to Rubisco and to minimise CO_2 leakage from the chloroplast.

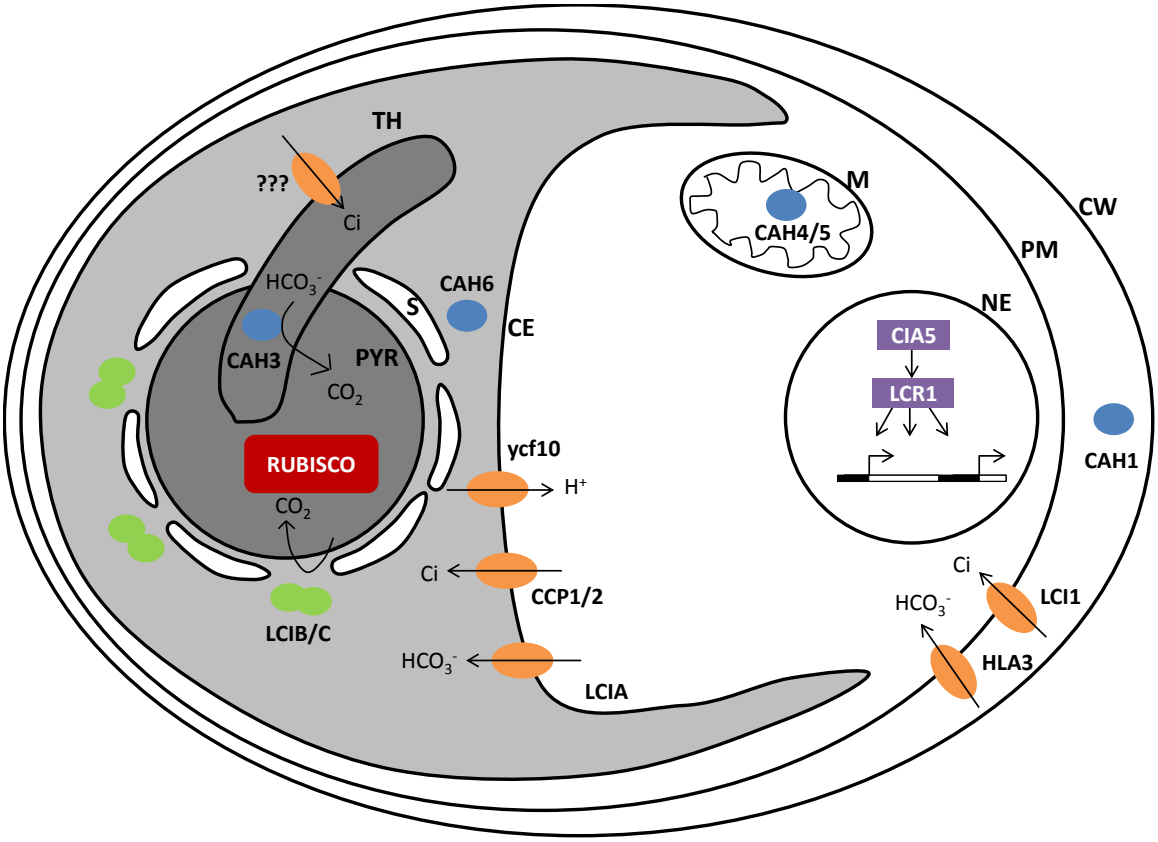


Figure 1.3 Model of the CCM in *Chlamydomonas reinhardtii*. Components include: transcription factors/regulators (purple), proteins involved in transmembrane transport of H^+ or Ci (orange), carbonic anhydrases (blue), CO_2 recapture proteins (green) and Rubisco (red). Gene name abbreviations: CAH, carbonic anhydrase; CCP, chloroplast carrier protein; CIA, inorganic carbon accumulation; HLA, high light-activated; LCI, low CO_2-inducible; LCR, low CO_2 stress response. Cell structure abbreviations: C, cell wall; PM, plasma membrane; M, mitochondrion; NE, nuclear envelope; CE, chloroplast envelope; TH, thylakoid membrane; PYR, pyrenoid; S, starch sheath.

1.3.1 Identification of CCM components and low CO_2-inducible genes

The CCM is induced at low CO_2 concentrations (in aqueous equilibrium with air levels) and results in an increased whole cell affinity for Ci and internal Ci accumulation (Badger et al., 1980), coincident with large transcriptional reorganisation (Miura et al., 2004; Yamano et al., 2008; Brueggeman et al., 2012; Fang et al., 2012) as well as the de novo synthesis of several proteins (Coleman and Grossman, 1984; Bailly and Coleman, 1988; Manuel and Moroney, 1988; Spalding and Jeffrey, 1989). Many putative CCM components were first identified as low CO_2-inducible (LCI) genes by smaller scale microarray studies. However, with an estimated third of the genome responding differentially to low CO_2 (Brueggeman et al., 2012), whole genome RNA sequencing experiments tend to yield too many, rather than too few, candidates for manual confirmation. The other drawback to this approach is that not all known elements of the CCM respond

transcriptionally to low CO_2 so it is possible that other important genes may be overlooked in these studies.

An alternative approach is to screen mutant populations for reduced growth at low CO_2 (for example, Spalding et al., 1983a; Moroney et al., 1989; Van et al., 2001; Colombo et al., 2002; Thyssen et al., 2003; Wang and Spalding, 2006; Ynalvez and Moroney, 2008; Jungnick et al., 2014). This has the advantage of identifying interesting candidates, that is, mutants with confirmed high CO_2-requiring phenotypes that are likely to be defective in the operation of the CCM. This approach can be useful to identify genes that are important for CCM function but that are not low CO_2-inducible. Mutagenic screens were used to identify the transcriptional regulator, CIA5 (inorganic carbon accumulation 5/CCM1), and thylakoid lumenal carbonic anhydrase, CAH3, which are essential to CCM function but are encoded by genes that are largely unresponsive to CO_2 (discussed in sections 1.3.2.5 and 1.3.2.2.2, respectively). However, the approach is generally labour-intensive, relying on phenotypic screening of up to tens of thousands of mutants and relatively low-throughput identification of the insertion site using PCR-based methods. A method for high-throughput genotyping has now been developed which, promisingly, is enabling the first genome-wide mutant screens (Zhang et al., 2014).

The following section briefly summarises our current understanding of the Chlamydomonas CCM, including the main genes investigated in this study, their function and regulation. Several reviews of the Chlamydomonas CCM have been published recently and it is not the author's intention to recapitulate any of these discussions. Rather, the aim is simply to provide enough detail to place the results of this study in context. For a more detailed discussion of the molecular components of the Chlamydomonas CCM in the wider context of algal diversity and evolution, please refer to Meyer and Griffiths (2013). Likewise, please refer to Wang et al. (2011) and Moroney et al. (2011) for comprehensive reviews of the Ci transporters and carbonic anhydrases, respectively. The gene nomenclature in this dissertation will be as defined in Wang et al. (2011) and Moroney et al. (2011) and is summarised in Table 1.1.

Table 1.1 Gene names and abbreviations of CCM components in *Chlamydomonas reinhardtii*

Abbreviated gene name	Full gene name	Functional category
CAH1/3/6	Carbonic anhydrase 1/3/6	Carbonic anhydrase
CCP1/2	Chloroplast carrier protein 1/2	Transmembrane Ci transport
CIA5	Inorganic carbon accumulation 5	Transcriptional regulator of low CO_2-responsive genes (CCM 'master switch')
HLA3	High light-activated 3	Transmembrane Ci transport
LCI1/A	Low CO_2-inducible 1/A	Transmembrane Ci transport
LCIB/C	Low CO_2-inducible B/C	Ci recapture
LCR1	Low CO_2 stress response 1	Transcription factor of *LCI* genes
rbcL	Rubisco large subunit	Carbon fixation/ pyrenoid component
RBCS	Rubisco small subunit	Carbon fixation/ pyrenoid component
ycf10	hypothetical chloroplast open reading frame 10	Transmembrane proton transport

1.3.2 Functional characterisation of CCM components

Large changes in metabolism and ultrastructure occur during acclimation to limiting Ci as well as the transcriptional downregulation of genes involved in photosynthesis, protein synthesis and other energy-consuming pathways (Brueggeman et al., 2012). Changes known to occur during CCM induction in Chlamydomonas include: differences in the vacuolisation of the cell and localisation of the mitochondria (Geraghty and Spalding, 1996), alterations in cyclic electron flow or other photosynthetic apparatus (Spalding et al., 1984), and increased photorespiration (Marek and Spalding, 1991; Renberg et al., 2010). However, the focus of this section is to introduce the components believed to be directly involved in the transport, recapture and accumulation of Ci, because expression of these elements provides a more direct means of investigating CCM regulation and thus forms the basis of the present study. The section will conclude with a discussion of the transcription and other regulatory factors responsible for induction and activation of these CCM genes. While homologues of most of the CCM components introduced in this section do exist in other species of algae (Meyer and Griffiths, 2013), this discussion will be limited solely to Chlamydomonas.

1.3.2.1 Inorganic carbon transport and recapture

Active transport of Ci is essential to CCM function and several studies have shown that Chlamydomonas is capable of transporting both CO_2 and HCO_3^- (Williams and Turpin, 1987; Sültemeyer et al., 1989; Palmqvist et al., 1994; Bozzo and Coleman, 2000), perhaps with a preference for CO_2 over HCO_3^- (Moroney and Tolbert, 1985; Aizawa and Miyachi, 1986). Katzman et al., (1994) presented evidence that HCO_3^- is the species accumulated in the chloroplast; however, alkaline conditions in the stroma and the presence of a carbonic anhydrase (CAH6) would quickly convert any CO_2 transported across the chloroplast envelope into HCO_3^-. The small, non-polar nature of CO_2 may allow it to passively diffuse across membranes or through protein channels whereas accumulation of the charged species, HCO_3^-, is likely to involve some kind of active uptake.

Multiple candidate Ci transporters have since been identified that are targeted to either the plasma membrane or chloroplast envelope (Wang et al., 2011; see also Fig. 1.3), but evidence is mixed as to their relative importance and substrate specificity. Identifying Ci transporter proteins has been generally challenging, in part due to the relative difficulty of isolating and analysing integral membrane proteins. CCP1 and CCP2 (chloroplast carrier proteins 1 and 2) are membrane-associated proteins but are still detectable in the soluble fraction, which may account for the early identification of these proteins. While the mRNAs encoding the Ci transport proteins identified thus far are all highly CO_2-responsive, which facilitates identification, other Ci transporter genes may be equally expressed under high and low CO_2. Multiple transport systems are likely based on the number of candidate transporters identified thus far as well as by analogy with the prokaryotic system. In cyanobacteria, high- and low-affinity Ci transporters work in parallel and may be inducible or constitutive. This apparent redundancy in transport systems may explain why no insertional mutant in a Ci transporter gene has been identified to date. Loss of one transporter might be compensated for under most conditions by the presence of other transporters in the same subcellular compartment.

A brief summary of the evidence for the role of each of the Chlamydomonas transporters in Ci uptake and accumulation will be presented in the following sections. Note that, although the current model of the CCM relies on a Ci

transporter in the thylakoid membrane, there is as yet no experimental evidence for its existence.

1.3.2.1.1 HLA3

HLA3 (high-light activated 3) was first identified as a high-light induced gene whose expression is repressed in photosystem I (PSI) and photosystem II (PSII) mutants and by the presence of PSII inhibitor 3-(3,4-Dichlorophenyl)-1,1-dimethylurea (DCMU). However, the induction of other CCM genes under high light, as well as the lack of *HLA3* induction in the *cia5* mutant, suggest that *HLA3* may in fact be responding to low CO_2 levels brought about by increased photosynthesis rather than the higher light intensities alone (Im and Grossman, 2002). HLA3 is predicted to be plasma membrane-localised and a member of the multi-drug resistance (MRP) family of the ATP-binding cassette (ABC-type) transporters (now known as the ABCC subfamily: Wanke and Kolukisaoglu, 2010). This ATP-dependence makes HLA3 the only energised Ci transporter identified in Chlamydomonas to date. Knock-down of *HLA3* results in a high pH-dependent reduction of photosynthetic affinity for Ci and Ci uptake under low CO_2, which suggests that HCO_3^- is the primary Ci species transported (Duanmu et al., 2009).

1.3.2.1.2 LCI1

LCI1 is a highly CO_2-responsive gene under the control of LCR1 (low-CO_2 stress response 1; Yoshioka et al., 2004) and CIA5 (Burow et al., 1996; Fukuzawa et al., 2001). LCI1 is predicted to contain four transmembrane domains and GFP fusion and cell fractionation experiments suggest that the protein is targeted to the plasma membrane. Expression of LCI1 in high CO_2-adapted (CCM-repressed cells) leads to increased internal Ci accumulation and affinity for Ci (Ohnishi et al., 2010). This strongly suggests that LCI1 is a Ci transporter, although no known transporter domains are predicted based on the amino acid sequence (SMART, http://smart.embl-heidelberg.de/). While the preference of LCI1 for a particular Ci species remains unclear, the LCI1-dependent mechanism has been shown to function in both acidic (pH 6.2) and alkaline (pH 7.8) conditions, distinguishing it from HLA3-dependent mechanisms of Ci transport that function mainly at alkaline pH (Ohnishi et al., 2010).

1.3.2.1.3 LCIA

LCIA (*NAR1.2*) is a member of the *Nar1* multigene family whose members have homology to the formate-nitrate transporter (FNT) family and whose expression is regulated by carbon as well as nitrogen in Chlamydomonas (Mariscal et al., 2006). *LCIA* is one of the most highly CO_2-responsive genes, encodes a protein that has six transmembrane domains and is predicted to be targeted to the chloroplast (ChloroP1.1, http://www.cbs.dtu.dk/services/ChloroP/). Expression of LCIA in a heterologous system (*Xenopus* oocyte) provided some evidence of low affinity HCO_3^- transport, although this was not the only substrate (Mariscal et al., 2006). More compelling evidence for a role in the CCM is that simultaneous knock-down of both *LCIA* and *HLA3* resulted in a strong decrease in growth, Ci uptake and photosynthetic Ci affinity at low CO_2. This phenotype was particularly apparent at high pH and suggests that HCO_3^- was the main Ci species transported (Duanmu et al., 2009). *LCIA* expression is absent in the *cia5* mutant (Miura et al., 2004), which also suggests that this gene is involved in carbon rather than nitrogen metabolism.

1.3.2.1.4 CCP1/2

CCP1 and CCP2 were first identified as membrane-associated proteins synthesised in response to low CO_2 (Spalding and Jeffrey, 1989) and were subsequently experimentally localised to the chloroplast envelope (Ramazanov et al., 1993). These 36 kDa proteins are 96% identical at the amino acid level and share sequence similarity with the mitochondrial carrier family of proteins (Chen et al., 1997), most of which transport metabolic intermediates, nucleotides and cofactors across the mitochondrial inner membrane (Monné and Palmieri, 2014). Both genes are highly responsive to low CO_2 and are not induced in the *cia5* mutant (Miura et al., 2004; Fang et al., 2012). *CCP1* and *CCP2* are clustered together in the genome with several other CO_2-responsive genes (*CAH1*, *CAH2*, *LCID*, *LCIE*; Supplemental Fig. S20 of Merchant et al., 2007). *LCID*/*CCP2* occur in a head-to-head conformation with *LCIE*/*CCP1* and their transcription is tightly co-regulated (Supplemental Fig. S6 of Brueggeman et al., 2012). This suggests that CCP1 and CCP2 might be involved in Ci transport into the chloroplast; however, RNAi mutants deficient in both CCP1 and CCP2 showed wild-type-like affinity for Ci despite exhibiting slower growth at low CO_2 (Pollock et al., 2004). This may be due to the activity of additional transporters compensating for the lack of CCP1/2 activity. Alternatively, CCP1 and CCP2 may be involved in the transport of other

metabolic intermediates, for example, photorespiratory compounds, that are required for long term growth at low CO_2.

1.3.2.1.5 ycf10

Disruption of *ycf10* results in reduced Ci uptake into the chloroplast (Rolland et al., 1997) and, apart from *rbcL*, *ycf10* is the only chloroplast-encoded gene with a known role in CCM function. Ycf10 has been localised to the chloroplast inner membrane (Sasaki et al., 1993) and, based on homology with the cyanobacterial protein cotA, is thought to only indirectly facilitate CO_2 transport (Katoh et al., 1996a; Katoh et al., 1996b). Ycf10 is likely to transport protons across the chloroplast envelope, which may be necessary for coupled transport or to maintain pH by balancing the hydroxyl ions released during the dehydration of HCO_3^-.

1.3.2.1.6 LCIB/C

Spalding et al. (1983c) identified a high CO_2-requiring mutant (*pmp-1*) that showed reduced capacity for Ci accumulation, suggesting that the lesion was in a gene involved in Ci transport. An allelic mutant, able to grow at high (>1%) and very low (<0.01%) but not low (0.04%) CO_2 concentrations (air-dier phenotype, *ad1*), was subsequently identified and the insertion was found to be in the low CO_2-responsive gene, *LCIB*. LCIB has no recognisable domains and has significant homology only to LCIC, LCID and LCIE (Wang and Spalding, 2006), which are similarly regulated at the transcriptional level during adaptation to low CO_2 (Brueggeman et al., 2012).

LCIB was first identified in a microarray experiment as a CIA5-regulated low CO_2-inducible gene (Miura et al., 2004) and the protein is 60% identical to LCIC at the amino acid level. When Yamano et al., (2010) identified a complex formed by LCIB and LCIC that appeared to accumulate in the peri-pyrenoidal space in response to low CO_2 and to light, they proposed a role in Ci recapture or forming a passive barrier to CO_2 (Yamano et al., 2010). There is now other evidence suggesting that the peri-pyrenoidal localisation of LCIB/LCIC actually occurs in response to very low rather than low CO_2 (Wang and Spalding, 2014); however, the essential role of LCIB to growth at low CO_2 is not in doubt. Finally, loss of thylakoid lumenal carbonic anhydrase (CAH3) function in the *LCIB* mutant restores growth at low

CO_2, but not to wild-type levels, indicating that LCIB functions downstream of CAH3 (Duanmu and Spalding, 2011).

1.3.2.2 Carbonic anhydrases

CCM function in cyanobacteria relies on only a single internal (carboxysomal) carbonic anhydrase, with ectopic expression of human carbonic anhydrase in the cytosol abolishing CCM activity (Price and Badger, 1989). Chlamydomonas, on the other hand, has at least twelve genes encoding carbonic anhydrases, which are present in most subcellular compartments (Moroney et al., 2011). At least four carbonic anhydrases have putative roles in CCM function, although the strength of the evidence is mixed. Periplasmic (*CAH1*) and mitochondrial (*CAH4/5*) carbonic anhydrase genes are strongly induced by low CO_2 but may be important only under specific conditions (*CAH1*) or are only indirectly implicated in CCM function (*CAH4/5*). CAH3 has been convincingly localised to the thylakoid lumen and mutant studies have demonstrated that it is essential to growth at low CO_2 but the role and localisation of the other chloroplast carbonic anhydrase, CAH6, is less well defined.

1.3.2.2.1 CAH1

CAH1 is a highly CO_2-responsive gene encoding a periplasmic carbonic anhydrase that is detectable in low but not high CO_2-adapted cells (Fukuzawa et al., 1990). Due to the early identification of this gene and its expression pattern, CAH1 is often used as a marker for CCM induction. However, the *cah1* mutant exhibits normal growth at low CO_2 over a range of pH values (Van and Spalding, 1999) and removing external CA activity from a wall-less strain does not affect its photosynthetic affinity for Ci (Williams and Turpin, 1987). While this indicates that CAH1 is not generally required for a fully functioning CCM, photosynthetic oxygen evolution curves suggest that CAH1 may still be important for CCM activity at very low Ci (Van and Spalding, 1999), such as when there is a strong boundary layer effect. Alternatively, other carbonic anhydrases located in the periplasm (CAH2) or plasma membrane (CAH8) might be able to compensate for the loss of CAH1 activity (Moroney et al., 2011).

1.3.2.2.2 CAH3

One carbonic anhydrase that is essential for CCM function is the thylakoid lumen-localised CAH3, which was first identified in mutants with severe high CO_2-requiring phenotypes (Spalding et al., 1983a; Moroney et al., 1986). In one mutant, CAH3 is truncated by a nonsense mutation (Funke et al., 1997) and in the second, mutations in the chloroplast transit peptide lead to mislocalisation and possibly degradation of the protein (Karlsson et al., 1998; Sinetova et al., 2012).

Two different roles for CAH3 have been suggested. The first, proposed by Pronina et al., is that CAH3 is needed to supply Rubisco with sufficient CO_2 by dehydrating HCO_3^- accumulated in the thylakoid lumen (Pronina and Semenenko, 1992; Pronina and Borodin, 1993). Calculations by Raven (1997) suggest that the relatively acidic conditions in the lumen, while favouring CO_2 over HCO_3^-, are inadequate for providing Rubisco with saturating levels of CO_2 and that the presence of a lumenal carbonic anhydrase would overcome this problem of supply.

The second role is that CAH3 supplies HCO_3^- to PSII reaction centres, which removes excess protons, stabilises the manganese cluster and enhances the water splitting reactions (Park et al., 1999; Villarejo et al., 2002; Shutova et al., 2008). Evidence for this hypothesis includes immunoblot analyses showing that CAH3 is absent in a mutant without the PSII reaction centre but present in a mutant lacking the PSII light-harvesting complex, which supports the association of CAH3 with polypeptides on the donor side of PSII (Markelova et al., 2009).

Hanson et al. (2003) attempted to distinguish between the two possible roles of CAH3 by analysing gas exchange and metabolite pools in the *cia3* mutant. The authors demonstrated that, under low CO_2, the mutant cannot supply Rubisco with enough CO_2 and that PSII function has no direct effect on CO_2 fixation. This indicates that the main function of CAH3 when the CCM is induced is supplying CO_2 to Rubisco and that stabilisation of PSII is a secondary effect.

Given the minimal transcriptional response of *CAH3* to low CO_2 and that the protein is expressed equally under high and low CO_2 (Karlsson et al., 1998), post-translational mechanisms of regulation are likely to be important during CCM induction and CAH3 activation. Indeed, new evidence indicates that the

phosphorylation of CAH3 in response to low CO_2 is correlated with increased carbonic anhydrase and CCM activity as well as preferential localisation to the pyrenoid (Blanco-Rivero et al., 2012). This highlights both the importance of protein localisation to CCM function and post-translational modification to CCM protein activity.

1.3.2.2.3 CAH6

The second chloroplastic carbonic anhydrase, CAH6, has been localised to the stroma and assigned two putative roles in the CCM. CAH6 may participate in Ci (re)capture by converting CO_2 to HCO_3^- in the alkaline conditions of the stroma because accumulation of the charged (HCO_3^-) rather than uncharged (CO_2) species limits diffusion of Ci back across the chloroplast envelope (Moroney et al., 2011). The CO_2 in this case may come either from transport across the chloroplast envelope or from leakage from the pyrenoid. Immunogold localisation suggests preferential localisation of CAH6 in the area around the pyrenoid (Mitra et al., 2004), which would be consistent with a role in the recapture of CO_2 as it leaks past Rubisco.

However, peri-pyrenoidal localisation of CAH6 would also be consistent with the alternative role of supplying Rubisco with CO_2 by converting some of the abundant HCO_3^- into CO_2 in the vicinity of the pyrenoid (Badger and Price, 1994). In either case, CAH6 might also be interacting indirectly with either the LCIB/LCIC complex or with the pyrenoidal thylakoid membranes (Yamano et al., 2010). Similar to *CAH3*, *CAH6* is only minimally induced during adaptation to low CO_2 (Mitra et al., 2004) so perhaps differential localisation and/or post-translational modification of CAH6 also occurs during CCM induction.

Nevertheless, future work should first focus on definitely localising CAH6 because this protein has also been identified in the flagellar proteome (Pazour et al., 2005). Immunogold localisation shows antibodies recognising sites that are mostly adjacent to starch granules (Mitra et al., 2004), which may instead be non-specific binding. If CAH6 is indeed flagellar, a different stromal carbonic anhydrase may instead perform the role in the CCM described above. CAH7 has not been experimentally localised but is predicted to be targeted to the chloroplast (TargetP, Emanuelsson et al., 2000; WoLF PSORT, Horton et al., 2007) and has a similar gene expression profile to CAH6 (Moroney et al., 2011).

1.3.2.2.4 CAH4/5

CAH4 and *CAH5* (formerly *Mca1* and *Mca2*) are 96% identical at the cDNA level and encode mitochondrial carbonic anhydrases that are identical at the amino acid level, suggesting that they are the result of gene duplication (Eriksson et al., 1996). *CAH4*/*5* transcription is strongly induced by low CO_2 conditions and is also regulated to some extent by light intensity and circadian rhythms (Villand et al., 1997; Eriksson et al., 1998). One model for CAH4/5 function is that it limits CO_2 leakage from the cell by converting CO_2 produced in the mitochondria to HCO_3^-, which may then return to the cytosol via channels in the inner mitochondrial membrane (Raven, 2001). An alternative hypothesis, based the observation that CAH4/5 expression is suppressed when nitrogen is limiting even under low CO_2, is that CAH4/5 is sometimes involved in the supply of HCO_3^- for anaplerotic assimilation of nitrogen (Giordano et al., 2003). These hypotheses are not mutually exclusive and the role of CAH4/5 may even extend to include regulation of pH (Moroney et al., 2011).

1.3.2.3 Aggregation of Rubisco and other proteins in or near the pyrenoid

In most single-celled algae, CCM activity is associated with the chloroplast pyrenoid, a specialised area of the chloroplast where Rubisco is aggregated (Morita et al., 1999; Thoms et al., 2001; Giordano et al., 2005; Moroney and Ynalvez, 2007). This structure is functionally analogous to the cyanobacterial carboxysome in that it is thought to limit diffusion of CO_2 away from Rubisco. In Chlamydomonas, CCM induction coincides with the aggregation of almost all Rubisco in the pyrenoid and around half of the cell's Rubisco is retained in the pyrenoid even at high CO_2 (Borkhsenious et al., 1998). Mutants with normal Rubisco levels but abnormal or absent pyrenoids also lack a fully functioning CCM (Ma et al., 2011; Meyer et al., 2012). The pyrenoid has no discernible outer protein or membrane structure but some thylakoid lamellae traverse the pyrenoid matrix where they become tubular. It is in these regions that CAH3 is preferentially localised when the CCM is induced (Sinetova et al., 2012).

The *cia6* mutant has reduced growth at low CO_2, reduced affinity for Ci and a disorganised pyrenoid. Together with the high CO_2-requiring phenotype, the mislocalisation of over half the Rubisco particles in the mutant is further evidence of the importance of the pyrenoid to CCM function. Although *CIA6* encodes a

protein with sequence similarity to a SET domain methyltransferase, no methyltransferase activity was detected in vitro, which leaves the exact function of CIA6 to be discovered (Ma et al., 2011). Localisation of CIA6 in high and low CO_2-adapted cells would be useful to determine whether this protein interacts directly or indirectly with the pyrenoid/Rubisco.

Other proteins that are associated with pyrenoid structure as well as CCM function include LCIB and LCIC. Mutants in LCIB localisation have either reduced Ci affinity and accumulation, abnormal numbers or sizes of pyrenoids or a combination of both, which further strengthens the putative link between LCIB function/localisation, pyrenoid development and CCM activity. Inhibitor studies have shown that accumulation of the LCIB/LCIC complex around the pyrenoid requires protein synthesis whereas diffusion away from the pyrenoid in the dark appears to be a passive process (Yamano et al., 2014).

Immunogold labelling experiments have also localised Rubisco activase to the pyrenoid (McKay et al., 1991) and some enzymes of the Calvin-Benson-Bassham (CBB) cycle to the stromal side of the starch sheath that surrounds the pyrenoid (Süss et al., 1995). Formation of the starch sheath itself coincides with CCM induction and was thought to participate in the CCM by acting as a barrier to CO_2 leakage (Ramazanov et al., 1994). However, physiological characterisation of a starch-less mutant found no difference in CCM phenotype between the mutant and wild-type strains (Villarejo et al., 1996a). Early attempts at pyrenoid purification identified the presence of several proteins other than Rubisco of varying molecular weights but these were not further characterised (Kuchitsu et al., 1988; McKay and Gibbs, 1991).

1.3.2.4 Proteins of unknown function

While thousands of genes are CO_2-responsive and, hence, indirectly associated with adaptation to low CO_2, several genes have been identified as having a more direct role in the CCM even though their exact function remains unknown (Jungnick et al., 2014). For example, transcripts of the CIA5-controlled gene, *LCI5,* are highly abundant in wild-type cells and show a relatively small but significant increase in abundance in response to low CO_2 (Brueggeman et al., 2012; Fang et al., 2012). LCI5 is chloroplast-localised and is phosphorylated specifically in response to low CO_2 but has no conserved domains or homology to other proteins

from organisms with a sequenced genome (Turkina et al., 2006). A mutant in the *LCI5* gene would be useful in determining its function.

1.3.2.5 Transcription factors/regulators

1.3.2.5.1 Master regulator CIA5

The CCM 'master switch' and transcriptional regulator of CO_2-responsive genes, CIA5/CCM1, was discovered independently by two different groups (Fukuzawa et al., 2001; Xiang et al., 2001). Mutants with insertions in this gene are unable to acclimate to low CO_2 and were originally identified as failing to induce a number of known CO_2-responsive genes, including *CAH1*, *CAH4, CAH5*, *CCP1*, *CCP2* and *LCI1*, as well as several low-CO_2 induced polypeptides (Moroney et al., 1989; Fukuzawa et al., 2001). CIA5 has a putative zinc finger (DNA binding) domain, is expressed equally under high and low CO_2 and is localised to the nucleus (Fukuzawa et al., 2001; Xiang et al., 2001; Wang et al., 2005). An enhancer element consensus sequence (EEC), present in the upstream region of several genes has also been implicated in CIA5-controlled gene expression (Yoshioka et al., 2004). This would suggest that CIA5 acts as a transcription factor that is activated under low CO_2; however, there is no experimental evidence for this protein directly binding DNA.

Although the exact mechanism of regulation remains unclear, microarray and transcriptomic studies have identified 2,787 genes (approximately 18% of the genome) whose expression is strongly correlated with CIA5 function and/or low CO_2 (Miura et al., 2004; Wang et al., 2005; Fang et al., 2012). A global analysis of gene expression during adaptation to low CO_2 comparing the *cia5* mutant to the wild-type identified genes showing altered expression in response to the loss of CIA5 function, strong responses to low CO_2 alone and responses to a combination of the two factors (Fang et al., 2012). CIA5 may thus interact with or act upstream of a number of other transcription factors and regulators to control gene expression during adaptation to low CO_2.

1.3.2.5.2 Other regulators of gene expression

The Myb-domain protein LCR1 operates downstream of CIA5 and has been shown experimentally to bind the *CAH1* promoter, controlling the transcription of *LCI1* and *LCI6* as well as *CAH1* (Yoshioka et al., 2004). Winck et al. (2013)

identified *LCR1* and a homologue, *LCR2*, as part of a large gene expression network using formaldehyde-assisted isolation of regulatory elements (FAIREseq). According to this model, *LCR1* and *LCR2* are both predicted to regulate several other transcription factors or regulators.

Analysis of the Chlamydomonas genome predicts around 230 transcription factors and regulators (Riaño-Pachón et al., 2008) and the expression of almost 60 of these changed at least two-fold during the CO_2 response (Winck et al., 2013), which further highlights the massive transcriptional rearrangement that occurs during CCM induction. Given that low CO_2-induced gene expression is predicted to be mediated by CIA5 via LCR1 and then via LCR1-regulated factors, the existence of one or more signalling cascades and feedback loops is plausible. As an example, Yoshioka et al. (2004) propose a model of CAH1 regulation in which the initial low CO_2 signal, transduced via CIA5, results in the transcription of both *LCR1* and *CAH1* through proteins that recognise a common motif upstream of these genes. Transcription of *CAH1* is then enhanced and maintained by the binding of LCR1 itself to as yet unknown enhancer elements in the *CAH1* promoter (Yoshioka et al., 2004). The presence of multiple regulatory sequences and multiple regulatory factors may provide a means of integrating different environmental and cellular signals to modulate CCM activity.

1.3.3 CCM regulation remains poorly understood

Despite our increasing knowledge of CCM components, the interaction and regulation of these components to produce the low CO_2/CCM-induced phenotype remains poorly understood. Although exposure to low CO_2 leads to many changes in gene expression, metabolism, physiology and even cellular ultrastructure, little is known about the signal perception and transduction that leads to these changes. To date, only two transcription factors/regulators have been identified and the regulatory elements required for low CO_2-induced gene transcription are largely unknown. The mobility and post-translational modification of some CCM components is intriguing and it remains to be determined how widespread these levels of control are. Finally, there exists the possibility of feedback between, and differential regulation of, CCM and other cellular components, for example, CCM mutants deficient in one element may alter the regulation of other elements in compensation.

1.4 General aims and hypotheses of this study

The general aim of this dissertation was to investigate regulation of the Chlamydomonas CCM by linking gene expression and molecular level changes with whole cell CCM physiology and growth in wild-type and mutant strains grown under contrasting conditions of light and CO_2. More specifically, the objectives of each area of research were to:

1. Identify mechanisms regulating CCM induction in synchronised cells during the dark-to-light transition (Chapter 3). This approach consisted of conducting detailed time course experiments measuring gene expression and physiology in synchronised wild-type cells during the dark-to-light transition. Unsynchronised cells adapting to low CO_2 were used a comparison because this is a well characterised response in Chlamydomonas.
2. Determine the effect of pyrenoid loss on CCM function and gene expression in pyrenoid-less Rubisco small subunit (RBCS) mutants (Chapter 4). This was a targeted approach and involved measuring growth of mutants, photosynthetic affinity for Ci and expression of known CCM components under high and low CO_2. The soluble proteomes of low CO_2-adapted RBCS pyrenoid-negative (spinach RBCS hybrid) and pyrenoid-positive (wild-type) strains were then analysed in more detail using 2D-DIGE and mass spectrometric identification of differentially expressed proteins.
3. Extend the analysis of differential protein expression in low CO_2-adapted RBCS pyrenoid-negative (spinach RBCS hybrid) and pyrenoid-positive (wild-type) cells to the genome level (Chapter 5). In this comprehensive, untargeted approach, shotgun LC-MS/MS was used to identify differentially expressed proteins in the total proteome and build up a general model of the cellular and metabolic response to low CO_2 of pyrenoid-negative cells compared to pyrenoid-positive cells.

2 MATERIALS AND METHODS

2.1 Author contributions

The author designed and carried out all experiments and data analysis except for some elements of the proteomic experiments described in section 2.6. The author designed both proteomics experiments (2D-DIGE and shotgun LC-MS/MS), grew and harvested the cells and performed the final data analysis. Sample preparation, mass spectrometric and statistical analyses were performed by specialist proteomics facilities.

Protein extraction, 2D-DIGE, image analysis, excision and mass spectrometric analysis of differentially expressed protein spots were carried out by Renata Feret and Dr Jo Rees (Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge). The author then analysed the mass spectrometric data (identified proteins and peptides) for each differentially expressed protein spot with some advice from Dr Mike Deery and Dr Jo Rees (Cambridge Centre for Proteomics).

Protein extraction, shotgun LC-MS/MS analysis and statistical analysis of the results was carried out by Dr Metodi Metodiev and Dr Gergana Metodieva (School of Biological Sciences, University of Essex). The author performed all subsequent analyses of the data including the collation of differentially expressed proteins and analysis of protein function.

2.2 Algal strains and growth conditions

The *Chlamydomonas reinhardtii* walled wild-type strain used for time courses of CCM induction in Chapter 3 was 2137 mt+ (Spreitzer and Mets, 1981), which is a commonly used wild-type strain for CCM studies (Spalding and Ogren, 1982, 1983; Spalding et al., 1983b; Coleman and Grossman, 1984; Marcus et al., 1986; Dionisio-Sese et al., 1990; Marek and Spalding, 1991; Palmqvist et al., 1994; Geraghty and Spalding, 1996; Bozzo and Colman, 2000). The other wild-type strains used for development of the oxygen evolution method were 137C mt+ (CC-125) from the Chlamydomonas Collection (http://chlamycollection.org/) and CMJ030 from Dr Martin Jonikas (Carnegie Institution for Science, Stanford, CA; Zhang et al., 2014). The strain CMJ030 is pyrenoid-positive and CCM-positive (this study, unpublished data).

For experiments investigating the co-regulation of the pyrenoid and CCM (Chapters 4 and 5), a set of near isogenic RBCS mutant strains was used. Strains with higher plant RBCSs (spinach, Arabidopsis or sunflower) or higher plant RBCS α-helices A and B (reciprocal) are all pyrenoid-negative. Strains with native Chlamydomonas RBCS (wild-type) or Chlamydomonas α-helices A and B (helixAB) are pyrenoid-positive (Genkov et al., 2010; Meyer et al., 2012; see also Fig. 4.1).

All algal strains were maintained in the dark on Tris-acetate (Spreitzer and Mets, 1981) medium 1.5% (w/v) agar plates supplemented with Kropat's trace elements (Kropat et al., 2011). Assays and extractions were performed on algae grown in liquid cultures in an Innova 42 incubator (New Brunswick Scientific, Enfield, CT, USA) at 25 °C with 50-100 µmol photons m^{-2} s^{-1} illumination and shaking (125 rpm). Starter cultures were inoculated into Tris-minimal medium (Tris-acetate without the acetate) from freshly replated strains. Experimental cultures were inoculated from starter cultures and harvested at mid-log phase (approximately 1-2 x 10^6 cells ml^{-1}).

Synchronised cultures were grown in 12 h:12 h dark:light cycles and harvested at mid-log phase during the third light period after inoculation. Cell concentration was determined at several points during the light period for three dark:light cycles in order to confirm that these conditions induced synchronicity. High CO_2 cultures

were bubbled with 5% (v/v) CO_2 for six days (two consecutive cultures) and low CO_2 cultures were harvested after bubbling with air (3 h for analysis of gene and protein expression). In all experiments, high CO_2 denotes air supplemented with 5% (v/v) CO_2 and low CO_2 denotes ambient levels of CO_2 i.e. 0.04% (v/v).

2.3 Physiological techniques

2.3.1 Qualitative growth analysis

Spot tests were performed to assess the relative growth of strains at low versus high CO_2 in order to infer CCM phenotype. Aliquots of cells (approximately 5 x 10^4) in Tris-minimal medium were spotted onto Tris-minimal plates and incubated at either high or low CO_2 in the light (30-50 µmol photons m^{-2} s^{-1}) for two weeks. Control aliquots were also spotted onto Tris-acetate plates, which were incubated in the dark and used to assess dilutions and cell viability.

2.3.2 Quantitative growth analysis

Growth curves were performed to quantify the high CO_2-requiring phenotype of Rubisco higher plant RBCS hybrid strains versus wild-type in a Multi-Cultivator MC 1000-OD connected to a Gas Mixing System GMS 150 (Photon Systems Instruments, Brno, Czech Republic). Test tubes of Tris-minimal medium (60 ml) were inoculated from Tris-acetate starter cultures to an optical density (OD) of 0.05 at 680 nm and allowed to grow for 48 h to mid-log phase (0.6-1.0 OD_{680nm}) at 25 °C. Cultures were then diluted to 0.05 OD_{680nm} and grown to mid-log phase a second time to ensure cells were acclimated to new growth conditions. Cultures were bubbled with either 5% (v/v) CO_2 or air and illuminated with 50 µmol photons m^{-2} s^{-1}. Optical density measurements at 680 and 730 nm were recorded every 10 min. Doubling times were calculated from the slope of the curve of time versus log_2OD for cells in the mid-log (linear) phase of growth.

2.3.3 Photosynthetic affinity for Ci as determined by oxygen evolution

Apparent affinity for Ci was determined using the oxygen evolution method described by Badger et al. (1980). Cells grown in Tris-minimal liquid medium were harvested by centrifugation at 3,500 x *g* for 5 min at 20 °C and resuspended in 25 mM HEPES-KOH (pH 7.3) to a density of 1.5 x 10^7 cells ml^{-1}. Aliquots of cells (1

ml) were added to a Clark type oxygen electrode chamber (Rank Brothers, Cambridge, UK) attached to a circulating water bath set to 25 °C. The chamber was closed and cells were illuminated with 200-300 µmol photons m^{-2} s^{-1} and allowed to consume internal Ci stores. When net oxygen evolution ceased, aliquots of sodium bicarbonate were added to the cells at 30-s intervals and the rate of oxygen evolution was recorded every second using a PicoLog 1216 data logger (Pico Technologies, St Neots, UK). Cumulative concentrations of HCO_3^- after each addition were as follows: 2.5, 5, 10, 25, 50, 100, 250, 500, 1000 and 2000 µM. To calculate $K_{0.5}$ values, the equation describing Michaelis-Menten kinetics was fitted to curves of external Ci versus photosynthetic rate and solved for $K_{0.5}$ (K_m) and maximum rate (V_{max}) using the least squares method. Some of the RNA and protein extraction and quantitation methods were conducted after 25 min illumination in the electrode chamber to mimic a standard light pre-treatment plus Ci addition protocol.

2.4 Chlorophyll extraction

Chlorophyll was extracted from cells for normalisation of oxygen evolution measurements (section 2.2.3) and loading of proteins for SDS-PAGE and immunoblots (section 2.4.2). After oxygen evolution measurements, cells were harvested by centrifugation at 6,200 x *g* for 5 min at 20 °C, resuspended in 90% (v/v) acetone and incubated for 10 min at 50 °C. For determination of chlorophyll concentration in protein extracts, 100 µl protein extract was mixed with 900 µl acetone. For all samples, cell debris was then pelleted by centrifugation at 2,400 x *g* for 5 min at 4 °C and the absorbance of the supernatant was measured at 647 and 664 nm. Chlorophyll concentration was calculated using the equations of Jeffrey and Humphrey (1975).

2.5 Molecular techniques

2.5.1 Analysis of gene expression by qRT-PCR

Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) was used to determine the relative abundance of CCM gene transcripts. Total RNA was extracted from 1 x 10^7 cells using TRIzol Reagent (Life Technologies, Carlsbad, CA, USA). RNA quality was verified by agarose gel electrophoresis,

while RNA quantity and purity was measured using a NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA). Complementary DNA was synthesised from 500 ng total RNA using SuperScript II reverse transcriptase (Life Technologies), RNaseOUT (Life Technologies) and oligo$(dT)_{18}$ primers (Thermo Scientific). Relative gene expression was determined using a Rotor-Gene Q Real-Time PCR Cycler (Qiagen, Venlo, Netherlands). Reactions (10 µl) used SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich, St Louis, MO, USA) and gene expression was calculated relative to the first time point and the reference gene *GBLP* (guanidine nucleotide binding subunit β-like protein, also known as Chlamydomonas β subunit-like polypeptide, *CBLP*, or receptor of activated protein kinase C, *RACK1*; Schloss, 1990) according to the method of Livak and Schmittgen (2001). Primers were designed using the web-based tools Primer3 (*CCP1*, *CCP2*, *GBLP*, *LCIA*, *LCIB*, *LCIC*, *rbcL, RBCS*; Rozen and Skaletsky, 2000) or QuantPrime (*CAH1*, *HLA3, LCI1*, *LCR1, RBCS1, RBCS2*; Arvidsson et al., 2008). Control primers were also designed using Primer3 to measure the expression of photosynthesis-associated genes: *GLYK* (glycerate kinase), *LHCB4* (chorophyll a-b binding protein of PSII) and *PSBO* (oxygen-evolving enhancer protein of PSII). Primer sequences are listed in Table 2.1.

Table 2.1 Primers used for qRT-PCR analysis of CCM gene expression. Control primers for *GBLP* and photosynthetic genes (*GLYK*, *LHC4*, *PSBO*) are also listed.

Gene	Forward primer (5' to 3')	Reverse primer (5' to 3')
CAH1	TGTGCACCAGGTGACTGAGAAG	TCGCAAAGATGGGCTCAAGCAG
CCP1	TACTCGTCCACGATGGACTG	ATGTGCTCCACGTTCTCCTC
CCP2	AGTACAGCACTACCATTGAC	ATGTGCTCCACGTTCTCCTC
GBLP	AACACCGTGACCGTCTCC	TGCTGGTGATGTTGAACTCG
GLYK	GCAGTGGGTGTTCAAGTGG	TTCTGGTCCACCTCGATGAT
HLA3	AGAAGCTTAAGGACCAGGATGGC	AGTTGACGTGGGACAGCAGA
LCI1	TCCAGTTCGAGCTGTTTGTGTTCC	AGCAGGAAGAAGATGGCGTTGATG
LCIA	CTCCTCCTCAAGTGTATGAGAACG	CAGAGCAAATAGCAGCTTGG
LCIB	GAGCTGATCAAGCACTTC	CCTCAATCTTGTCCTTCA
LCIC	CGTCCATGGAGTTCATTGC	ATGCGCTTCTCCAGGTAGC
LCR1	GCGCACAGTCCCTTCGTCTAATTC	AGCGATCTCGGTCCAACTGTTAC
LHCB4	TCCCATCTTCACCAACAACA	GGTGGGTCGACAGGTTCTTA
PSBO	GTGGCTTTCCTGTTCACCAT	GCCTTGGTGATCTTCACGTT
rbcL	GGTGACCACCTTCACTCTGG	TCACCGAAGATTTCAACTAAAGC
RBCS	GCGTGTCTTGCCTGTACTACG	CTGCTTCTGGTTGTCGAAGG
RBCS1	AAGCGTGATCGCATGACGCAAG	ACCGCATAACCTCAGAACACCAG
RBCS2	ATACTGCTCTCAAGTGCTGAAGCG	AAAGACTGATCAGCACGAAACGG

2.5.2 SDS-PAGE and detection of proteins using immunoblots

Immunoblots were used to detect CCM proteins in both wild-type and RBCS mutant strains to determine whether changes in mRNA abundance were mirrored at the protein level. Cells (2.5×10^7) were harvested by centrifugation at 3,500 x *g* for 5 min at 4 °C, resuspended in 500 µl ice-cold protein extraction buffer (50 mM Bicine pH 8.0, 10 mM $NaHCO_3$, 10 mM $MgCl_2$, and 1 mM dithiothreitol) and sonicated in 15 ml Falcon tubes on high power for 15 min at 4 °C using a Diagenode Bioruptor Standard (Diagenode, Liège, Belgium). Optimal lysis conditions were previously determined by visually inspecting samples under a light microscope. Unbroken cells were pelleted by centrifugation at 4,400 x *g* for 10 min at 4 °C and the supernatant (soluble protein fraction) was stored at -80 °C until analysis. Protein levels were quantified using Bradford reagent (Sigma-Aldrich; Bradford, 1976) and the crude protein fraction (1.5 µg total chlorophyll or 5 µg total protein) was separated on a 5% (w/v) stacking and 12% (w/v) resolving

polyacrylamide gel using SDS-PAGE (Laemmli, 1970). Gels were prepared and run according to the Mini-Protean 3 Cell instruction manual (Bio-Rad, Hercules, CA, USA). Total protein was visualised using GelCode Blue Stain (Thermo Scientific).

For detection of CAH1 in wall-less RBCS strains, growth medium (10 ml) was collected after cells were pelleted and was concentrated using Amicon Ultra-15 centrifugal filter units (10,000 MWCO, EMD Millipore, Billerica, MA, USA). Protein levels in the supernatant were too low to be detected using the Bradford assay so growth medium was concentrated to equal volumes (approximately 500 µl) before loading aliquots (24 µl) on SDS-PAGE gels.

For immunodetection of specific CCM components, proteins were transferred electrophoretically to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad; Towbin et al., 1979). Non-specific binding sites on membranes were blocked by incubation in Tris buffered saline (TBS) plus 5% (w/v) skim milk powder for 1 h at 20 °C. Membranes were probed with rabbit primary antibodies and donkey secondary antibodies (enhanced chemiluminescence anti-rabbit IgG, horseradish peroxidase-linked whole antibody, GE Healthcare, Amersham, UK) diluted in TBS plus 2.5% (w/v) skim milk powder. Membranes were washed after each incubation with TBS plus 0.1% (v/v) Tween20. Primary and secondary antibody incubations were for 1 h at 20 °C. Membranes were stripped using Restore Plus Western Blot Stripping Buffer (Thermo Scientific) and reprobed to allow the detection of multiple proteins on a single membrane.

The Rubisco and PSBO antibodies were gifts from Prof. John Gray (University of Cambridge; 1:5,000 and 1:2,000 dilutions, respectively), LCIB and LCIC antibodies were from Asst. Prof. Takashi Yamano (University of Kyoto; 1:5,000 and 1:10,000 dilutions, respectively) and the CAH1 antibody was from Prof. Martin Spalding (Iowa State University; 1:1,500 dilution). Histone H3 antibody (abcam, Cambridge, UK) was used as a loading control (1:2,000 dilution).

2.6 Proteomic analyses

2.6.1 Two-dimensional differential in gel electrophoresis (2D-DIGE) analysis of soluble proteins

Triplicate cultures of wall-less wild-type and spinach RBCS hybrid strains, adapted to low CO_2 for 3 h, were harvested and lysed in 2D gel protein extraction buffer (6 M urea, 2 M thiourea, 4% [w/v] CHAPS, $Mg(CH_3COO)_2$, 10 mM Tris pH 8.5) by freeze-thaw cycles. Unbroken cells and cell debris were pelleted by centrifugation at 22,000 x *g* for 10 min at 4 °C.

The supernatants (crude soluble protein fractions) were submitted to the Cambridge Centre for Proteomics (Department of Biochemistry, University of Cambridge) for subsequent preparation and analysis, according to the method described below.

Samples were vortexed and sonicated for 1 min on high power in a cold water bath followed by a 1 min break for a total of 15 min sonication. Proteins were then precipitated by adding 5 volumes of 10% (v/v) cold trichloroacetic acid in acetone. The samples were incubated on ice for 1 h with intermittent vortexing. Proteins were collected by centrifugation at 16,000 x *g* for 10 min at 4 ^{o}C and the supernatant was discarded. The pellet was washed with 80% (v/v) 0.1 M ammonium acetate in methanol, vortexed or sonicated to break up the pellet and then centrifuged at 16,000 x *g* for 5 min at 4 ^{o}C. The pellet was washed with 80% (v/v) cold acetone and air dried for 5 min.

The pellet was resuspended in extraction buffer (0.1 M Tris HCl pH 8.0, 30% [w/v] sucrose, 2% [w/v] SDS, 1 mM phenylmethylsulfonyl fluoride, 2% [v/v] 2-mercaptoethanol). An equal volume of 1 M Tris-saturated phenol (pH 8.0) was added and the mixture was vortexed on ice for 5 min. The sample was centrifuged at 16, 000 x *g* for 5 min at 4 ^{o}C, the upper phenol phase was collected and the extraction step was repeated once more.

Proteins were precipitated from the phenol phase by adding 5 volumes of 0.1 M ammonium acetate in methanol followed by an overnight incubation at -20 ^{o}C. Samples were centrifuged at 16,000 x *g* for 10 min at 4 ^{o}C and the supernatant was discarded. As above, the pellet was washed with 80% (v/v) 0.1 M ammonium

acetate in methanol, vortexed or sonicated to break up the pellet and then centrifuged at 16,000 x *g* for 5 min at 4 ^{o}C. The pellet was washed with 80% (v/v) cold acetone and air dried for 5 min. Pellets were dissolved by vortexing in the original urea/thiourea extraction buffer.

Proteins (50 µg) were fluorescently labelled and separated in the first dimension by isoelectric focusing using a non-linear immobilised pH gradient strip (pH 3-11) and in the second dimension by 12% (w/v) SDS-PAGE. Sample fluorescence was analysed using DeCyder software (GE Healthcare) and proteins that were differentially expressed between wild-type and spinach RBCS hybrid strains (t-test $p<0.05$, expression ratio > 1.3) were excised from the gel for identification by LC-MS/MS.

2.6.2 Identification of differentially expressed proteins from 2D-DIGE using LC-MS/MS

Mass spectrometric analysis of differentially expressed protein spots was performed by the Cambridge Centre for Proteomics, as described below.

Gel spots were excised from the gels using a 10 ml pipette tip and were placed into a 96 well PCR plate. The gel spots were destained, reduced using dithiothreitol and alkylated using iodoacetamide, and subjected to enzymatic digestion with trypsin overnight at 37 °C. After digestion, the supernatant was pipetted into a sample vial and loaded onto an autosampler for automated liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

All LC-MS/MS experiments were performed using a nanoAcquity UPLC (Waters Corp., Milford, MA, USA) system and an LTQ Orbitrap Velos hybrid ion trap mass spectrometer (Thermo Scientific). Separation of peptides was performed by reverse-phase chromatography using a Waters reverse-phase nano column (BEH C18, 75 µm i.d. x 250 mm, 1.7 µm particle size) at flow rate of 300 nL min^{-1}. Peptides were initially loaded onto a pre-column (Waters UPLC Trap Symmetry C18, 180 µm i.d x 20mm, 5 µm particle size) from the nanoAcquity sample manager with 0.1% (v/v) formic acid for 3 minutes at a flow rate of 10 µL min^{-1}. After this period, the column valve was switched to allow the elution of peptides from the pre-column onto the analytical column. Solvent A was water + 0.1% (v/v)

formic acid and solvent B was acetonitrile + 0.1% (v/v) formic acid. The linear gradient employed was 5-40% (v/v) solvent B in 60 minutes.

The LC eluant was sprayed into the mass spectrometer by means of a New Objective nanospray source. All mass-to-charge ratios (m/z) values of eluting ions were measured in the Orbitrap Velos mass analyzer, set at a resolution of 30,000. Data dependent scans (Top 20) were employed to automatically isolate and generate fragment ions by collision-induced dissociation in the linear ion trap, resulting in the generation of MS/MS spectra. Ions with charge states of 2^+ and above were selected for fragmentation. Post-run, the data were processed using Protein Discoverer (version 1.2., ThermoFisher). Briefly, all MS/MS data were converted to mgf files and these were submitted to the Mascot search algorithm (Matrix Science, London UK) and searched against the *Chlamydomonas reinhardtii* database (RefSeq Jan 2014), using a fixed modification of carbamidomethyl (C), a variable modification of oxidation (M) and in specific cases, phosphorylation (Y,S,T) using a peptide tolerance of 20 ppm (MS) and 0.1 Da (MS/MS). Peptide identifications were accepted if they could be established at greater than 95.0% probability.

2.6.3 Shotgun LC-MS/MS analysis of total proteome

Wild-type and spinach RBCS hybrid strains were grown to mid-log phase, adapted to low CO_2 for 3 h and harvested as described above. Cell pellets were snap frozen in liquid nitrogen and stored at -80 °C until analysis.

Extraction and mass spectrometric analysis of proteins, as well as statistical analyses, were carried out by Dr Metodi Metodiev and Dr Gergana Metodieva (School of Biological Sciences, University of Essex) essentially as described in McKew et al. (2013).

Briefly, proteins were extracted in a buffer containing 2% SDS (w/v), Tris-HCl (50 mM, pH 6.8) and protease inhibitors (Roche Diagnostics, Burgess Hill, United Kingdom). Cells lysates were centrifuged at 16,000 x *g* for 15 min at 4 °C and the supernatant was stored at -80 °C until analysis. Protein samples were digested in-gel with trypsin and the tryptic peptides were extracted, dried and reconstituted. Peptides were separated on a 15-cm-long pooled-tip nanocolumn and analysed by electrospray ionisation-tandem mass spectrometry on a hybrid resolution

LTQ/Orbitrap Velos instrument (Thermo Scientific). This method is described in more detail in Metodieva et al. (2013). Data were processed as described in McKew et al. (2013) using UniProt and the Chlamydomonas genome (Merchant et al., 2007) to identify proteins based on MS/MS spectra.

Protein abundance was quantified in two ways: label-free ion intensity and spectral counts. Label-free ion intensity (a measure of the number of ions of a given m/z detected during a particular time interval) is proportional to the absolute abundance of the protein but the relationship is non-linear and this measure can only be used to compare the same ion species between different samples. The number of peptide-identifying spectra can be linearly correlated with the relative abundance of the protein in the sample but this measure is less sensitive to changes in proteins of low abundance. Different statistical tests must be used for each of these measures because ion intensity is a continuous variable while spectral count is a discrete variable (Schulze and Usadel, 2010, and references therein).

To identify proteins that were differentially expressed between the wild-type and spinach RBCS hybrid strains, a t-test was performed on the label-free intensity values, summed for the two technical replicates, for the three biological replicates of each strain. The t-test was used to identify significant ($p<0.05$) differences between the mean intensities for each protein in wild-type and spinach RBCS hybrid strains in relation to the variation in the data. Imputation was also used to avoid the use of zero values in the analysis but this did not significantly affect the results of the t-test.

Secondly, a G-test (goodness-of-fit test) was performed on the spectral counts generated for each protein. This gave a different, but overlapping, list of differentially expressed proteins. Subsequent normalisation of spectral counts to a control protein (PSBO or GBLP) did not significantly affect the results of the G-test.

2.7 Transmission electron microscopy and immunogold localisation

For fixation of Chlamydomonas cells for electron microscopy, glutaraldehyde (electron microscopy grade) was added to a final concentration of 0.5% (v/v) to cultures prior to harvest by centrifugation at 3,500 x *g* for 5 min at 4 °C. Cells were

then resuspended in 0.5% (v/v) glutaraldehyde and 1% (v/v) H_2O_2 in Tris-minimal medium and fixed for 30 min to 2 h at 4 °C. All subsequent steps were carried out at room temperature (approximately 20 °C).

Samples were postfixed and osmicated for 1 h in 1% (v/v) OsO_4, 1.5% (w/v) $K_3[Fe(CN)_6]$ and 2 mM $CaCl_2$. After osmication, samples were washed four times by harvesting cells (7,000 x *g* for 5 min at 20 °C), resuspending them in distilled water, and incubating them for 5 min. Samples were stained in 2% (w/v) uranyl acetate for 1 h and then washed three times in distilled water as described above.

Samples were then serially dehydrated for 5 min each in 75% (v/v), 95% (v/v), and 100% ethanol followed by two incubations in 100% acetonitrile. Samples were embedded in epoxy resin containing 34% (v/v) Quetol 651, 44% (v/v) nonenyl succinic anhydride, 20% (v/v) methyl-5-norbornene-2,3-dicarboxylic anhydride and 2% (v/v) catalyst dimethyl-benzylamine (Agar Scientific, Stansted, UK). The cells were pelleted (8,000 x *g* for 5 min at 20 °C) and the resin was refreshed four times over 2 d. Samples were degassed and cured at 60 °C for at least 3 h. Sections (50 nm) were obtained with a Leica Ultracut UCT Ultramicrotome (Leica Microsystems) and mounted on bare 300 mesh copper grids for straight imaging or bare 300 mesh nickel grids for immunogold labeling.

For immunogold labelling of proteins, fixed and embedded samples were treated to remove superficial osmium and unmask epitopes (Skepper, 2000). Grids were incubated face down on droplets of 4% (w/v) metaperiodate for 15 min and then washed on droplets of distilled water (6 x 1 min). Samples were then incubated in 1% (w/v) periodic acid for 5 min and washed in distilled water (6 x 1 min). Non-specific binding sites were blocked by incubating grids in Tris buffered saline with 0.001% (v/v) Triton X-100 and 0.001% (v/v) Tween20 (TBS-TT) containing 0.1% (w/v) bovine serum albumin (BSA) for 5 min. Primary antibody incubations were in TBS-TT plus 0.1% (w/v) BSA for 16 h and grids were washed in TBS-TT (3 x 1 min) and then distilled water (3 x 1 min). Grids were incubated for 1 h with secondary antibodies (Goat anti-rabbit 15 nm gold conjugates, BBI Solutions, Cardiff, UK) diluted 1:200 in TBS-TT plus 0.1% (w/v) BSA. Grids were washed with TBS-TT and distilled water (as above) and then dried.

Samples were examined with a Technai G^2 transmission electron microscope at 120 kV (FEI Company, Hillsboro, OR, USA) and imaged with AMT Image Capture Engine software (Advanced Microscopy Techniques, Woburn, MA, USA). Image analysis was performed using ImageJ (Abramoff et al., 2004). The fraction of Rubisco or CAH3 in the pyrenoid was calculated for cells in which the pyrenoid and much of the chloroplast was visible. For each cell, the number of gold particles in the pyrenoid (density multiplied by area) was expressed as a percentage of the total number of particles in the chloroplast (stroma plus pyrenoid). Nonspecific labelling was determined by calculating the density of labelling in the cytosol and this value was deducted from both the pyrenoidal and stromal densities. The Rubisco antibody was used at 1:5,000 dilution. The CAH3 antibody (1:100 dilution) was obtained from Agrisera AB (Vännäs, Sweden).

3 CCM INDUCTION IN SYNCHRONISED WILD-TYPE CELLS DURING THE DARK-TO-LIGHT TRANSITION

This chapter contains material that was published in the article:

Mitchell MC, Meyer MT, Griffiths H (2014) Dynamics of carbon concentrating mechanism induction and protein relocalisation during the dark-to-light transition in synchronised *Chlamydomonas reinhardtii*. Plant Physiology **166:** 1073-1082

3.1 Introduction

Two key environmental signals that modulate CCM induction are light and low CO_2. While low CO_2 induction is the most obvious and well characterised response, there is good evidence of the role of light and photosynthesis in the regulation of CCM activity. In particular, accumulation of Ci in Chlamydomonas cells and chloroplasts depends on light and is driven by ATP, which may be derived from cyclic electron transport (Spalding and Ogren, 1982; Spalding et al., 1984). The inorganic carbon transporter HLA3 is a member of the ABCC superfamily, which suggests that it is also energised by photosynthesis-derived ATP, and *HLA3* transcription increases in response to high light as well as to low CO_2 (Im and Grossman, 2002). Similarly, *CAH1* transcription depends on light, low CO_2 and photosynthetic electron flow (Dionisio-Sese et al., 1990; Kucho et al., 1999). In synchronised cells, many CCM-related gene transcripts show oscillations in abundance during dark/light cycles with transcripts for inorganic carbon

transporters (*CCP1*, *CCP2*, *HLA3*, *LCI1*, *LCIA*) peaking during the light period (Tirumani et al., 2014).

At the protein level, the LCIB-LCIC complex forms a ring around the pyrenoid under low CO_2 in the light but diffuses away when cells are shifted to high CO_2 or the dark. Aggregation near the pyrenoid is blocked by the addition of DCMU, which suggests that the process depends on photosynthesis and an active CCM (Yamano et al., 2010). CAH3 preferentially relocalises to the pyrenoid during CCM induction in asynchronous cells (Blanco-Rivero et al., 2012), and in synchronised cells, CAH3 is also pyrenoid-localised during the light period and dispersed during the dark period (Tirumani et al., 2014).

However, while light is an important regulator of CCM activity, it is not an absolute requirement for the expression of some CCM components. For example, some CCM-related transcripts (*CAH3*, *CAH6*, *LCIB*) peak during the dark period in synchronised cells, although levels of CAH3 protein do not peak until later in the light period (Tirumani et al., 2014). In addition, cells acclimating to low CO_2 induce external carbonic anhydrase activity and active HCO_3^- transport even in the dark, although there is a long lag in carbonic anhydrase activity compared to cells switched to low CO_2 in the light (Bozzo and Colman, 2000).

CCM activity is also reduced in the dark in synchronised cells (Marcus et al., 1986), likely due to the dependence of the CCM on light and photosynthesis. However, the mechanisms governing the repression and induction of the CCM during the dark/light cycle are unknown. Estimates of protein half-lives in Chlamydomonas vary from several minutes to several days (Pootakham et al., 2010; Bienvenut et al., 2011; Mastrobuoni et al., 2012), and we hypothesised that not all CCM components would be degraded during the dark period. So, unlike the CO_2 response, which requires strong transcriptional upregulation and de novo protein synthesis, induction of the CCM during the dark-to-light transition might involve non-transcriptional or non-translational mechanisms.

The majority of experiments investigating CCM induction have avoided the potentially confounding effect of light by analysing asynchronous cultures grown in continuous light. However, this approach may not capture effects of light or the cell cycle on CCM induction. The aim of the present study was to investigate the

dynamics of CCM induction and molecular regulation in synchronised cells grown in dark/light cycles by undertaking detailed time course experiments covering the dark-to-light period transition. The focus was on the dark-to-light transition as this is most analogous to CCM induction in response to low CO_2. Measurements were made of overall CCM activity ($K_{0.5}$ for Ci) as well as the relative abundance of known CCM (CO_2-responsive) gene transcripts and proteins in actively growing cultures. For each of these time courses, cells adapting to low CO_2 were used as a control. The extent that the thylakoid lumenal CAH3 and Rubisco relocalised from the stroma to the pyrenoid during the dark-to-light transition was also determined, since this preferential localisation appears to be required for a fully functioning CCM (Blanco-Rivero et al., 2012; Meyer et al., 2012).

In synchronised cells, CCM activity was initially downregulated in the dark even though CCM-related proteins were relatively abundant at this point. Full CCM activity was subsequently inducible before dawn and also prior to maximum gene transcription and protein accumulation. Rubisco and CAH3 were both found to have relocalised to the pyrenoid, coincident with increased CCM activity. In contrast, unsynchronised cells adapting to low CO_2 showed rapid increases in mRNA levels of CCM genes, followed by increased protein abundance and then increased CCM activity.

3.2 Results

3.2.1 Synchronisation of cells grown in dark/light cycles

In order to confirm the growth conditions required for synchronisation of the cell cycle, cell counts were performed on 12 h:12 h dark/light-grown cultures during the light period until the third light period after inoculation. Counts from duplicate flasks indicated that cells divided during the dark period so that the cell density doubled from one light period to the next and remained constant during the light (Fig. 3.1). Cells reached mid-log phase during the third light period after inoculation and this time was chosen for future harvests.

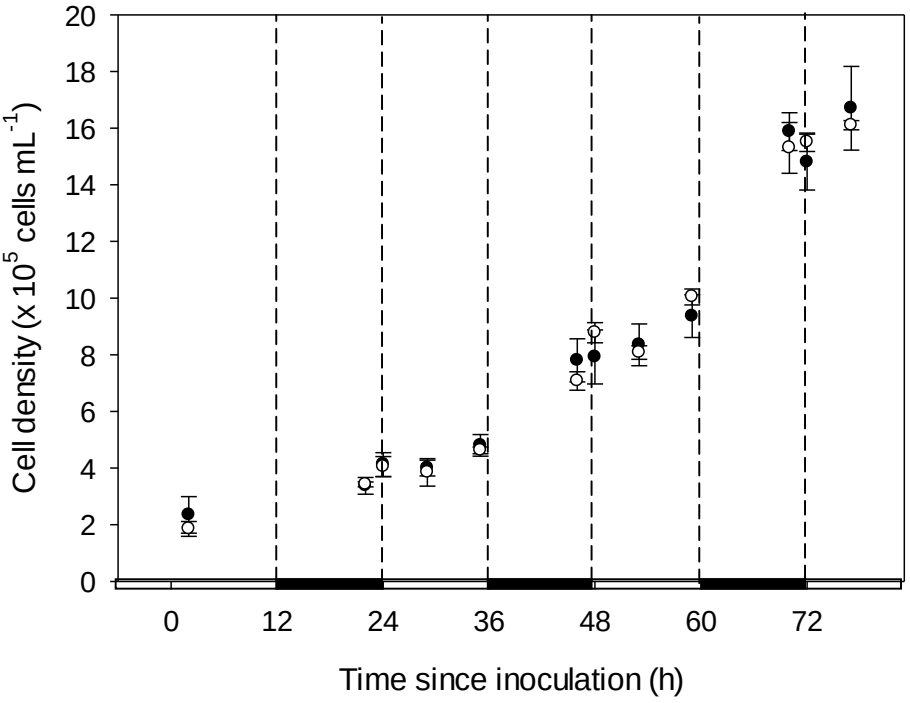


Figure 3.1 Synchronisation of wild-type (2137 mt+) cells grown in 12 h:12 h dark/light cycles. Cells were inoculated at the beginning of the first light period (0 h). Open and closed circles represent mean ± 1 s.e. of counts from two separate flasks.

3.2.2 Determination of photosynthetic affinity for Ci using the oxygen evolution method

Photosynthetic affinity for Ci (a proxy for CCM activity) was determined by adding increasing concentrations of bicarbonate to Ci-depleted cells and measuring rates of oxygen evolution using an oxygen electrode. In high CO_2-adapted cells, high external concentrations of Ci (≥ 1000 µM HCO_3^-) are required to reach maximum photosynthetic rates and this is reflected in the shallow curve of external Ci versus photosynthetic rate (Fig. 3.2, closed circles). In contrast, low CO_2-adapted cells (Fig. 3.2, open circles) are able to photosynthesise efficiently even at low external Ci, which results in a steep initial gradient and cells reaching maximum photosynthetic rates at much lower concentrations of Ci (approximately 250 µM HCO_3^-). The shift in cells' photosynthetic response to external concentrations of Ci when grown at low versus high CO_2 indicates that the CCM has been induced.

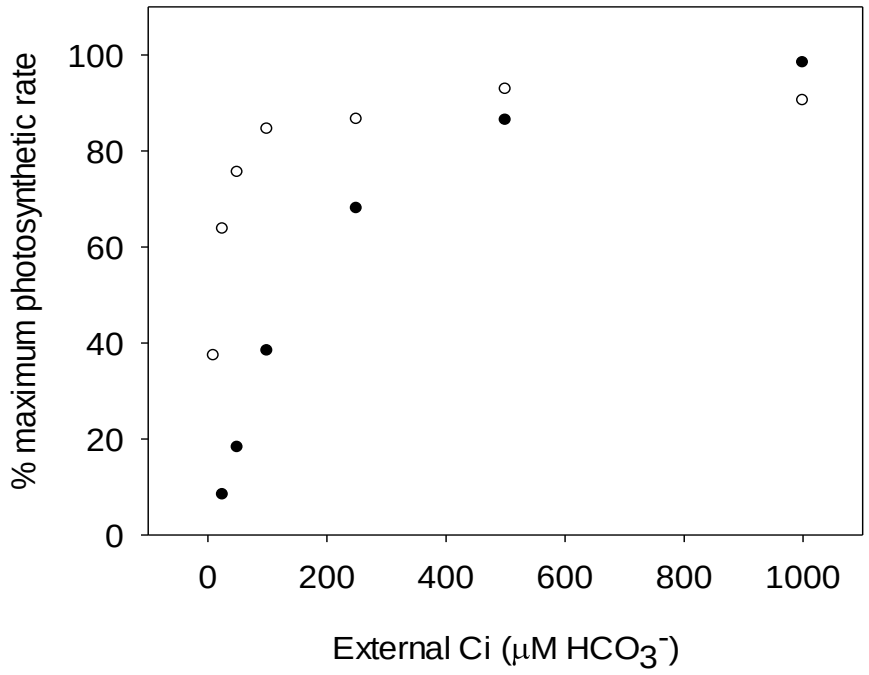


Figure 3.2 Representative oxygen evolution response curves for asynchronous cultures of wild-type strain 2137 mt+. Oxygen evolution rates in response to external levels of Ci were measured for high CO_2-adapted cells (closed circles) and low CO_2-adapted cells (open circles).

Quantification of CCM activity is also possible by calculating the $K_{0.5}$ for Ci, that is, the external concentration of Ci required for cells to reach half maximal rates of photosynthesis. Three different wild-type strains showed 9-fold to 16-fold reductions in $K_{0.5}$ when shifted from high to low CO_2, indicating greatly increased photosynthetic affinity for Ci (Fig. 3.3). While the absolute $K_{0.5}$ values varied slightly from strain to strain, the overall shift in $K_{0.5}$ between high and low CO_2-adapted cells was large, making this a good indicator of relative CCM activity in each strain. The strain used in this set of experiments, 2137 mt+, showed a mean high CO_2-adapted $K_{0.5}$ of 150 µM HCO_3^- and a mean low CO_2-adapted value of 15 µM HCO_3^-.

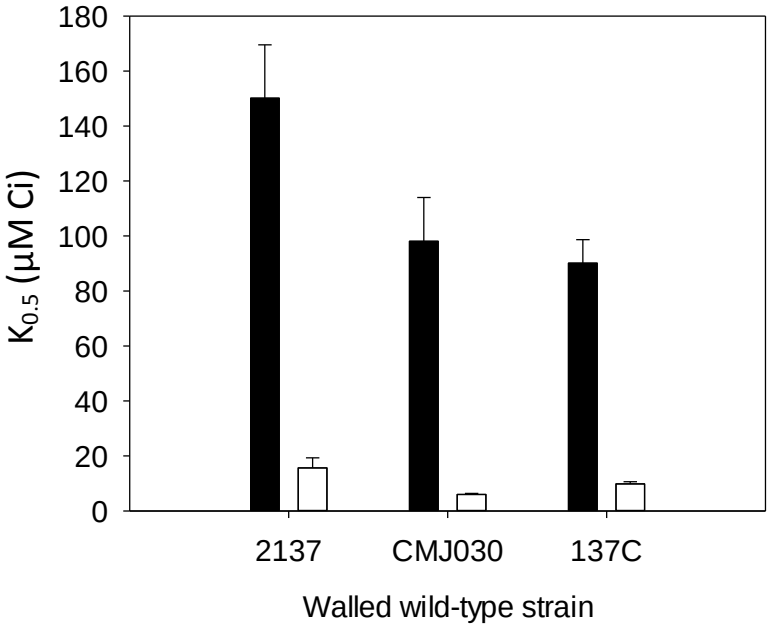


Figure 3.3 Relative affinity for Ci of three walled wild-type strains adapted to high CO_2 (black bars) or low CO_2 (white bars). Cells were adapted to low CO_2 for 11-18 h. Values are mean ± 1 s.e. of 3 or 4 separate oxygen evolution curves.

3.2.3 The CCM is partially repressed in the dark but can be induced one hour before dawn

CCM induction during the dark-to-light transition was measured at the whole cell and molecular level to identify factors affecting CCM regulation in synchronised cultures. In synchronised cells, $K_{0.5}$ for Ci was measured from two hours before dawn to six hours after dawn (Fig. 3.4A). A relatively high $K_{0.5}$ value (60 µM Ci) was found towards the end of the dark period (-2 h), indicating that the CCM was at least partially repressed in the dark, even after the short light pre-treatment required for these measurements. From this time point, CCM inducibility rapidly increased, so that one hour before dawn, cells were able to express a fully functioning CCM (as indicated by a low $K_{0.5}$ of 11 µM Ci). This high affinity for Ci was maintained through dawn and the first six hours of the light period (mean $K_{0.5}$ of 11 to 19 µM Ci).

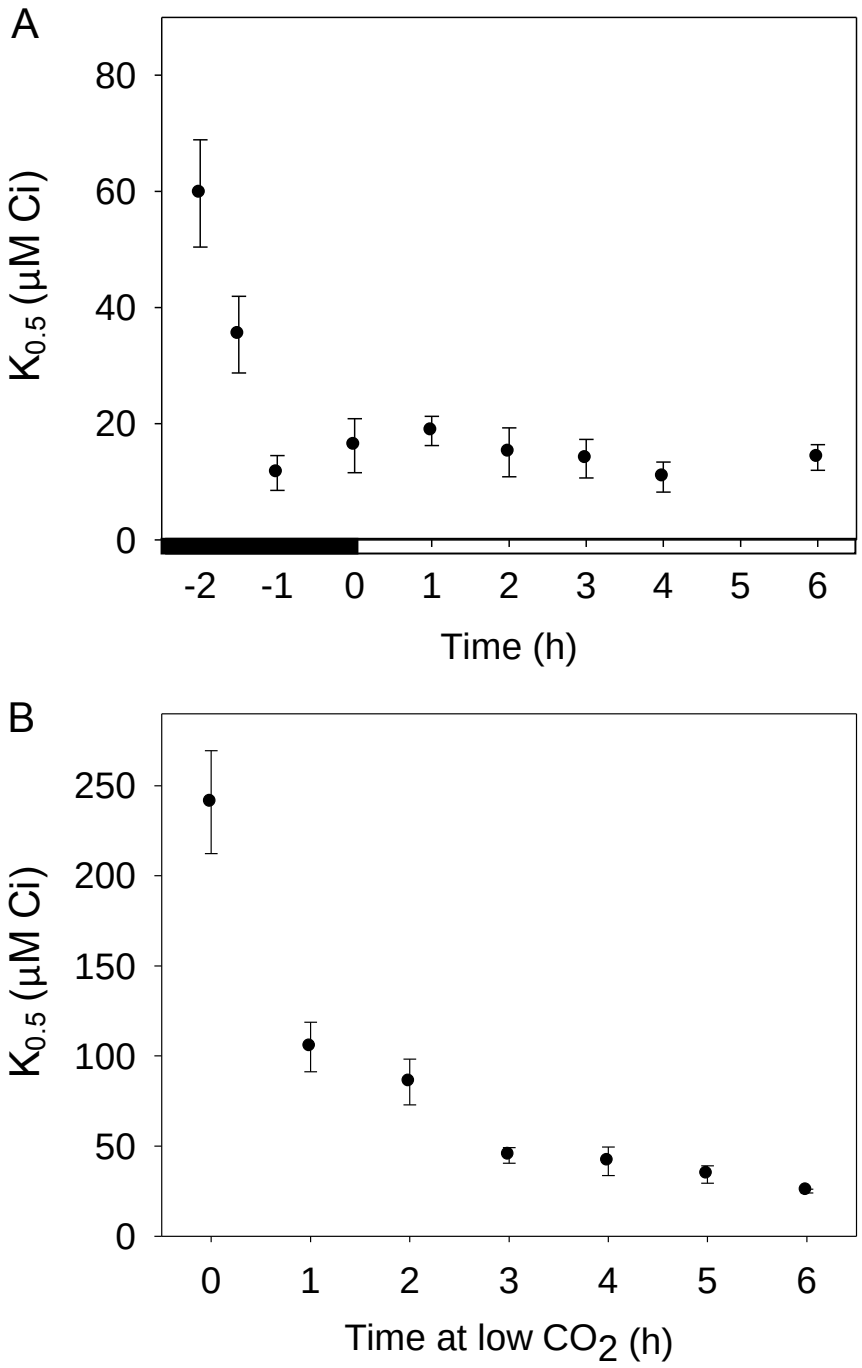


Figure 3.4 Whole cell affinity for Ci ($K_{0.5}$) measured in wild-type cells during CCM induction in synchronised cells during the dark-to-light transition grown continuously at low CO_2 (A) and asynchronous cultures adapting to low CO_2 (B). Synchronised cells were grown in 12 h:12 h dark/light cycles under low CO_2 and harvested during the third dark-to-light transition after dilution (dawn = 0 h). Asynchronous cells were grown to mid-log phase at high CO_2 and harvested following the switch to low CO_2 (t = 0 h). Values are mean ± 1 s.e. of three to five independent experiments.

In order to provide a direct comparison for dark-to-light time courses, CCM induction in asynchronous cultures grown in continuous light and switched to low

CO_2 was also investigated. The very high $K_{0.5}$ value (240 µM Ci) of high CO_2-adapted cells indicated that the CCM was fully repressed at the beginning of the time course (Fig. 3.4B). Affinity for Ci rapidly increased following the transfer to low CO_2, with the measured $K_{0.5}$ reduced to 105 µM Ci within one hour of the shift to low CO_2. Cells showed maximum CCM induction (highest affinity for Ci) after 4-6 h at low CO_2, as indicated by low $K_{0.5}$ values (25-40 µM Ci). The maximum $K_{0.5}$ value measured for high CO_2-adapted cells was four times greater than that observed in dark-adapted (-2 h) cells, as was the overall decrease in $K_{0.5}$ during each time course (ten-fold and six-fold, respectively), which suggests that the CCM was repressed to a greater extent at high CO_2 than before dawn.

3.2.4 Maximum gene expression occurs after CCM induction during the dark-to-light transition

In this comparative study between CCM regulation during the dark-to-light transition and the conventional CO_2 response, qRT-PCR was used to track expression of CO_2-responsive genes. Gene expression was quantified relative to the control gene *GBLP* and to the first time point. In synchronised cells two hours before dawn, *CCP1*, *CCP2*, *LCI1*, *LCIA*, *LCIB* and *LCIC* were all present at similar levels to high CO_2-adapted cells (Table 3.1). *CAH1*, *LCR1* and *rbcL* were present in greater abundance while *HLA3* and *RBCS* were present in lesser abundance in dark/light-grown compared to high CO_2-grown cells. Maximum levels of *CAH1*, *CCP1*, *LCI1*, *LCIA*, *LCIB* and *LCIC* mRNA were not significantly different between dark/light and CO_2 time courses (Table 3.2). Some differences were observed in the maximum abundance of *CCP2, HLA3*, *LCR1*, *rbcL* and *RBCS* but these were below one order of magnitude and smaller than the differences observed at the beginning of the time courses. Overall, this indicates that transcriptional regulation during CCM induction in response to light is similar to that of CO_2.

Indeed, in synchronised cells, many CO_2-responsive genes (*CCP1*, *CCP2*, *LCI1*, *LCIA*, *LCIB* and *LCIC*) were upregulated during the light period to a similar extent to the CO_2 response (Fig. 3.5, Tables 3.3 and 3.4). However, the majority of transcriptional upregulation did not occur until several hours after maximum CCM activity was observed (Fig. 3.4A , Fig. 3.5; see also summary Fig. 3.11 and Table 3.2). This contrasted with the concurrent CCM activation and low CO_2-induced

gene induction observed in cells adapting to low CO_2 (Fig. 3.4B, Fig. 3.6; see also summary Fig. 3.11 and Table 3.2).

Table 3.1 Relative mRNA levels of CCM genes at the first time point of dark/light and CO_2 response time course experiments. Mean ΔCt ($Ct_{(CCM\ gene)} - Ct_{(GBLP)}$) values of three biological replicates are shown with ± 1 s.e. The final column compares ΔCt values for the dark/light time course to the CO_2 response. Transcripts were determined to be present in greater ('>', $p<0.05$), lesser ('<', $p<0.05$) or equal ('=', $p\geq0.05$) abundance in D/L (-2D) cells compared to CO_2 (0 h) cells, as supported by a t-test. Note that a ten-fold increase in mRNA abundance will result in a decrease in Ct of approximately 3.3.

Gene	**D/L time course, ΔCt**	**CO_2 time course, ΔCt**	**D/L vs CO_2 (-2D vs 0 h)**
CAH1	7.90 ± 0.22	12.67 ± 0.66	>
CCP1	13.79 ± 0.80	14.00 ± 0.69	=
CCP2	12.32 ± 0.50	11.25 ± 0.45	=
HLA3	12.56 ± 0.17	8.83 ± 0.96	<
LCI1	10.99 ± 0.11	10.75 ± 0.31	=
LCIA	13.20 ± 0.51	13.66 ± 0.16	=
LCIB	4.24 ± 0.16	5.88 ± 0.61	=
LCIC	8.15 ± 0.22	8.59 ± 0.47	=
LCR1	6.48 ± 0.96	10.15 ± 0.52	>
rbcL	-5.51 ± 0.02	-3.36 ± 0.52	>
RBCS	-5.06 ± 0.05	-5.45 ± 0.16	<

Although CO_2-responsive genes were consistently upregulated in the light, there were systematic differences in the timing of maximum mRNA expression compared to the CO_2 response. For example, inorganic carbon transporter transcripts (*CCP1*, *CCP2*, *HLA3*, *LCI1*, *LCIA*), *LCR1* and *CAH1* reached maximum levels between two and four hours after dawn, whereas *LCIB* and *LCIC* levels were maximal just one hour into the light period (Fig. 3.5, Table 3.2). In contrast, CCM induction following transfer from high to low CO_2 showed a more coordinated response. CO_2-responsive transcripts accumulated rapidly, reaching maximum levels after approximately two hours and then showing a steady decline over the next four hours (Fig. 3.6, Table 3.2).

Table 3.2 Relative mRNA levels of CCM genes at maximum expression during dark/light (excluding the pre-dawn light treatment) and CO_2 response time course experiments. Mean ΔCt ($Ct_{(CCM\ gene)} - Ct_{(GBLP)}$) values of three biological replicates are shown with one standard error of the mean (s.e.). The time when maximum mRNA expression occurred is also shown as Max. (h). The final column compares ΔCt values for the dark/light time course to the CO_2 response. Transcripts were determined to be present in greater ('>', $p<0.05$), lesser ('<', $p<0.05$) or equal ('=', $p\geq0.05$) abundance in D/L compared to CO_2 cells as supported by a t-test.

Gene	D/L time course		CO_2 time course		D/L vs CO_2 (max. exp.)
	ΔCt	Max. (h)	ΔCt	Max. (h)	
CAH1	-3.20 ± 0.17	3	-3.12 ± 0.41	2	=
CCP1	0.34 ± 0.29	4	1.04 ±0.44	2	=
CCP2	8.39 ± 0.27	2	6.94 ± 0.89	2	<
HLA3	6.64 ± 0.26	4	5.19 ± 0.36	4	<
LCI1	5.94 ± 0.26	2	4.39 ± 0.24	2	=
LCIA	0.64 ± 0.34	3	0.85 ± 0.36	2	=
LCIB	3.45 ± 0.27	1	3.99 ± 0.32	2	=
LCIC	3.51 ± 0.10	1	4.52 ± 0.27	2	=
LCR1	1.61 ± 0.18	4	0.64 ± 0.58	2	<
rbcL	-5.76 ± 0.04	-1	-5.36 ± 0.17	5	>
RBCS	-3.80 ± 0.09	6	-5.51 ± 0.02	5	>

Additionally, because $K_{0.5}$ measurements required an initial period of pre-acclimation in the light, gene expression was measured in synchronised cells harvested before dawn after a brief period of illumination in the oxygen electrode chamber (Fig. 3.5: -2L and -1L). This light pre-treatment resulted in the strong upregulation of the Ci transporter genes *CCP1*, *CCP2*, *HLA3*, *LCI1* and *LCIA* but did not affect the expression of *CAH1* and *LCR1*. *LCIB* and *LCIC*, although having a lower level of induction to begin with, also responded to the pre-dawn light exposure.

RBCS (both *RBCS1* and *RBCS2*) mRNA levels remained largely unchanged throughout the dark/light time course, including the pre-dawn light treatment (Fig. 3.5). *rbcL* transcripts appeared to decrease in abundance by up to 80% both in response to the light treatment and from dawn to six hours into the light, however overall abundance of both *RBCS* and *rbcL* transcripts remained very high (at least one order of magnitude higher than the highly abundant reference gene *GBLP*; Tables 3.1 and 3.2).

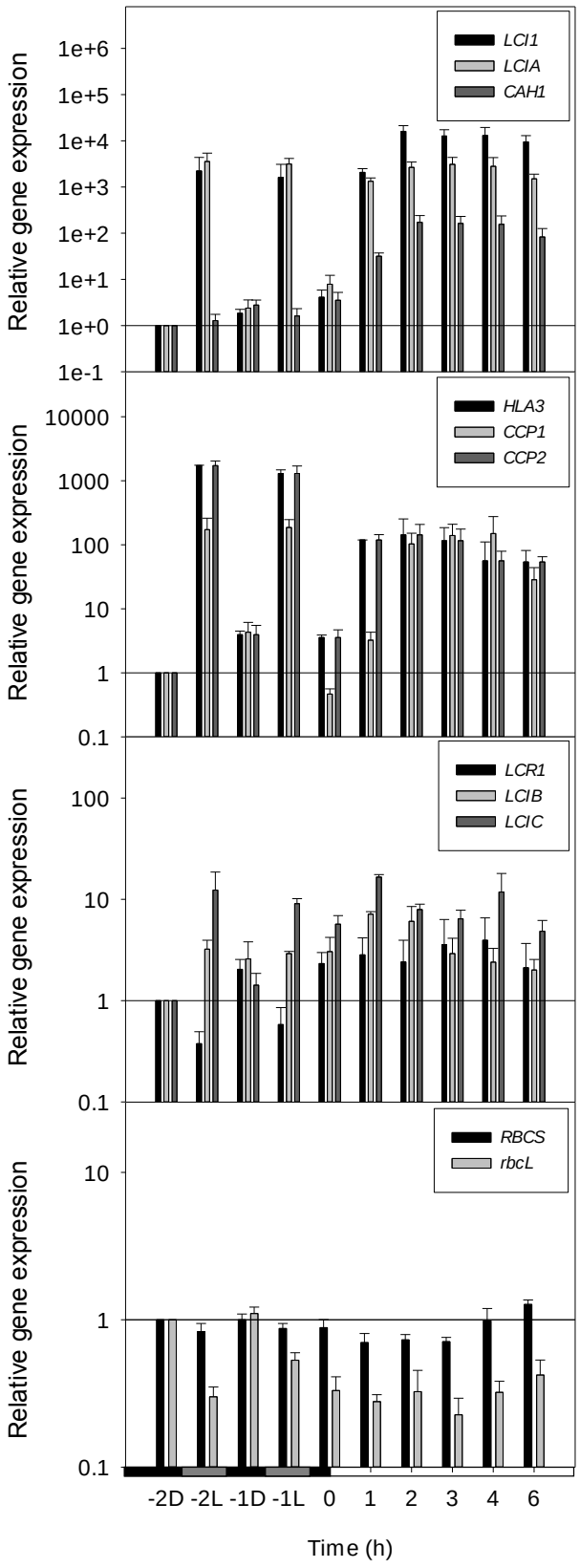


Figure 3.5 Expression profiles of CCM genes in synchronised cells during the dark-to-light transition. mRNA abundance was determined using qRT-PCR and normalised to the first time point (-2D) and to *GBLP* (control gene) expression. Growth and harvest conditions were as described in Fig. 3.4. During the dark period (-2 and -1 h), mRNA and protein were harvested from cells taken either straight from the dark (D) or after a brief illumination in the oxygen electrode chamber (L) to mimic the light pre-treatment necessary for $K_{0.5}$ measurements. Values are mean ± 1 s.e. of three to seven separate flasks harvested across at least three independent experiments.

Table 3.3 Relative expression of CCM genes in synchronised cells during the dark-to-light transition. mRNA abundance was determined using qRT-PCR and values shown are normalised to mRNA levels from the first time point (-2D) cells and to *GBLP* (control gene) expression. See also Fig. 3.5 and Fig. 3.11.

Gene	**Time (h)**								
	-2L	**-1D**	**-1L**	**0**	**1**	**2**	**3**	**4**	**6**
CAH1	1.3	2.8	1.6	3.3	26	130	160	120	64
CCP1	170	3.9	190	0.7	100	140	180	200	120
CCP2	1700	3.4	1300	2.9	90	97	92	61	48
HLA3	61	2	250	1.9	4.1	150	150	200	58
LCI1	2200	1.3	1600	5.9	2300	12,000	10,000	11,000	7000
LCIA	3600	2.1	3200	7.4	970	1800	2400	2000	1100
LCIB	3.2	2	2.9	2.4	5.3	4.4	2.4	1.8	1.4
LCIC	12	1.3	9.1	4.3	13	5.7	5.4	7.7	3.5
LCR1	0.38	2	0.58	2.3	2.8	2.4	3.6	4	2.1
rbcL	0.3	1.1	0.53	0.33	0.28	0.33	0.23	0.32	0.42
RBCS	0.83	1	0.87	0.89	0.7	0.73	0.71	0.98	1.3

Similar to the dark/light time course, only minimal changes were detected in the overall abundance of genes encoding Rubisco (*RBCS* and *rbcL*) throughout the CO_2 time course (Fig. 3.6). Differential expression of *RBCS1* but not *RBCS2* was observed in synchronous but not asynchronous cultures as discussed in section 3.2.7.

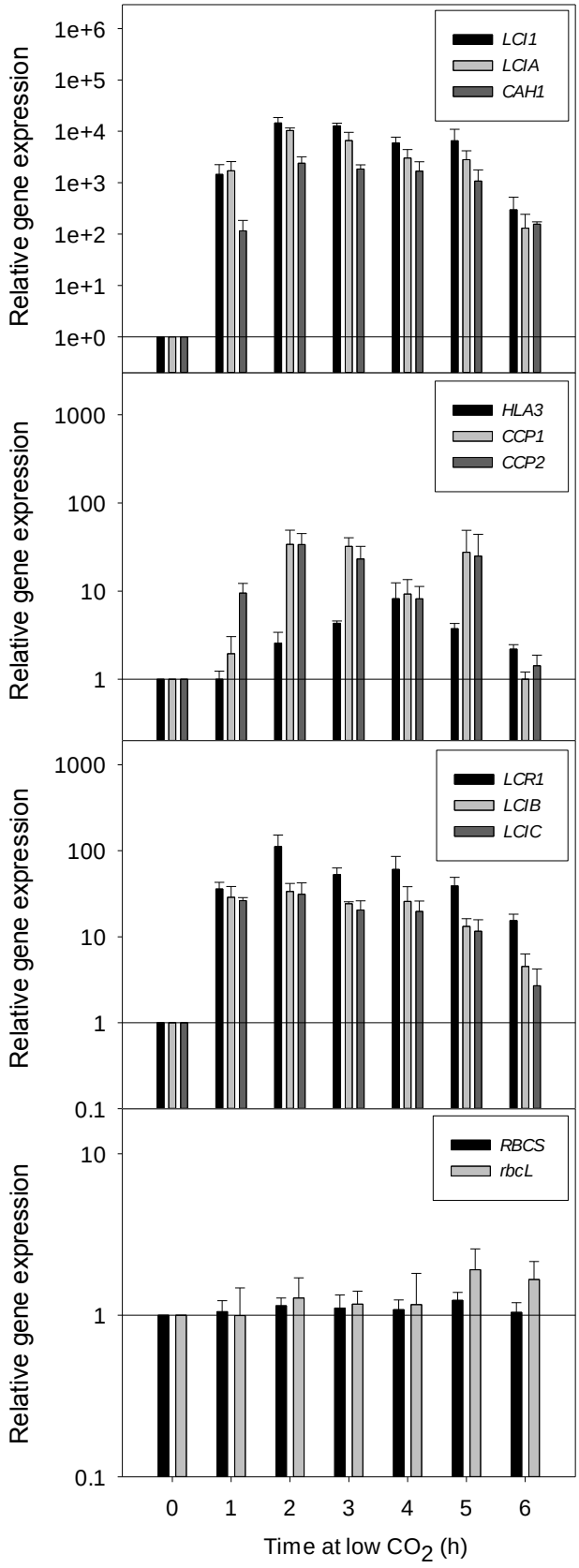


Figure 3.6 Expression profiles of CCM genes in asynchronous cells adapting to low CO_2. mRNA abundance was determined using qRT-PCR and normalised to mRNA levels in high CO_2-adapted cells (0 h) and to control gene (*GBLP*) expression. Values are mean ± 1 s.e. of three separate flasks harvested during a single experiment.

Table 3.4 Relative expression of CCM genes in asynchronous cells adapting to low CO_2. mRNA abundance was determined using qRT-PCR and normalised to mRNA levels in high CO_2-adapted cells (0 h) and to control gene (*GBLP*) expression. Values listed under Bruegg. refer to the relative induction of CCM genes observed in wild-type Chlamydomonas cells after 180 min at very low CO_2 (Brueggeman et al., 2012, Supplemental Dataset 1). See also Fig. 3.6 and Fig. 3.11.

Gene	Time at low CO_2 (h)						Bruegg.
	1	**2**	**3**	**4**	**5**	**6**	
CAH1	120	2400	1800	1700	1100	160	670
CCP1	4.7	29	22	8.6	22	0.86	2000
CCP2	17	32	20	8.2	16	1.4	120
HLA3	0.93	3.4	6	10	5.2	3.4	40
LCI1	1500	15,000	13,000	5900	6500	300	3000
LCIA	1700	10,000	6600	3000	2800	130	4200
LCIB	29	34	24	26	13	4.5	26
LCIC	26	31	20	20	12	2.7	16
LCR1	36	110	53	61	39	15	53
rbcL	1	1.3	1.2	1.2	1.9	1.7	-
RBCS	1.1	1.1	1.1	1.1	1.2	1	-

3.2.5 CCM repression in the dark is independent of key CCM protein abundance

The abundance of several CCM proteins during the dark-to-light transition was measured using immunoblots to determine whether increased mRNA abundance resulted in increased protein levels (Fig. 3.7). To complement qRT-PCR data, pre-dawn samples (-2, -1 h) were harvested either directly from the dark (D) or after light pre-treatment (L). Although CO_2-responsive transcripts were present at low levels two hours before dawn, their corresponding proteins were easily detectable at this time.

LCIB and LCIC were present at high levels two hours before dawn, while CCM activity was still low. These proteins reached maximum levels two to four hours into the light period, several hours after maximum CCM activity occurred. CAH1 was also detectable in the dark and did not show major changes in abundance during the dark-to-light transition. No major change in either rbcL or RBCS abundance was detected during this time course. Pre-dawn exposure to light (-2L and -1L) had minimal effect on the abundance of all proteins probed.

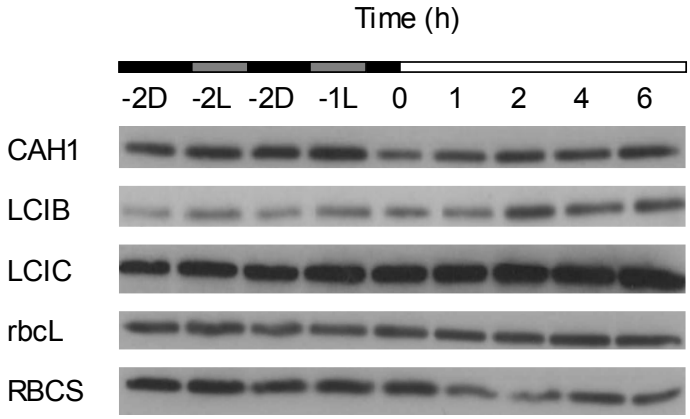


Figure 3.7 Expression profiles of CCM-related proteins in synchronised cells during the dark-to-light transition. Soluble protein extracts were separated using SDS-PAGE and used for immunoblot analyses with antibodies raised against CAH1, LCIB, LCIC and Rubisco (rbcL and RBCS). Sample loading was normalised by chlorophyll content. During the dark period (-2 and -1 h), mRNA and protein were harvested from cells taken either straight from the dark (D) or after a brief illumination in the oxygen electrode chamber (L) to mimic the light pre-treatment necessary for $K_{0.5}$ measurements.

In contrast, for cells adapting to low CO_2, immunoblot analyses of LCIB, LCIC, CAH1 and Rubisco (RBCS and rbcL) showed coordinated increases in mRNA abundance and protein concentration (Fig. 3.8). Unlike in dark-harvested (partially CCM-repressed) synchronised cells, LCIB, LCIC and CAH1 were almost undetectable in high CO_2-adapted cells (0 h). However, all three proteins rapidly accumulated in response to low CO_2, reaching maximum abundance after approximately three (LCIB and LCIC) to five hours (CAH1) and maintaining these high levels for the remainder of the time course. No changes in rbcL or RBCS abundance were observed during the acclimation to low CO_2.

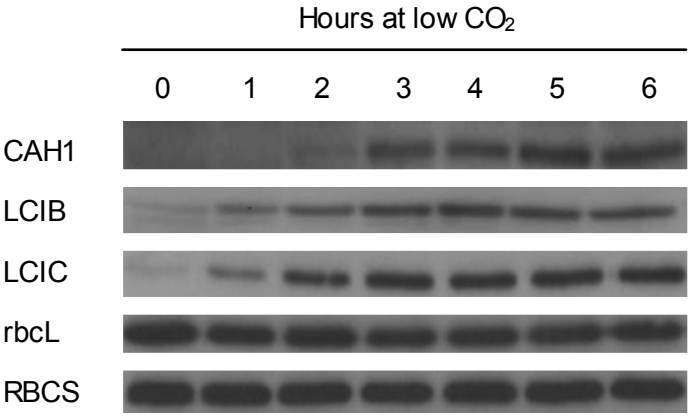


Figure 3.8. Expression profiles of CCM-related proteins in asynchronous cells adapting to low CO_2. Soluble protein extracts were separated using SDS-PAGE and used for immunoblot analyses with antibodies raised against CAH1, LCIB, LCIC and Rubisco (rbcL and RBCS). Sample loading was normalised by chlorophyll content.

3.2.6 Relocalisation of Rubisco and CAH3 to the pyrenoid coincides with CCM induction at the end of the dark period

Immunogold labelling was used to determine the localisation of Rubisco and CAH3, which are both preferentially localised in the pyrenoid in response to low CO_2, in synchronised cells during the dark-to-light transition (Fig. 3.9).

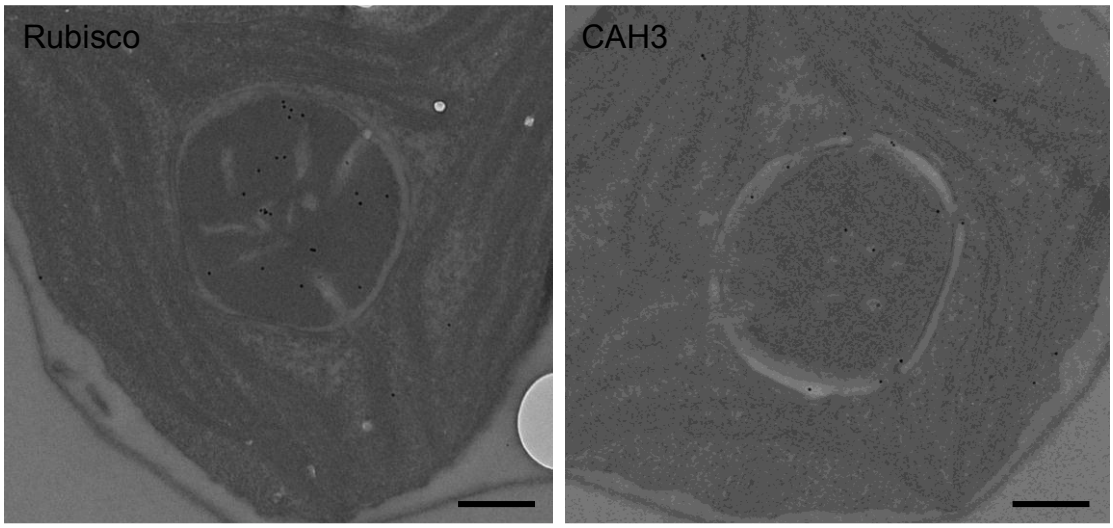


Figure 3.9 Detail of representative transmission electron micrographs showing the pyrenoid and part of the chloroplast of synchronised wild-type cells with immunogold labelling of Rubisco and CAH3 (black dots). Cells were harvested two hours after dawn. Scale bars = 500 nm. Note that Rubisco is present mainly in the pyrenoid matrix whereas CAH3 is present in the lumen of thylakoid tubules.

Two hours before dawn (-2D), approximately 75% of Rubisco-labelled particles were present in the pyrenoid while, one hour later, almost 90% of labelled Rubisco was found in the pyrenoid (-1D) and this high count was maintained two hours into the light period (2 h; Fig. 3.10A). This change in Rubisco localisation coincided with the measured change in $K_{0.5}$ but was independent of the dark period light pre-treatment (-2L and -1L).

In contrast, relative CAH3 localisation changed both before dawn in the dark and in response to pre-dawn light exposure (Fig. 3.10B). Two hours before dawn approximately 22% of CAH3 particles were in the pyrenoid (-2D) but this proportion increased to 35% one hour before dawn (-1D) and reached 40% two hours into the light period (2 h). The percentage of CAH3 in the pyrenoid also responded directly to the pre-dawn light treatment (increasing to 33% at -2L and 39% at -1L).

In general, mean cross-sectional cell area increased in response to light exposure both before and after dawn (Fig. 3.10C). Pyrenoid area also increased in line with cell area across the time course (Fig. 3.10D).

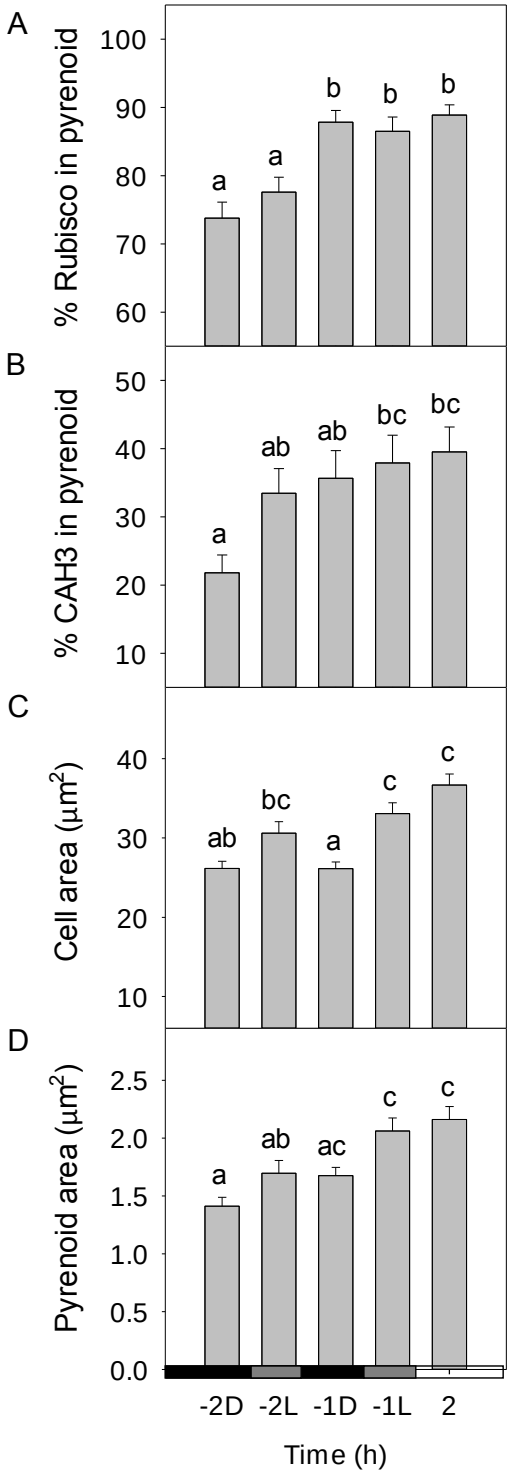


Figure 3.10 Localisation of mobile CCM components in synchronised cells during the dark-to-light transition. Immunogold localisation was used to determine the relative abundance in the pyrenoid of primary photosynthetic carboxylase and pyrenoid component, Rubisco (A) and thylakoid lumen-localised carbonic anhydrase, CAH3 (B; n=53). Cell area (C) and pyrenoid area (D) were also determined (n=80). Bars represent mean ± 1 s.e. Means with different letters are significantly different (Kruskal-Wallis ANOVA and Tukey test, $p<0.05$).

3.2.7 Differences in CCM induction in synchronous compared to asynchronous cultures

Synchronised cultures are able to induce a CCM before dawn and before maximum accumulation of mRNA and proteins, whereas, in asynchronous cultures, CCM induction follows mRNA and protein synthesis (Fig. 3.11). In synchronised cells at the end of the dark period, LCIB and LCIC are present at relatively high levels and steadily increase throughout the time course while mRNA levels peak one hour into the light period. *CAH1* mRNA follows a similar pattern, peaking three hours into the light period, while the protein levels remain relatively constant. Asynchronous cells, on the other hand, rapidly accumulate *LCIB*, *LCIC* and *CAH1* transcripts as well as proteins from very low levels and protein accumulation reaches a maximum shortly after mRNA levels peak.

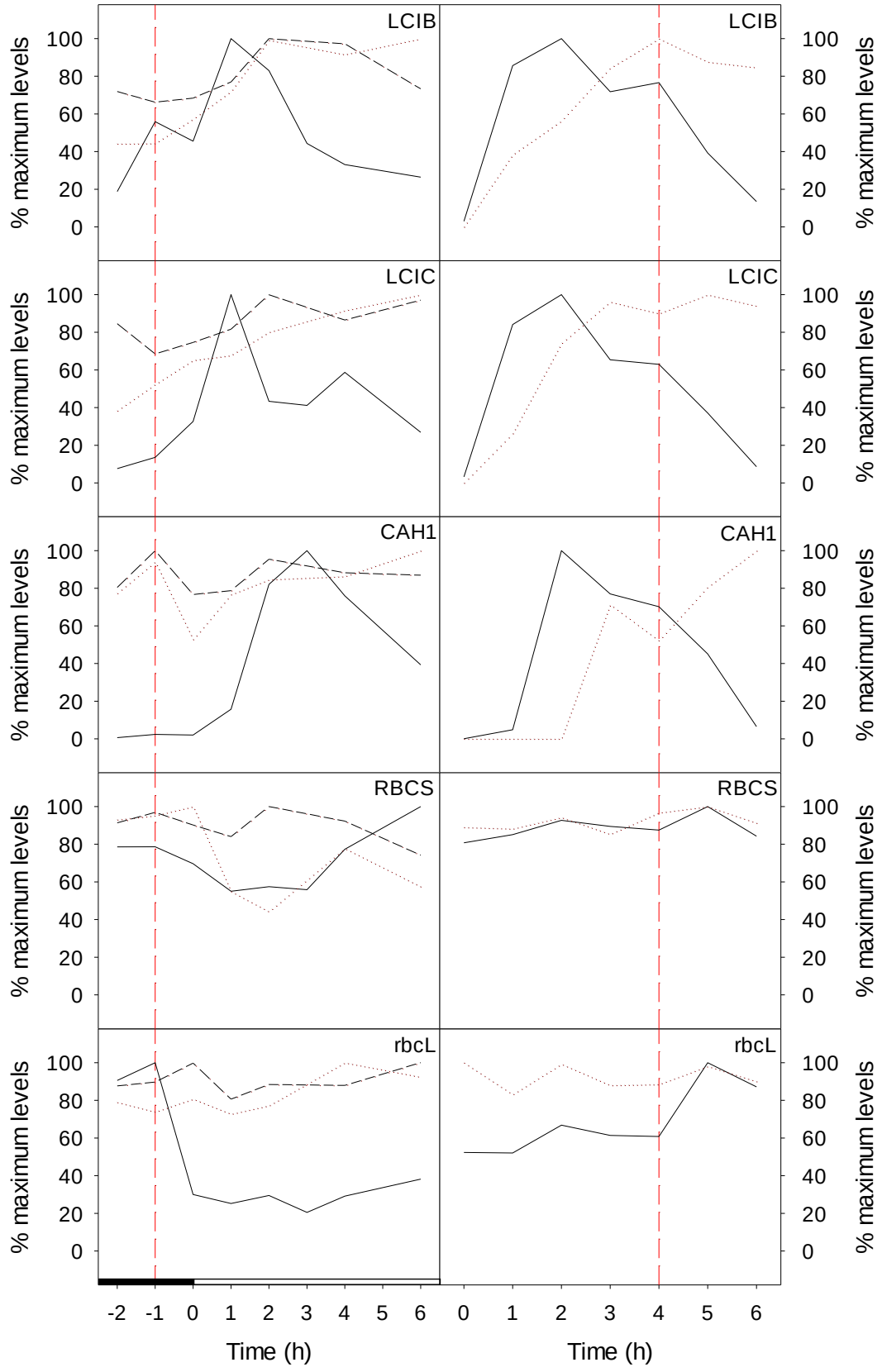


Figure 3.11. Summary of relative changes in mRNA (solid line) and protein normalised by chlorophyll (dotted line) accompanying CCM induction in (left) synchronous cultures during the dark-to-light transition and (right) asynchronous cultures adapting to low CO_2. The black dashed line (left) represents relative protein levels determined by protein-normalised SDS-PAGE/immunoblots. The vertical red dashed lines indicate the time at which CCM activity reached maximum levels as determined by oxygen evolution measurements (Fig. 3.4).

In contrast to the inducible components of the CCM, Rubisco mRNA and protein levels remain relatively high throughout both dark/light and CO_2 response time courses. However, some differences in the expression patterns of the two genes encoding the Rubisco small subunit, *RBCS1* and *RBCS2*, were observed in synchronous cultures (Fig. 3.12). *RBCS1* abundance increased up to 12-fold during the light period whereas *RBCS2* abundance did not change (Fig. 3.12A). However, this did not greatly affect overall *RBCS* abundance because *RBCS1* was much less abundant than *RBCS2* throughout the time course, only accumulating to between 1% and 11% of *RBCS2* levels (Fig. 3.12B). No differences in the abundance or expression patterns of *RBCS1* compared to *RBCS2* were observed in asynchronous cultures (data not shown).

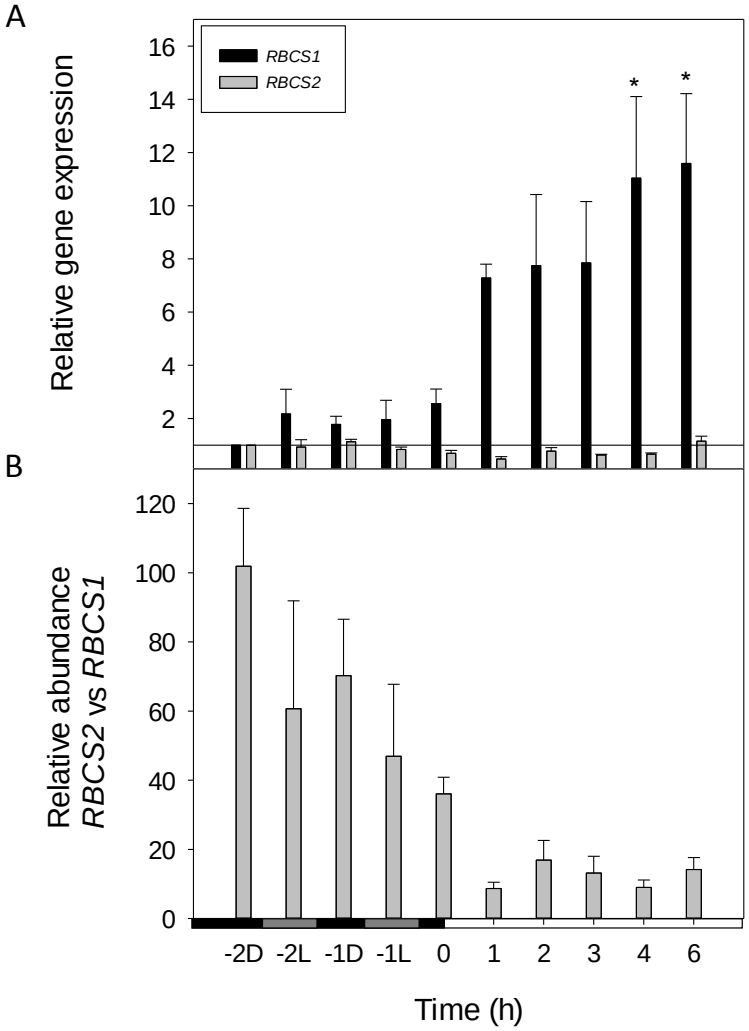


Figure 3.12. Expression of genes encoding the Rubisco small subunit, *RBCS1* and *RBCS2*, in synchronised cells during the dark-to-light transition. Abundance (fold-change) of *RBCS1* and *RBCS2* relative to expression levels at the first time point (-2D; A) and abundance of *RBCS2* relative to *RBCS1* (fold-difference) across the same time points (B) are shown. Bars are mean ± 1 s.e. of three separate cultures. Asterisks indicate significantly different mRNA levels compared to the first time point (-2D; one way ANOVA, $p<0.05$).

3.3 Discussion

Since Marcus et al. (1986) identified oscillations in both CCM activity and photosynthetic characteristics in synchronous cells, the Chlamydomonas CCM has been characterised to a much greater extent in asynchronous cells (see reviews: Moroney and Ynalvez, 2007; Wang et al., 2011). The aim of this study was thus to reconcile changes at the molecular level with regulation of overall CCM activity and cellular dynamics of key CCM constituents in a detailed time course during the dark-to-light transition. Several main differences in CCM induction were observed in synchronous, as compared to asynchronous cultures, as well as in the differential expression of known CCM genes in response to light. Since the CCM was inducible before dawn, without concomitant increases in mRNA or protein levels of known CCM components, this study highlights the importance of additional regulatory mechanisms.

3.3.1 Comparison of CCM induction in response to light and CO_2

This study systematically characterised an eight hour period covering the dark-to-light transition as well as the first six hours of acclimation to low CO_2 for comparison. Coordinated increases in CCM mRNA and protein abundance support the idea that CCM induction in response to low CO_2 relies on the de novo synthesis of many components. In contrast, synchronised cells, which divide at the start of the dark period, can induce CCM activity in advance of the light period and prior to increased expression of CCM gene transcripts and proteins. Some CCM gene transcripts, mostly encoding Ci transporters, were also found to respond differentially to pre-dawn light exposure.

The genes analysed by qRT-PCR in the current study (*CAH1*, *CCP1*, *CCP2*, *HLA3*, *LCI1*, *LCIA*, *LCIB*, *LCIC*, *LCR1*) have all been assigned putative roles in the CCM and identified as transcriptionally upregulated in response to low CO_2, in at least three of four microarray or transcriptomic studies published to date (Supplemental Dataset 1 of Brueggeman et al., 2012). Maximum abundance of these gene transcripts occurred after two hours at low CO_2 and the average fold-change from high CO_2 levels was generally within five-fold of the values obtained by Brueggeman et al. (2012) (see also Table 3.4).

In synchronised cells, the level of gene induction (fold-increase) during the light period was generally within one order of magnitude of that observed in cells adapting to low CO_2 (Tables 3.3 and 3.4). Perhaps more importantly, the maximum levels of mRNA abundance (relative to *GBLP*) were also within 4-fold of values obtained for the CO_2 time course (Table 3.2). Strong upregulation of CCM gene transcripts was thus observed in synchronised cells but, crucially, this occurred only during the light period. Maximum mRNA levels were reached approximately two hours into the light period when cells had been expressing maximum CCM activity for the previous three hours. Tirumani et al., (2014) also showed that Ci transporter transcripts were upregulated in the light, although gene expression for carbonic anhydrases (CAH3, CAH6) and LCIB was higher in the dark.

In the CO_2 response, de novo CCM protein expression was correlated with a large decrease in $K_{0.5}$. qRT-PCR and immunoblot analyses for LCIB and LCIC showed maximum protein abundance occurring approximately one hour after maximum mRNA abundance in both CO_2 and dark/light time courses. However, LCIB and LCIC were only slightly induced in the light period compared to the large induction observed in response to low CO_2. In contrast, CAH1 levels showed no increase during the light period despite increased transcription, which again distinguishes this response from the large increases in transcription and translation observed in response to low CO_2. Overall, during the dark-to-light transition, the minimal induction of CAH1, LCIB and LCIC and the relatively high protein content in the dark were consistent with both the lower $K_{0.5}$ and the inducibility of the CCM in the dark (Fig. 3.4 and Fig. 3.7).

3.3.2 Differential expression of *RBCS1* and *RBCS2* genes in response to light and CO_2

The clear transcriptional response of *RBCS1* to the light period in synchronised cells is similar to previous observations. Hwang and Herrin (1994) showed fluctuations in both *RBCS1* and *RBCS2* abundance during dark/light cycles, although further investigation did not show *RBCS2* expression to be truly circadian and *RBCS1* levels were too low to draw conclusions. Additional evidence that *RBCS1* but not *RBCS2* is likely to be under the control of circadian rhythms is that the circadian clock-controlled RNA-binding protein CHLAMY1 binds *cis*-elements

in the 3' UTR of *RBCS1* (Waltenberger et al., 2001). While the biological relevance of this differential regulation may be worth pursuing, the timing of increased *RBCS1* transcription does not support a direct role in Rubisco relocalisation.

In asynchronous cells, mRNA accumulation of *RBCS1* and *RBCS2* was found to be largely independent of CO_2 availability, when cells were acclimated to low CO_2 for up to 6 h. This is consistent with previously published Northern blot data (Winder et al., 1992) as well as transcriptomic studies. Fang et al. (2012) found that *RBCS1* showed a two-fold decrease from high to low CO_2, clustering with genes involved in cell redox/cellular homeostasis, while *RBCS2* abundance increased by a mere 10%, and was not assigned to a cluster of significantly differentially expressed genes. A second transcriptomic study of CCM induction also found both *RBCS1* and *RBCS2* to be minimally affected by low CO_2 (Brueggeman et al., 2012).

The effect of any altered accumulation of Rubisco-encoding gene transcripts on holoenzyme accumulation is unclear. Winder et al. (1992) detected a transient decrease in RBCS and rbcL synthesis during CCM induction (between 4 and 8 h after the switch to low CO_2). However, the current study did not detect differences in RBCS or rbcL abundance during the first six hours of adaptation to low CO_2 conditions. Overall, this suggests that altered transcription of *RBCS1* is not likely to be related to CCM activation or pyrenoid formation

3.3.3 Transcriptional response to light pre-treatment and inducibility of the CCM in the dark

Both the timing and speed of CCM induction in synchronised cells suggest that regulation during the dark-to-light transition may not depend on the de novo synthesis of known CCM proteins, in contrast to the low CO_2 response. The differential regulation of mRNA expression and protein levels was further illustrated by the light pre-treatment used for comparability with the protocol for $K_{0.5}$ measurements. Transcripts for inorganic carbon transporters (*CCP1*, *CCP2, HLA3*, *LCI1*, *LCIA*) all responded strongly to the light pre-treatment, generally reaching levels at least as high as those achieved one hour into the light period of a conventional dark/light cycle. *LCIB* and *LCIC* transcripts, although induced to a far lesser extent during the light period, showed a similar level of induction under

pre-dawn illumination. The speed and magnitude of this transcriptional response to light may be in part due to the greater light levels in the oxygen electrode chamber (photon flux density 200-300 µmol photons m^{-2} s^{-1}) compared to growth conditions in the incubator (50 µmol photons m^{-2} s^{-1}). For example, mRNA levels of a CO_2-responsive mitochondrial carbonic anhydrase (CAH4) increase linearly with increasing light intensity up to 500 µmol photons m^{-2} s^{-1} (Eriksson et al., 1998), well above the intensities used in this study.

However, not all CCM-related gene transcripts showed the same pattern of induction in response to pre-dawn light exposure. *LCR1*, although slightly induced during the light period, did not respond to light before dawn. More notably, *CAH1* mRNA levels also did not change in response to pre-dawn light exposure, despite being strongly upregulated during the following light period. This indicates that *LCR1* and *CAH1* transcription may be regulated in a different manner to other CO_2-responsive genes.

The evidence presented thus far suggests that the increased affinity for Ci ($K_{0.5}$) and inducibility of the CCM in advance of the light period is uncoupled from the transcriptional upregulation of key CCM elements which occurred in the light, as distinct from low CO_2-induced CCM activity. Secondly, the transcriptional responsiveness of inorganic carbon transporter genes to pre-dawn light treatment, relative to other components, shows there to be differentially regulated elements of the CCM. This, together with the relatively high concentrations of key CCM proteins in cells growing through dark and into light periods, leads us to suggest that post-translational mechanisms might be important for CCM induction in synchronised cells during the dark-to-light transition.

3.3.4 Dynamics of protein abundance and relocalisation during CCM induction across the dark-to-light transition

The energetic costs of synthesising and maintaining the proteins required for a fully functioning CCM may prohibit cells from actively degrading proteins when they are not needed during the dark period. This is also consistent with the limited information on CCM protein stability. For example, CAH1 is still detectable after 48 h at high CO_2 (Ramazanov et al., 1994) even though the transcript is undetectable after only one to two hours (Fujiwara et al., 1990; Yoshioka et al.,

2004). In an analogous system, *Synechocystis* CCM proteins are not actively degraded even when they are no longer needed after cells are shifted to high CO_2 (Zhang et al., 2007).

The relative localisation of mobile CCM components in synchronised cells was investigated as a possible means of modulating CCM activity, independently of protein abundance. Importantly, Rubisco and CAH3 were found to be differentially localised during the dark-to-light transition, although with certain key differences in this response.

Rubisco accumulated within the pyrenoid during the two hours before dawn, coinciding with measured changes in $K_{0.5}$ values and suggesting that the aggregation of Rubisco in the pyrenoid is associated directly with increased CCM activity. Brief exposure to light prior to dawn did not affect Rubisco localisation, which suggests that changes in localisation during the dark-to-light transition may rely on an endogenous (light-independent) signal. Nevertheless, Rubisco is likely to respond to multiple signals because aggregation is also known to occur in response to low CO_2 (Borkhsenious et al., 1998). This is consistent with other species of microalgae in which localisation of Rubisco to the pyrenoid is either circadian-regulated (Nassoury et al., 2001) or associated with irradiance and the active growth phase of cells (Lin and Carpenter, 1997a, 1997b). In *Chlorella vulgaris*, pyrenoid formation also depends on a range of factors including light and low CO_2 as well as reduced electron carriers and cytoplasmic translation (Kuchitsu et al., 1991).

CAH3 localisation, like Rubisco, may also be controlled by an endogenous signal that coordinates CCM induction. Two hours prior to dawn, 20% of CAH3 was localised to the pyrenoid but one hour later the proportion of CAH3 in the pyrenoid had nearly doubled and this high proportion was maintained two hours into the light period. However, in cells harvested two hours before dawn and exposed briefly to light, the proportion of CAH3 in the pyrenoid was also high. This suggests that CAH3 localisation could also be directly regulated by light.

CAH3 relocalisation has also been observed in cells acclimating to low CO_2 and in the current study, CAH3 was present in the pyrenoid at similar relative abundance in dark- versus light-adapted cells, compared to those adapted to high versus low

CO_2 (19% and 37% respectively; Blanco-Rivero et al., 2012). In the CO_2 response, this shift in localisation was linked to the phosphorylation of CAH3 (Blanco-Rivero et al., 2012) but it has not been determined whether a similar mechanism is operating in air-grown synchronised cells. The low CO_2-inducible protein, LCI5, is also rapidly phosphorylated in response to low CO_2 and this has been shown to occur in a redox-dependent manner (Turkina et al., 2006). A similar redox-dependent kinase may be responsible for phosphorylation of CAH3 in response to low CO_2 and light; however, this would not explain the shift in CAH3 localisation in the dark. Whatever the mechanism, the high relative localisation of CAH3 in the pyrenoid after the light pre-treatment two hours before dawn may be partly responsible for the relatively low $K_{0.5}$ value (incomplete repression of the CCM) at this point.

3.4 Conclusions and future work

Although there are similarities in the transcriptional response of genes during CCM induction in response to low CO_2 and during the dark-to-light transition, this study has identified several key differences that distinguish the response of synchronised cells from that of asynchronous cultures. For example, differential regulation of CCM genes suggests that CO_2-responsive and light-responsive regulatory elements and their respective transcription factors should form the basis of future studies. This may enable the identification of gene (regulatory) networks and additional CCM candidates based on the presence of one or more regulatory elements or transcription factor binding sites.

In addition, the inducibility of the CCM before dawn and in advance of mRNA and protein accumulation was directly linked to increased localisation of Rubisco and CAH3 to the pyrenoid and highlights the potential importance of post-translational regulation of the CCM. Further work will be needed to determine the mechanism of relocalisation of Rubisco and CAH3 as well as to elucidate the pathways triggered in response to the different signals. Studies using inhibitors of transcription, translation and post-translational modification could prove informative in this respect. Transcriptomic analyses of synchronised cells might also identify genes involved in pre-dawn regulatory processes.

In any case, studies focussing on how components are modified, including their subcellular localisation and interactions, will be necessary to further our understanding of the CCM. Investigating these changes in synchronised cultures of cells during the dark-to-light transition will provide additional means of probing these important aspects of CCM regulation.

4 REGULATION OF THE CCM IN PYRENOID-LESS RBCS MUTANTS

This chapter contains material that was published in the article:

Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ, Griffiths H (2012) Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. Proceedings of the National Academy of the United States of America **109:** 19474-19479

4.1 INTRODUCTION

Although the pyrenoid has long been identified as a characteristic structure of eukaryotic algal cells, and is associated in most cases with a high affinity CCM (Giordano et al., 2005), there is currently no molecular definition of this chloroplastic microcompartment (Raven et al., 2012). Unlike the ordered proteinaceous shell of the cyanobacterial carboxysome (Rae et al., 2013), the pyrenoid has no discernible outer membrane or protein structure to define its boundary. Early isolation studies identified Rubisco as the major pyrenoid component in several species of algae (for example, Holdsworth, 1971; Kerby and Evans, 1978; Kuchitsu et al., 1988; Okabe and Okada, 1988) but confirmation of the pyrenoid as the site of carbon fixation and not Rubisco storage did not occur until later (Borkhsenious et al., 1998). Crucial to our understanding of the relationship of the pyrenoid to the CCM was the observation that the amount of Rubisco in the pyrenoid is higher in cells grown at low CO_2 compared to those

grown at high CO_2 (Borkhsenious et al., 1998), which suggested that aggregation of Rubisco might also function to increase carboxylation efficiency under CO_2-limiting conditions.

In Chlamydomonas, additional evidence for aggregation of Rubisco in the pyrenoid came from the identification of mutants that were pyrenoid-negative under conditions leading to reduced accumulation of the Rubisco holoenzyme. In one mutant, chloroplast ribosomes are downregulated in the presence of acetate, leading to reduced synthesis of the Rubisco large subunit (Goodenough and Levine, 1970). Pyrenoids are also absent in a nonsense Rubisco mutant and a temperature-sensitive Rubisco mutant grown at a non-permissive temperature, but a mutant with an inactive enzyme still appears to form a normal pyrenoid (Rawat et al., 1996). However, the loss of photoautotrophic growth of these strains meant that the effect of pyrenoid presence/absence on the CCM could not be determined.

Subsequently, pyrenoid-less mutants with hybrid or chimeric Rubisco enzymes were generated and these provided a tool for investigating the importance of the pyrenoid for growth at low CO_2. Deletion mutants in the Rubisco small subunit (Δ*RBCS1and2*) were rescued with cDNA encoding the *RBCS* genes of higher plants (spinach, Arabidopsis and sunflower RBCS hybrid mutants). These three mutant strains are able to grow photoautotrophically but have a high CO_2-requiring phenotype. In addition, despite having wild-type amounts of catalytically competent Rubisco, the strains also lack pyrenoids. So it was hypothesised that loss of the pyrenoid might affect at least some elements of the carbon concentrating mechanism (Genkov et al., 2010).

Building on the observation that sequence elements of the Chlamydomonas RBCS are required for pyrenoid formation, Dr Moritz Meyer (Griffiths lab, Department of Plant Sciences, University of Cambridge) and Dr Todor Genkov (University of Nebraska, Lincoln) developed mutants in which two α-helices of the RBCS were substituted (Fig. 4.1). The helixAB mutant has a predominantly spinach RBCS with Chlamydomonas α-helices, which is sufficient to restore the pyrenoid. Conversely, the reciprocal mutant, which has a Chlamydomonas RBCS with spinach α-helices, does not form a pyrenoid.

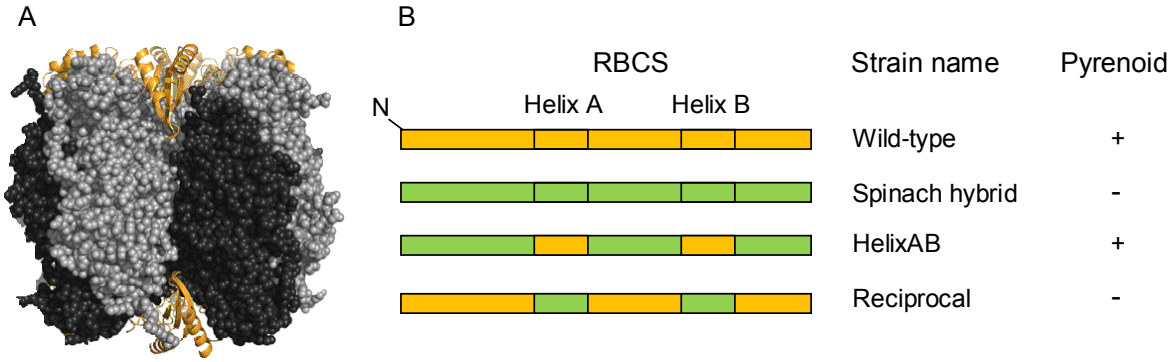


Figure 4.1 Schematic for understanding the RBCS mutants used in this study. (A) The L_8S_8 Rubisco holoenzyme showing the large subunits (rbcL) in grey and the small subunits (RBCS) in yellow. (B) Strains differ only in their respective RBCSs. Chlamydomonas sequences are shown in yellow and spinach sequences in green. Nomenclature (strain name) of these mutants follows Meyer et al. (2012). Arabidopsis and sunflower higher plant RBCS hybrid strains resemble the spinach RBCS hybrid but with the substitution of the respective RBCS genes (Genkov et al., 2010).

This set of near isogenic pyrenoid-lost and pyrenoid-restored mutants, with their constitutive pyrenoid 'off' or 'on' phenotype, established a unique system to investigate both the relationship between Rubisco localisation and CCM efficacy and the molecular role of RBCS in pyrenoid biogenesis. This chapter reports the characterisation of these mutants using the same physiological and molecular tools developed and optimised for the characterisation of a wild-type strain (Chapter 3). In particular, the aims were to measure:

(i) the extent of CCM activity, using an oxygen electrode to assay whole cell photosynthetic affinity for Ci in each mutant line;

(ii) relative expression of the known CCM genes investigated in Chapter 3 as well as other low CO_2-induced and photosynthesis-related genes in each mutant line;

(iii) differences in accumulation of known CCM proteins (CAH1, LCIB and LCIC) in each mutant line;

(iv) differences in total soluble protein between the wild-type and one pyrenoid-less mutant with a hybrid Rubisco enzyme (spinach).

These experiments were designed to test the hypothesis that loss of the pyrenoid itself is sufficient to abolish internal Ci accumulation/CCM activity and that the expression of inducible CCM components is unaffected in higher plant RBCS hybrid strains.

4.2 RESULTS

4.2.1 Pyrenoid-negative mutants have high CO_2-requiring phenotypes and reduced photosynthetic affinity for Ci

Qualitative growth analyses suggested that pyrenoid-negative RBCS mutants (spinach RBCS hybrid and reciprocal) grew poorly at low CO_2 and that restoration of the pyrenoid and CCM partially restored growth in the helixAB mutant (Meyer et al., 2012). To quantify this difference, growth rates (doubling times) were calculated for these strains at low and high CO_2. All strains grew well in medium aerated with 5% CO_2 (Fig. 4.2). There was no statistically significant difference between wild-type (mean doubling time of 11.5 h) and spinach RBCS hybrid (10.7 h) strains or between helixAB (13.9 h) and reciprocal (14.7 h) strains.

In contrast, relative growth of the four strains at low CO_2 followed a different pattern. Wild-type grew slower than at high CO_2 (doubling time of 15.3 h) but still grew the fastest of the four strains. HelixAB grew even more slowly at low CO_2 compared to both high CO_2 and wild-type with a doubling time of 28.3 h. Notably, spinach RBCS hybrid and reciprocal appeared unable to grow at low CO_2 as no increase in cell density was measured over a 48 h period.

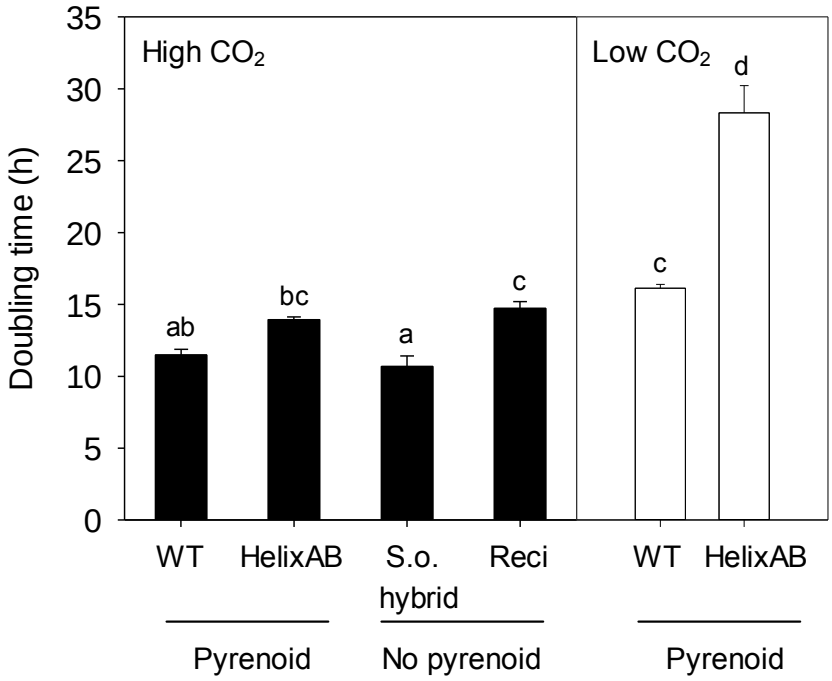


Figure 4.2 Relative growth of RBCS mutant strains grown in continuous light at high CO_2 (black bars) or low CO_2 (air; white bars). No growth of spinach (S.o.) RBCS hybrid or reciprocal strains was observed over 48 h at low CO_2 so doubling times were not calculated. Bars denote the mean ± 1 s.e. of 4-6 separate cultures during 2-3 consecutive growth experiments (dilutions). Means with different letters are significantly different (one way ANOVA and Holm-Sidak test, $p<0.05$).

The photosynthetic response of cells to external bicarbonate addition was measured (as described in Chapter 3) in wild-type and mutant strains to determine whether the lack of growth of the spinach RBCS hybrid and reciprocal strains at low CO_2, relative to wild-type and helixAB, was due to reduced CCM activity. When grown at high CO_2, all four strains showed similar responses to Ci with saturated rates of photosynthesis occurring only at very high concentrations of external Ci (≥ 1000 uM Ci; Fig. 4.3). Under these conditions, wild-type and helixAB had relatively low photosynthetic affinity for Ci ($K_{0.5}$ values of 120 and 82 µM Ci, respectively; Table 4.1) while spinach RBCS hybrid and reciprocal had even lower affinities (350 ± 43 and 500 ± 47 µM Ci, respectively; Table 4.1). Cells were acclimated to low CO_2 for at least 6 h to achieve maximum CCM expression as defined in Chapter 3. Wild-type and helixAB showed at least a ten-fold reduction in $K_{0.5}$ values after the shift from high to low CO_2 (9.4 ± 4.4 and 4.7 ± 0.9 µM Ci, respectively). In contrast, spinach RBCS hybrid and reciprocal showed no change in affinity for Ci ($K_{0.5}$) after the same period of acclimation at low CO_2 (310 ± 37 and 530 ± 130 µM Ci, respectively; Table 4.1).

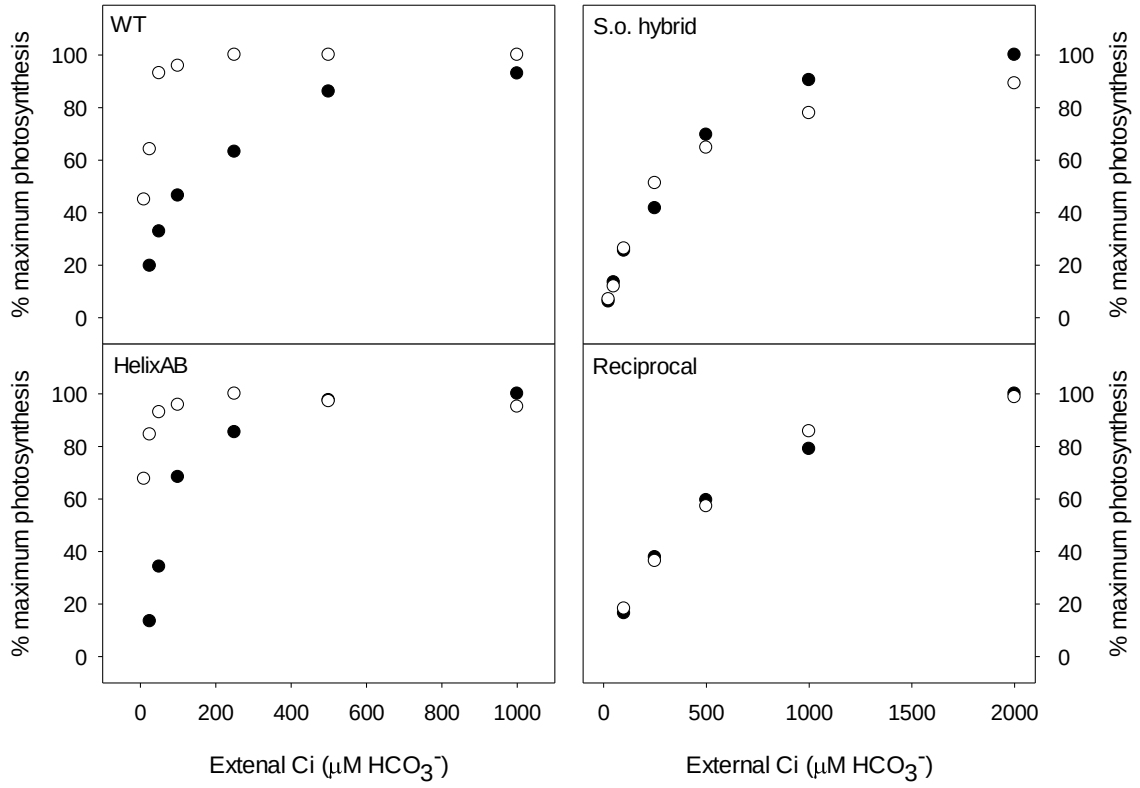


Figure 4.3 Oxygen evolution responses to external Ci additions for RBCS mutants adapted to either high CO_2 (closed circles) or low CO_2 (open circles). Note that different x-axis scales were used to represent CCM-positive (left hand panels; WT and helixAB) and CCM-negative (right hand panels; S.o. hybrid and reciprocal) strains. Each point represents the mean of three separate oxygen evolution response curves except WT high CO_2 (n = 2) and helixAB high CO_2 (n = 1). Curves for low CO_2-adapted strains were previously published in Meyer et al. (2012). See also Table 4.1.

Table 4.1 Photosynthetic affinity for Ci ($K_{0.5}$ for Ci) of RBCS mutants as determined by the oxygen evolution method. $K_{0.5}$ values were calculated by fitting the Michaelis-Menten equation to plots of external Ci versus photosynthetic rate and solving for K_m and V_{max}. Values shown are the mean ± 1 s.e. of three separate curves except WT high CO_2 (n = 2) and helixAB high CO_2 (n = 1).

	Relative affinity for Ci, $K_{0.5}$ (μM Ci)			
Strain	WT	HelixAB	S.o. hybrid	Reciprocal
High CO_2	120	82	350 ± 43	500 ± 47
Low CO_2	9.4 ± 4.4	4.7 ± 0.9	310 ± 37	530 ± 130

4.2.2 RBCS substitutions alter the expression of selected CCM genes and proteins

The loss of CCM activity observed in pyrenoid-negative mutants may be due solely to the redistribution of Rubisco throughout the stroma or it may be that expression of other CCM components was also altered in these mutants in response to pyrenoid loss. To distinguish between these two possibilities, qRT-PCR was used to probe the CCM gene expression of higher plant RBCS hybrid

strains in response to low CO_2, relative to the wild-type (Fig. 4.4; Table 4.2). To factor out positional effects associated with the random insertion of a foreign RBCS cDNA, mRNA expression analysis was compared in all three higher plant RBCS hybrids available at the time of the study (spinach, Arabidopsis, and sunflower). Data are presented in Figure 4.4 for two independent experiments, showing gene expression for the three higher plant RBCS hybrid strains, relative to the wild-type. The first dataset (A, Figure 4.4) also compared cells which were grown under high and low CO_2, whereas the second data set (B, Figure 4.4 and Table 4.2) was generated under low CO_2 only.

Of the CCM genes, transcripts for the periplasmic carbonic anhydrase *CAH1* and the Ci transporters (*CCP1*, *CCP2* and *LCI1*) were detected in much lower abundance in low CO_2-adapted higher plant RBCS hybrid strains (see Fig. 1.3 for an explanation of CCM gene function and probable subcellular localisation). In the two independent experiments (A and B), *CAH1* was present at 60% and 8% of wild-type levels while *LCI1* was detected at 200% and 9% of wild-type levels. *CCP1* and *CCP2* were only probed at low CO_2 and were detected at 0.2% and 4% of wild-type levels, respectively. *LCIA* expression was reduced in mutant strains at low CO_2 (29% and 19% of wild-type in independent experiments). Gene transcripts encoding the transcription factor, LCR1, were also less abundant in low CO_2-adapted mutants (82% and 17% of wild-type levels in independent experiments) and, similarly, *LCIC* mRNA was present at around 63% of wild-type levels.

However, not all CCM-related genes showed reduced expression in low CO_2-adapted higher plant RBCS hybrid strains compared to the wild-type. *CIA5* and *LCI5* transcripts were somewhat more abundant in the higher plant RBCS hybrid strains relative to the wild-type (3.1-fold and 7.4-fold more abundant, respectively). *CAH3*, *CAH6*, *LCIB* and *rbcL* are genes that are generally less responsive to low CO_2 and were not differentially regulated in the higher plant RBCS hybrid strains.

To test the specificity of differential gene expression to low CO_2-adapted cells, a number of CCM genes were also probed in cells grown at high CO_2. In high CO_2-adapted wild-type and mutant strains, all the CCM gene transcripts probed were found to be present in similar abundance, except perhaps *LCIA,* which was present in higher plant RBCS hybrid mutants at approximately 50% of wild-type levels, and *CAH6*, which was present at approximately 70% of wild-type levels. As

an additional control to confirm the specificity of differential expression of CCM genes, three photosynthetic genes were also probed. Expression of *GLYK* (glycerate kinase), *LHCB4* (chorophyll a-b binding protein of PSII) and *PSBO* (oxygen-evolving enhancer protein of PSII) in pyrenoid-less mutant strains was found to be similar to the wild-type at both high and low CO_2.

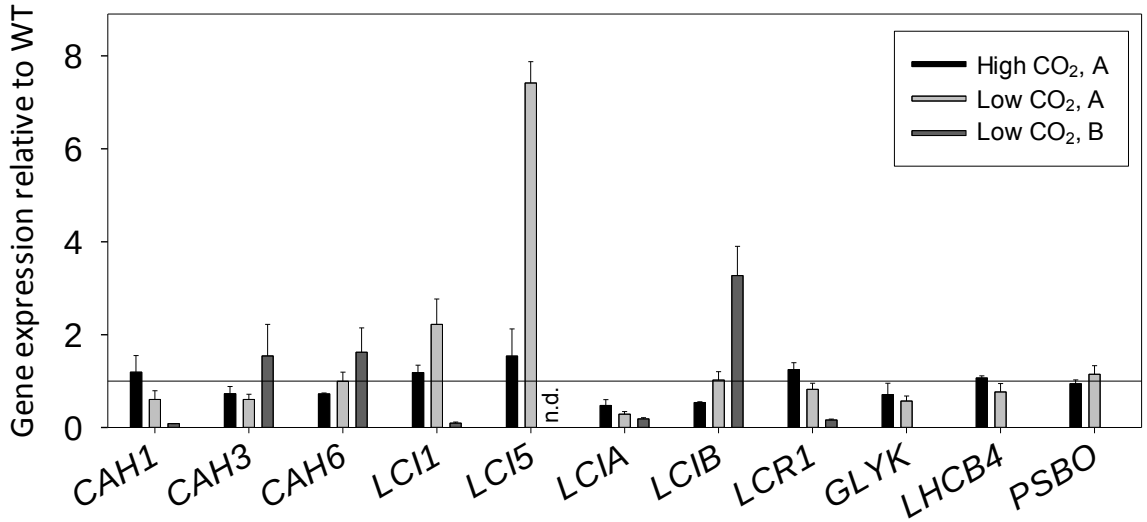


Figure 4.4 Expression of CCM genes in higher plant RBCS hybrid strains relative to wild-type. Bars denote mean and standard error of spinach, Arabidopsis and sunflower RBCS hybrid strains. CCM genes (*CAH1*, *CAH3*, *CAH6*, *LCI1*, *LCI5*, *LCIA* and *LCIB*) were probed as well as photosynthetic genes (*GLYK*, *LHC4*, *PSBO*) not associated with the CCM. Data are presented for two independent experiments (A) and (B). Raw qRT-PCR data are courtesy of Dr Moritz Meyer. See also Table 4.2.

Table 4.2 Expression of additional CCM genes in higher plant RBCS hybrid strains relative to WT at low CO_2. Values were obtained from low CO_2 (B) cDNA but were not plotted in Fig. 4.4 because there were no equivalent values for high CO_2-adapted cells.

Gene	Gene expression relative to WT (mean ± 1 s.e.)
CCP1	0.0023 ± 0.0004
CCP2	0.038 ± 0.005
CIA5	3.1 ± 0.33
LCIC	0.63 ± 0.19
rbcL	0.68 ± 0.45

Where possible, the expression of CCM and control genes was also measured at the protein level using immunoblots (Fig. 4.5). This showed that differential expression of CCM mRNA may be generally correlated with differences in protein abundance between higher plant RBCS hybrid and wild-type strains. CCM proteins (CAH1, LCIB, LCIC) were undetectable at high CO_2 but were induced to

some extent in all strains by exposure to low CO_2. As at the mRNA level, LCIB protein accumulated to a similar extent in wild-type and higher plant RBCS hybrid strains (mean fold-difference in abundance was 1.10 ± 0.019). LCIC was induced to a lesser extent in the higher plant RBCS hybrid strains (fold-difference compared to wild-type was 0.74 ± 0.03), as was also seen at the mRNA level. *CAH1* mRNA was detected at very low levels in the higher plant RBCS hybrids and, similarly, CAH1 only accumulated to very low levels in the mutant strains (fold-difference compared to wild-type of 0.13 ± 0.13). Rubisco (rbcL and RBCS) and control proteins (His H3 and PSBO) were not differentially expressed between high and low CO_2.

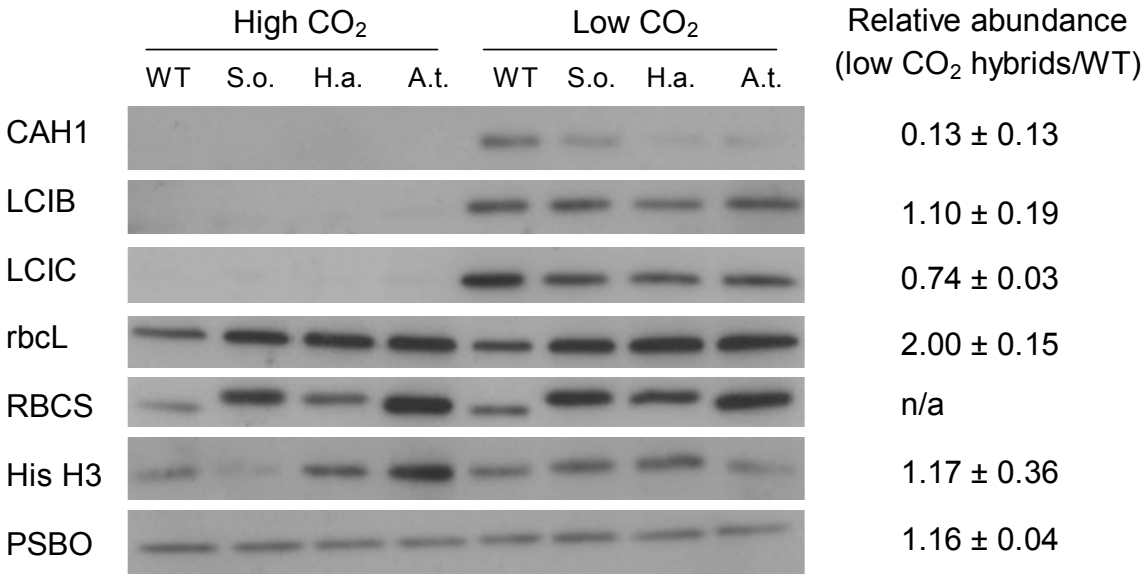


Figure 4.5 Expression levels of CCM-related proteins in WT and RBCS hybrid (S.o.: spinach, H.a., sunflower, A.t.: Arabidopsis) strains adapted to high CO_2 or after 3 h adaptation to low CO_2. Soluble proteins were separated using SDS-PAGE and probed with antibodies against Rubisco (RBCS and rbcL), CCM proteins (LCIB and LCIC), a photosynthetic protein (PSBO) and control protein Histone H3 (His H3). The abundance of the low CO_2-inducible periplasmic carbonic anhydrase, CAH1, was determined by probing concentrated growth medium (see Methods for more detail). Fold-difference in mean protein abundance/band size (± 1 s.e.) in higher plant RBCS hybrid strains relative to wild-type at low CO_2 is shown to the right of the blots.

These initial studies of CCM gene expression in the pyrenoid-less higher plant RBCS hybrid mutants identified some differential expression of CCM genes, particularly the Ci transporters (*CCP1*, *CCP2*, *LCI1*, *LCIA*) and a highly CO_2-responsive carbonic anhydrase gene (*CAH1*) at the mRNA level. However, these results were not supported by subsequent analysis of the transcriptome of these strains (Fig. 5.8, p108). Immunoblot analysis also provided preliminary evidence of

differential expression of two CCM proteins (CAH1 and LCIC) but this was limited by the availability of suitable antibodies.

4.2.3 Identification of soluble proteins differentially expressed in low CO_2-adapted pyrenoid-negative and pyrenoid-positive strains using 2D-DIGE and MS analysis

Given the reduced accumulation of CCM-related mRNA, and the potential for differentially expressed CCM proteins in pyrenoid-negative compared to pyrenoid-positive strains, a broader comparative approach was taken to identify differentially expressed proteins in wild-type and higher plant RBCS hybrid mutant strains. In a pilot study with the Cambridge Centre for Proteomics, focusing on a single higher plant RBCS hybrid mutant, soluble proteins were extracted from three separate cultures of low CO_2-adapted wild-type and spinach RBCS hybrid strains. Strains were adapted to low CO_2 for three hours, which enabled comparison with previous data (qRT-PCR and immunoblots) and was deemed sufficient for CCM-related proteins to accumulate to near-maximal abundance (Fig. 3.11). Proteins analysed by two-dimensional PAGE were separated in the first dimension by isoelectric point (pI) and by molecular weight in the second dimension. This resolved 2,500 individual fluorescently-stained protein spots across the six gels (three wild-type and three spinach RBCS hybrid; see Fig. 4.6 for representative gel). Only 12 of these spots were identified as differentially expressed based on statistical analysis (t-test) of spot size, as determined by automated image analysis. All 12 spots were picked for identification using mass spectrometry (numbered in Fig. 4.6).

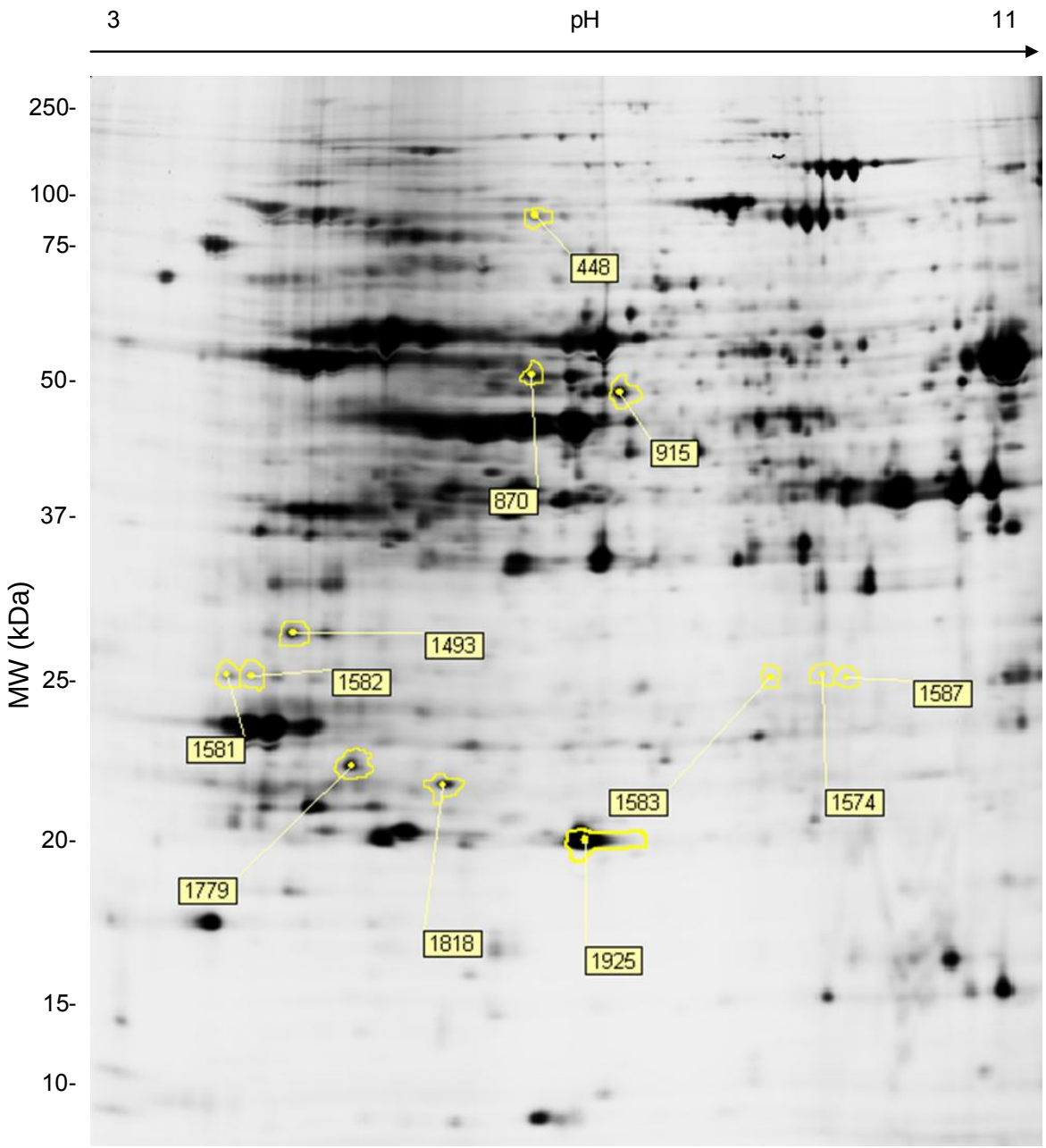


Figure 4.6 Example of low CO_2-adapted spinach RBCS hybrid soluble proteins separated by two-dimensional gel electrophoresis. Protein spots identified as differentially expressed between wild-type and spinach RBCS hybrid strains were picked for mass spectrometric analysis and are numbered on the gel. Proteins were separated in the first dimension on an immobilised non-linear pH gradient (pH 3-11) and in the second dimension by 12% SDS-PAGE. Approximate molecular weights (kDa) are shown to the left of the gel. Proteins were extracted from three separate cultures for each strain (six gels in total) and differential expression was determined relative to a control sample (mixture of all samples). Image courtesy of Cambridge Centre for Proteomics.

The most likely candidate(s) for each spot were decided based on an analysis of: the number of unique peptides mapping back to the protein (≥ 3 where possible); the percent of the total amino acid sequence covered by the detected peptides (≥ 15% where possible); the estimated molecular mass from the gel; the calculated molecular mass; the estimated isoelectric point and calculated isoelectric point.

The most likely candidates were subsequently reanalysed in light of the results of a second proteomics experiment which compared the total protein of the same strains under identical growth conditions (described in Chapter 5). The probable identities of each spot based on this analysis are summarised in Table 4.3. Predictions were classified as high confidence if the above thresholds for the number of unique peptides and percent protein coverage were met and estimated and calculated molecular weights and isoelectric points were similar.

For example, peptides matching 50 unique proteins were identified in spot 915, with the number of unique peptides assigned to each protein ranging from one to 26 and the percent coverage ranging from 0.2% to 76.8%. Of these, three proteins were identified by at least three unique peptides covering at least 15% of the amino acid sequence and fell within the range of 43 and 53 kDa (estimated molecular mass was 48 kDa). While these would all be considered high confidence hits, the protein with the highest coverage and number of unique peptides (S-adenosylmethionine synthetase) was chosen as the most likely candidate. This was also supported by the identification of several long peptides assigned to this protein.

After an initial round of mass spectrometric analysis, a single protein was identified with high confidence for ten of the twelve spots although no known CCM-associated protein was identified as a likely candidate in this analysis. Two spots (1574 and 1583) yielded interesting candidates, albeit detected with low likelihood. Spot 1925 was identified with some confidence as oxygen evolving enhancer protein 2 of PSII but was thought to perhaps be the misidentified spinach RBCS, based on estimated pI and molecular weight. Spot 870 was more abundant in the spinach RBCS hybrid and was also identified with high confidence as S-adenosylmethionine synthetase; however, this same protein was identified with an even greater coverage and number of unique peptides in spot 915, which was more abundant in the wild-type. These four spots were therefore reanalysed to resolve these uncertainties in identification but no additional candidates were identified for spots 1925, 1574 and 1583. Reanalysis confirmed spot 870 as being S-adenosylmethionine synthetase. Further analysis of mass spectrometric data for spots 915 and 870 identified the presence of different sets of unique peptides and

confirmed the likely presence of differentially modified forms of S-adenosylmethionine synthetase (Fig. 4.7).

Mass spectrometric data was also reanalysed to look for phosphopeptides in spots 1925, 1574, 1583 and 1587. Spots 1574, 1583 and 1587 have similar molecular weights but different isoelectric points, which could be due to different phosphorylation states of a single protein. One of the top candidates for all these spots is oxygen-evolving enhancer protein 2 of PSII (PSBP), which is known to be phosphorylated in Arabidopsis and *Sequoia* (Yang et al., 2003; Chang et al., 2010). Eleven unique phosphopeptides comprising eight different phosphorylated serine or threonine residues were identified in the largest spot (1925) containing PSBP (Fig. 4.8); however, no phosphopeptides were detected in the top protein hits of spots 1574, 1583 or 1587.

Table 4.3 Mass spectrometric identification of differentially expressed proteins in low CO_2-adapted wild-type and spinach RBCS hybrid strains. Spots are numbered as shown in Fig. 4.6. The relative abundance of proteins in wild-type compared to spinach RBCS hybrid based on image analysis of replicate gels is given (WT/Sp spot) and t-tests were used to determine whether spot sizes were significantly different between samples (p-value). NCBI accession numbers, Uniprot IDs and descriptions are provided for the most likely candidate(s) for each spot based on an analysis of: the number of unique peptides (Unique pep.); the % coverage of the protein sequence (% cover); the estimated molecular weight from the gel (Est. MW); the calculated molecular weight (MW); the estimated isoelectric point (Est. pI) and calculated isoelectric point (pI). The number of unique phosphopeptides (Unique phosST pep.) and the number of unique phosphorylation sites (PhosST sites) are shown. LC-MS/MS refers to the second proteomics experiment, described in Chapter 5. Proteins are listed as detected (Det.) and present in significantly different abundance (Sig. diff.) according to a t-test or G-test, while relative protein abundance is shown as the ratio of the total spectral counts between the two strains (WT/Sp S.C.). Where more than one candidate is listed, the most likely candidate is listed first (in bold). No likely (high confidence) hits were detected for spots 1583 or 1587.

2D-DIGE analysis													**LC-MS/MS analysis**		
Spot	**WT/Sp (spot)**	**p-value**	**NCBI acc. no. (Uniprot ID)**	**Description**	**Unique pep.**	**% cover**	**Est. MW**	**MW**	**Est. pI**	**pI**	**Unique phosST pep.**	**PhosST sites**	**Det.**	**Sig. diff.**	**WT/Sp (S.C.)**
915	**1.49**	**0.014**	**gi|159477124 (A8HYU5)**	**S-adenosylmethionine synthetase**	**28**	**73.8**	**48**	**43**	**6.3**	**6.0**	**n/a**	**n/a**	**Y**	**Y (T)**	**1.16**
			gi|159477301 (A8J506)	argininosuccinate synthase	19	46.5	48	49	6.3	8.4	n/a	n/a	Y	N	-1.05
			gi|159468534 (A8ITU2)	predicted protein	14	40.2	48	43	6.3	6.3	n/a	n/a	Y	N	-1.03
448	**1.36**	**0.017**	**gi|159484368 (A8JCQ8)**	**acetyl CoA synthetase**	**22**	**45.8**	**80**	**78**	**5.8**	**6.4**	**n/a**	**n/a**	**Y**	**N**	**1.20**
			gi|159488061 (A8JFR9)	acetyl CoA synthetase	15	26.1	80	74	5.8	7.0	n/a	n/a	Y	Y (G)	-1.58
1493	**-1.22**	**0.017**	**gi|159478202 (A8J6D1)**	**chlorophyll a-b binding protein of PSII**	**6**	**30.4**	**28**	**30**	**5.0**	**6.2**	**n/a**	**n/a**	**Y**	**Y (G)**	**-1.22**
			gi|159489184 (A8ICC8, Q93Y52)	inorganic pyrophosphatase	12	43.6	28	31	5.0	6.3	n/a	n/a	Y	N	1.25
			gi|159488705 (A8JGJ0)	hypothetical protein	6	27.9	28	31	5.0	5.6	n/a	n/a	N	n/a	n/a
			gi|159475641 (Q9FEK6)	minor chlorophyll a-b binding protein of PSII	6	19.0	28	31	5.0	5.4	n/a	n/a	Y	N	1.05

Table 4.3 continued

2D-DIGE analysis													LC-MS/MS analysis		
Spot	**WT/Sp (spot)**	**p-value**	**NCBI acc. no. (Uniprot ID)**	**Description**	**Unique pep.**	**% cover**	**Est. MW**	**MW**	**Est. pI**	**pI**	**Unique phosST pep.**	**PhosST sites**	**Det.**	**Sig. diff.**	**WT/Sp (S.C.)**
1779	**-1.30**	**0.040**	**gi|159463028 (A8HN52)**	**glutaredoxin, CGFS type**	**11**	**44.8**	**23**	**31**	**5.2**	**8.9**	**n/a**	**n/a**	**N**	**n/a**	**n/a**
			gi|159466246 (A8IPC9)	riboflavin synthase	6	34.0	23	23	5.2	5.4	n/a	n/a	Y	N	(1.25)
			gi|159491518 (A8JIE0)	chaperonin 20	5	37.5	23	23	5.2	8.4	n/a	n/a	Y	N	1.23
1818	**-1.32**	**0.054**	**gi|159482564 (A8JBF2)**	**organellar elongation factor P**	**9**	**46.2**	**22**	**21**	**5.5**	**6.0**	**n/a**	**n/a**	**Y**	**N**	**(1.15)**
1581	-1.33	0.032	gi|159485778 (A8I5F7)	translocon-associated protein beta	3	12.8	25	25	4.5	5.5	n/a	n/a	N	n/a	n/a
			gi|159463778 (A8HMA8)	acetyl-coa biotin carboxyl carrier	3	11.6	25	25	4.5	6.5	n/a	n/a	Y	N	(2.25)
1925	**-1.36**	**0.006**	**gi|159471964 (A8IYH9, P11471)**	**oxygen-evolving enhancer protein 2 of PSII**	**21**	**58.5**	**20**	**30**	**6.0**	**9.3**	**11**	**8**	**Y**	**Y (G)**	**1.18**
			gi|159480962 (Q19VH4)	flagellar flavodoxin	5	26.4	20	21	6.0	6.5	0	0	Y	N	(1.08)
			gi|159481799 (A8JAI8)	20S proteasome beta subunit D, type 2	5	25.9	20	21	6.0	6.2	0	0	N	n/a	n/a
870	**-1.41**	**0.059**	**gi|159477124 (A8HYU5)**	**S-adenosylmethionine synthetase**	**24**	**57.9**	**50**	**43**	**5.8**	**6.0**	n/a	n/a	**Y**	**Y (T)**	**1.16**

Table 4.3 continued

2D-DIGE analysis													LC-MS/MS analysis		
Spot	**WT/Sp (spot)**	**p-value**	**NCBI acc. no. (Uniprot ID)**	**Description**	**Unique pep.**	**% cover**	**Est. MW**	**MW**	**Est. pI**	**pI**	**Unique phosST pep.**	**PhosST sites**	**Det.**	**Sig. diff.**	**WT/Sp (S.C.)**
1582	**-1.47**	**0.004**	**gi|159473144 (A8J0E4, P12853)**	**oxygen-evolving enhancer protein 1 of PSII**	**12**	**48.5**	**25**	**31**	**4.6**	**8.3**	**n/a**	**n/a**	**Y**	**N**	**1.07**
			gi|159484978 (A8JDN2)	adenylate kinase 3	10	40.2	25	29	4.6	5.8	n/a	n/a	Y	N	(-1.30)
			gi|159486567 (A8JES1)	mitochondrial grpE-type co-chaperone of the HSP70 system	7	24.2	25	29	4.6	5.8	n/a	n/a	N	n/a	n/a
			gi|159463778 (A8HMA8)	acetyl-coa biotin carboxyl carrier	4	19.4	25	25	4.6	6.5	n/a	n/a	Y	N	(2.25)
			gi|159485778 (A8I5F7)	translocon-associated protein beta	5	17.7	25	25	4.6	5.5	n/a	n/a	N	n/a	n/a
			gi|159491492 (Q93WL4)	light-harvesting complex II chlorophyll a-b binding protein M3	3	19.8	25	27	4.6	5.7	n/a	n/a	Y	N	1.05
1583	-1.62	0.018	gi|159483497 (A8I4P5)	ribosomal protein S3, component of cytosolic 80S ribosome and 40S small subunit	3	12.6	25	26	7.7	9.7	0	0	Y	Y (G, T)	1.62
			gi|159471964 (A8IYH9, P11471)	oxygen-evolving enhancer protein 2 of PSII	2	9.6	25	30	7.7	9.3	0	0	Y	Y (G)	1.18
			gi|159464581 (A8II85)	hypothetical protein	2	12.0	25	18	7.7	9.4	0	0	N	n/a	n/a
1574	**-1.69**	**0.014**	**gi|159471964 (A8IYH9, P11471)**	**oxygen-evolving enhancer protein 2 of PSII**	**4**	**18.8**	**25**	**30**	**8.4**	**9.3**	**0**	**0**	**Y**	**Y (G)**	**1.18**
1587	-2.05	0.002	gi|159471964 (A8IYH9, P11471)	oxygen-evolving enhancer protein 2 of PSII	3	13.1	25	30	8.5	9.3	0	0	Y	Y (G)	1.18

870 MAALESFLYTSESVNEGHPDK**ICDQVSDAVLDACLEQDPDSKVACETCCKTGMVMVFGEITTKAKVDYEAI**
915 MAALESFLYTSESVNEGHPDK**ICDQVSDAVLDACLEQDPDSKVACETCCKTGMVMVFGEITTKAKVDYEAI**

870 **VR**KVCK**EIGFTSEDVGLDADKCK**VLVHIEEQSPDIGQGVHGMGTKALEEIGAGDQGHMFGYATDETPELMP
915 **VR**KVCK**EIGFTSEDVGLDADKCKVLVHIEEQSPDIGQGVHGMGTKALEEIGAGDQGHMFGYATDETPELMP

870 LTHVLATQLGYKLTEVRK**NGVCPWLRPDGKTQVTVEYK**R**EGGAMIPQRVHTILISTQHNPDVTNEKIREDL**
915 **LTHVLATQLGYK**LTEVR**KNGVCPWLRPDGKTQVTVEYKREGGAMIPQRVHTILISTQHNPDVTNEKIREDL**

870 **MEHVIKPVVPAK**YLDDK**TIFHLNPSGRFVIGGPHGDAGLTGRKIIIDTYGGWGAHGGGAFSGKDPTKVDR**S
915 **MEHVIKPVVPAK**YLDDK**TIFHLNPSGRFVIGGPHGDAGLTGRKIIIDTYGGWGAHGGGAFSGKDPTKVDR**S

870 GAYIARQAAK**SVVASGLAK**RCLVQVSYSIAVAEPLSVFVDTYGTGTMPDAEILKLIRK**HFDFRPGLIGK**NL
915 GAYIARQAAK**SVVASGLAK**RCLVQVSYSIAVAEPLSVFVDTYGTGTMPDAEILKLIRK**HFDFRPGLIGK**NL

870 DLKRGGNKRYQK**TAAYGHFGRDDPDFTWETVK**KLE
915 DLKRGGNKRYQK**TAAYGHFGRDDPDFTWETVK**KLE

Figure 4.7 Unique peptide coverage of S-adenosylmethionine synthetase in protein spots 870 and 915. Proteins were picked from 2D-DIGE gels and identified using mass spectrometry. The coverage of the protein sequence in each spot is highlighted in red and bold font.

MATALCNKAFAAAPVARPASRRSAVVVRASGSDVSRRAALAGFAGAAALVSSSPANAAYGDSANVFGKVTNKS
GFVPYAGDGFALLLPAKWNP**S**KENDFPGVILRYEDNFDAVNNLVVIAQDTDKKAIADFGSQDKFLE**S**VSYLLG
KQAY**S**GE**T**QSEGGFAPNRV**S**AASLLDVSTTTDKKGK**T**YYKYELLVR**S**ADGDEGGRHQLIGATVG**S**DNKLYIIK
IQIGDKRWFKGAKKEAMGAFDSFTVV

Figure 4.8 Phosphorylation sites (yellow, bold) of the oxygen-evolving enhancer protein 2 of photosystem II (PSBP) detected in spot 1925. Mass spectra were reanalysed to identify likely phosphopeptides in all four spots containing PSBP but were not identified in spots 1574, 1583 or 1587. The chloroplast transit peptide sequence is underlined.

4.2.4 Loss of the pyrenoid leads to specific changes in protein abundance

Protein spots identified by mass spectrometry were further analysed based on manual curation of broad and specific functional groups, protein localisation and whether the genes had been found to be CO_2-responsive in previous genome-wide RNA sequencing experiments (Table 4.4). Of the nine unique proteins identified, three were associated with photosynthesis, three were involved in protein synthesis, two were involved in primary metabolism and one was possibly involved in stress/redox homeostasis.

Of the photosynthesis-associated proteins, two oxygen-evolving enhancer proteins associated with PSII (PSBO and PSBP) were found to be more abundant in the

spinach RBCS hybrid and PSBP was present in at least three separate spots. The largest spot (1925) contained a smaller and highly phosphorylated isoform of PSBP (Fig. 4.6 and 4.8). The reduced size could be due to cleavage of the 5.5 kDa chloroplast transit peptide while phosphorylation of the protein may substantially reduce the pI. The slower migration in the second dimension of PSBO (spot 1582) and of PSBP in spots 1583, 1574 and 1587 compared to spot 1925 (PSBP) might be due to these proteins retaining their transit peptides, although there was no evidence for this in the mass spectrometry data. The third PSII-associated protein was a chlorophyll a-b binding protein (LHCB4), which was also present in greater abundance in spinach RBCS hybrid. Comparison of the estimated and calculated molecular weights suggests that this protein may also retain its transit peptide in this spot.

The primary metabolic enzymes acetyl CoA synthetase and S-adenosylmethionine synthetase were both found to be more abundant in wild-type cells (spots 448 and 915, respectively). However, in a different spot (870), S-adenosylmethionine synthetase was more abundant in spinach RBCS hybrid. The small differences in estimated molecular weight and pI (Table 4.3) as well as the different peptide patterns (Fig. 4.7) of these two spots suggests that the protein may have different post-translational modifications and hence, different activity levels or functions in the two strains. The remaining protein, glutaredoxin (GRX6; spot 1779), was also present in greater abundance in the spinach RBCS hybrid strain. Overall, six of the nine proteins identified as differentially expressed were localised to the chloroplast.

Table 4.4 Summary of soluble proteins identified as differentially expressed in low CO_2-adapted wild-type and spinach RBCS hybrid strains. Gene ontology and localisation (Local.) were manually determined based on UniProt annotation. The number of amino acids (aa) in the chloroplast transit peptides (cTP) was predicted using ChloroP 1.1 (Emanuelsson et al., 1999) except for PSBP and PSBO, which were annotated on UniProt. Whether the genes encoding these proteins were differentially regulated in wild-type cells in response to low CO_2 according to Brueggman et al., 2012 (Brueg.) and Fang et al., 2012 (D indicates downregulation), and whether the proteins were identified in the chloroplast proteome by Terashima et al., 2011 (Tera.) was also recorded.

Spot	WT/Sp (spot)	NCBI acc. no. (Uniprot ID)	Description	Gene ontology	Role/pathway	Local.	cTP (aa)	Brueg.	Fang	Tera.
915 870	W S	gi|159477124 (A8HYU5)	S-adenosylmethionine synthetase (METM)	primary metabolism	synthesis of AdoMet (methyl donor)	Cy	n/a	Y (D)	N	N
448	W	gi|159484368 (A8JCQ8)	acetyl CoA synthetase (ACS2)	primary metabolism	carbon metabolism	Cp	51	Y (D)	N	Y
1493	S	gi|159478202 (A8J6D1)	chlorophyll a-b binding protein of PSII (LHCB4)	photosynthesis	PSII	Cp	38	Y (D)	N	Y
1779	S	gi|159463028 (A8HN52)	glutaredoxin, CGFS type (GRX6)	stress	redox homeostasis	Cp	52	Y (D)	Y (D)	N
1818	S	gi|159482564 (A8JBF2)	organellar elongation factor P (EFP1)	protein synthesis	translation	Cp	N	N	N	N
1581	S	gi|159485778 (A8I5F7)	translocon-associated protein beta	protein synthesis	protein targeting/processing	Cy	n/a	Y? (D)	N	N
1925 (1583) 1574 1587	S	gi|159471964 (A8IYH9, P11471)	oxygen-evolving enhancer protein 2 of PSII (PSBP)	photosynthesis	PSII	Cp	57	Y (D)	N	Y
1582	S	gi|159473144 (A8J0E4, P12853)	oxygen-evolving enhancer protein 1 of PSII (PSBO)	photosynthesis	PSII	Cp	52	Y (D)	N	Y
(1583)	S	gi|159483497 (A8I4P5)	ribosomal protein S3, cytosolic 80S ribosome and 40S small subunit (RPS3)	protein synthesis	translation	Cy	n/a	Y (D)	N	N

4.3 DISCUSSION

4.3.1 Pyrenoid loss in RBCS mutants is associated with loss of whole cell CCM activity and reduced CCM gene expression

Mutant Chlamydomonas strains with higher plant RBCSs (or higher plant-like RBCS α-helices) are unable to form pyrenoids even under CCM-induced conditions. Loss of the pyrenoid leads to reduced photosynthetic affinity for Ci and impaired growth at low CO_2 as well as altered light use characteristics and thylakoid arrangement (to be discussed further in Chapter 5). In contrast, under high CO_2 or in heterotrophic conditions, these strains exhibit normal growth. The somewhat (but not significantly) slower growth rate of the helixAB mutant compared to the wild-type is likely due to the relatively poor kinetics of the chimeric Rubisco, which in turn may lead to a reduction in maximum photosynthetic rate relative to the wild-type (Meyer et al., 2012). The reciprocal mutant also grew more slowly at high CO_2 than its counterpart, the spinach RBCS hybrid strain. However, no difference was observed in the maximum photosynthetic rates of reciprocal and spinach RBCS hybrid strains suggesting that the kinetics of the reciprocal enzyme are not limiting at saturating concentrations of Ci (Meyer et al., 2012). Spinach RBCS hybrid and reciprocal strains also had higher calculated $K_{0.5}$ values compared to wild-type and helixAB strains at high as well as low CO_2, which suggests that the presence of a minimal pyrenoid may boost affinity for Ci even under conditions in which the CCM is repressed. This is based on the observation that, while spinach RBCS hybrid and reciprocal mutants entirely lack a pyrenoid in both CCM-induced and CCM-repressed conditions, wild-type cells retain at least 50% of Rubisco in the pyrenoid even at high CO_2 or in the dark (Chapter 3; Mitchell et al., 2014; Borkhsenious et al., 1998).

Further characterisation of the CCM phenotype, at the molecular rather than whole cell level, in higher plant RBCS hybrid strains was performed using qRT-PCR and immunoblots to probe the expression of known CCM components. The expression of several highly CO_2-responsive genes, especially the Ci transporters *CCP1*, *CCP2*, *LCI1* and *LCIA* as well as *CAH1*, was reduced in higher plant RBCS hybrids compared to the wild-type after adaptation to low CO_2. The expression of other CCM-related genes such as *CAH3* and *CAH6* was not altered in the mutants but these genes are minimally responsive to low CO_2. Comparison of two

independent experiments analysing gene expression at low CO_2 revealed some variability but, more importantly, also supported the role of transcription factor LCR1 in CCM gene expression. In one experiment, *LCR1* mRNA levels were much lower in the mutants than in the wild-type and this was correlated with lower expression of CO_2-responsive genes including those already shown to be under the control of LCR1 (*CAH1* and *LCI1*). In contrast, when *LCR1* expression was at wild-type level, the expression of *CAH1* and *LCI1* were also close to wild-type levels. Although only a limited number of CCM proteins were able to be probed, immunoblot analyses indicated that reduced CCM mRNA abundance was reflected in reduced CCM protein abundance.

Similar reduced expression of CCM genes in mutants, unrelated to the primary lesion or mutation, has been observed in other CCM mutants, although this phenomenon remains largely unexplored. For example, the *ca-1* (CAH3) mutant shows reduced accumulation of 21 kDa and 36 kDa (CCP) polypeptides in response to low CO_2. *CCP* mRNA is present at similar levels to the protein (21% and 17% of wild-type levels, respectively), suggesting that loss of CAH3 function may lead to impaired low CO_2-induced gene transcription (Spalding et al., 1991). The *pmp1* (*LCIB*) mutant was also shown to lack *LCIA* and *HLA3* mRNA accumulation in response to low CO_2 (Miura et al., 2004); however, this was not the case in later studies (Wang and Spalding, 2006; Duanmu and Spalding, 2011). The *cia6* mutant exhibits abnormal pyrenoid morphology and CCM activity as well as reduced accumulation of *CAH4*, *CCP1* and *LCIB* mRNA (Ma et al., 2011), although the exact function of *CIA6* is not clear. Finally, screens of high CO_2-requiring mutants have identified strains that do not appear to be allelic to *cia-5*/*ccm1* yet still fail to accumulate significant levels of mRNA or protein of more than one CCM component, although the genes responsible for these phenotypes are yet to be confirmed (Fukuzawa et al., 1998; Van et al., 2001; Wang et al., 2014b).

Mutants that lack Rubisco or produce only an inactive enzyme also fail to accumulate three low CO_2-inducible proteins (Villarejo et al., 1996b). Similarly, temperature sensitive Rubisco mutants induce CO_2-responsive proteins only when grown at permissive temperatures and mutants with reduced amounts of Rubisco show lower levels of induction of CAH1 (Chen et al., 1990; Tavío et al., 1996). The

CCM phenotype of the RBCS mutants in this study resembles these other Rubisco mutants, which may be due to the requirement of photosynthesis for full CCM expression. Alternatively, the RBCS mutant phenotype may resemble that of a regulatory mutant or mutant with otherwise impaired CO_2 sensing, signalling or induction.

4.3.2 Preliminary functional analysis of changes to the soluble proteome of wild-type versus spinach RBCS hybrid

Analysis of the known CCM components is important for understanding their relationship to pyrenoid formation and CCM activity but this will not provide any further information on additional as yet unidentified elements of the CCM. This pilot proteomic study was used to test the possibility of using pyrenoid-negative mutants, deficient in CCM activity, to probe for novel CCM candidates and mechanisms of CCM regulation. Discussion of the proteomics results presented in this chapter will be brief because a full discussion of the functional implications of pyrenoid loss based on the combined results of 2D-DIGE and genome-scale LC-MS/MS experiments will follow in Chapter 5.

In this pilot study, differential gene expression in a pyrenoid-less RBCS mutant (spinach RBCS hybrid) compared to the wild-type was further investigated at the level of whole cell soluble protein content using 2D-DIGE and mass spectrometric identification of proteins spots. The nine proteins identified as more abundant in the spinach RBCS hybrid after adaptation to low CO_2 were generally associated with PSII and protein synthesis and, notably, almost all of these proteins have been shown to be downregulated at the mRNA level in wild-type cells in response to low CO_2 (Miura et al., 2004; Brueggeman et al., 2012).

One chlorophyll-binding (LHCB4) and two oxygen evolving enhancer proteins (PSBP and PSBO), all associated with PSII, were more abundant in spinach RBCS hybrid cells. LHCB4 and PSBO were not found to be differentially expressed at the mRNA level (Fig. 4.4) but post-transcriptional mechanisms may be controlling protein accumulation. Alternatively, differential post-translational modification of these proteins may be occurring, which would be consistent with the presence of multiple spots in the gel with similar molecular masses and slightly different pIs (Fig. 4.6). The presence of PSBP in at least three spots (Table 4.4)

also supports the idea of multiple states of protein modification. Alterations in PSII-associated protein abundance might be related to changes in the thylakoid membrane organisation and light harvesting characteristics of low CO_2-adapted pyrenoid-less cells. Investigation of the regulation of the light reactions in pyrenoid-less RBCS mutant strains is being undertaken by Oliver Caspari (Griffiths lab, Department of Plant Sciences, University of Cambridge) and will be discussed in Chapter 5.

Assuming the proteins associated with the photosystems are generally present in greater abundance in the spinach RBCS hybrid strain, the increased abundance of organellar elongation factor P (EFP1) as well as cytosolic translation machinery (ribosomal protein S3 and translocon-associated protein beta) could be related to higher rates of protein synthesis. This would also be consistent with most of the differentially expressed proteins being more abundant in spinach RBCS hybrid than in wild-type cells.

A glutaredoxin (GRX6) was also more abundant in spinach RBCS hybrid cells and may be involved in redox control or signalling in response to changes in photosynthesis/light requirements due to low CO_2. However, this annotation is provisional because GRX6 was not found to have classical glutaredoxin activities (disulphide reductase or dehydroascorbic acid reductase activity) in vitro and an alternative function in iron-sulphur cluster assembly has been proposed (Gao et al., 2010). The N-terminal domain of GRX6 also shares strong similarity with the sequence of slr1035 from *Synechocystis* sp. PCC6803, which is a protein of unknown function. Redox regulation and the potential for light stress in pyrenoid-negative cells will also be discussed in greater detail in Chapter 5.

Acetyl CoA synthetase and S-adenosylmethionine synthetase were both found to be more abundant in wild-type cells (spots 448 and 915, respectively). Acetyl CoA synthetase uses acetate as a substrate to produce acetyl CoA, which is involved in many metabolic pathways, including amino acid biosynthesis, carbon metabolism and fatty acid biosynthesis. Upregulation of the enzyme responsible for acetyl CoA production is perhaps associated with more active metabolism in wild-type cells, sustained by CCM induction and higher rates of carbon fixation.

S-adenosylmethionine synthetase catalyses the synthesis of S-adenosylmethionine, which is the primary methyl donor in cells (Takusagawa et al., 1996). S-adenosylmethionine is a substrate for DNA and protein methylation and is therefore implicated in gene expression as well as other regulatory and metabolic functions (Ravanel et al., 1998). Differential expression of S-adenosylmethionine synthetase may thus be associated with low CO_2-induced changes in protein function or cellular metabolism.

Although the sample size is small, this suggests that the chloroplast may be the main site of acclimation to low CO_2 and/or regulation of pyrenoid biogenesis. Seven of the eight proteins that were more abundant in the spinach RBCS hybrid mutant were also found to be transcriptionally downregulated in wild-type cells in response to low CO_2 in at least one transcriptome study. Together with expression data on the CCM genes, this suggests that the spinach RBCS hybrid mutant neither fully induces CCM components nor reduces expression of the pathways normally downregulated in response to low CO_2. This presents the intriguing possibility that Rubisco localisation and/or activity control/s CCM induction and acclimation to low CO_2. Nevertheless, this picture is incomplete, in part due to limitations inherent in the more targeted analyses (qRT-PCR, immunoblots) and the 2D-DIGE analysis used thus far.

4.3.3 Sensitivity and limitations of 2D-DIGE analysis

Given the apparent lack of transcriptional induction of the (membrane-localised) Ci transporters, one of the major limitations of using 2D-DIGE is that it is only effective in analysing soluble proteins. An alternative approach would be to extract membrane proteins by incubating samples sequentially with detergents of increasing strength but then a different means of protein separation must be used.

In addition, while 2D gels are capable of resolving large numbers of proteins, the gels do not capture proteins of lower abundance well and individual protein spots may be contaminated by highly abundant proteins such as histones, which were identified in all twelve spots analysed. While using estimated molecular mass and pI to filter mass spectrometry data can yield likely candidates, post-translational modification can alter both these parameters. In the case of PSBP, which was present in one large and two small spots, phosphorylation of the protein and

cleavage of the transit peptide may explain shifts in the migration pattern. The small spots might represent the less abundant or immature forms of the protein, which may not be the most biologically relevant. In some cases, distinguishing between differently modified forms of a single protein may be an advantage but the data may be difficult to interpret without measures of the overall abundance of the protein.

4.3.4 Conclusions and future work

Detailed analysis of the phenotype of RBCS mutants has provided strong evidence for the proposed link between CCM activity and pyrenoid presence. Mutants with higher plant RBCSs or higher plant-like RBCS α-helices completely lack pyrenoids and were also shown to lack a functioning CCM. Further investigation of CCM expression revealed specific changes in the transcription of several low CO_2-inducible genes as well as differences in the abundance of a small number of photosynthetic and metabolic proteins. Further experiments that could independently confirm 2D-DIGE results include labelling and/or quantitation of proteins of interest, for example, using immunoprecipitation or SILAC (stable isotope labelling by/with amino acids in cell culture). However, it seems more promising to first extend the analysis to the genome level. In light of this, the transcriptomes of the higher plant RBCS hybrid mutants under CCM-repressed and CCM-induced conditions are currently being analysed (by Dr Moritz Meyer and Richard Smith-Unna, Hibberd lab, University of Cambridge). One aim of this approach is to identify significantly repressed genes in the higher plant RBCS hybrids compared to the wild-type under low CO_2 which may represent other candidate transporters. In addition to the transcriptomics experiment, large-scale shotgun proteomics was also carried out with the aim of identifying changes to metabolism and gene expression that are required for acclimation to low CO_2.

Crosstalk between the light harvesting and carbon fixation pathways of photosynthesis as well as the energy-requiring pathways such as protein synthesis and primary metabolism is likely given the large-scale changes to ultrastructure, growth and metabolism that must occur to enable growth at low CO_2. This model of CCM regulation is supported by the initial proteomics data presented in this chapter and will be discussed further in the context of the follow-up proteomics experiment in Chapter 5.

5 USING GENOME-SCALE PROTEOMICS TO INVESTIGATE LOW CO_2 ADAPTATION AND PYRENOID FUNCTION IN THE PYRENOID-LESS SPINACH RBCS HYBRID MUTANT

5.1 Introduction

Advances in genome sequencing and annotation have enabled analysis of CCM induction to be extended from small-scale or targeted approaches such as Northern blots, qRT-PCR, immunoblots and radiolabelling of proteins (for example, Spalding and Jeffrey, 1989; Spalding et al., 1991; Fukuzawa et al., 2001; Xiang et al., 2001; Mitchell et al., 2014; Tirumani et al., 2014), to larger scale experiments such as microarrays (Miura et al., 2004; Yamano et al., 2008), and finally, to experiments at the whole genome level such as RNA sequencing and formaldehyde-assisted isolation of regulatory elements (for example, Brueggeman et al., 2012; Fang et al., 2012; Winck et al., 2013). These genome-scale studies have uncovered large-scale transcriptional remodelling in cells acclimating to low CO_2 and a comparison of early microarray studies with new transcriptomic data has identified a core set of 27 low CO_2-inducible genes (Supplemental Figure S7 of Brueggeman et al., 2012). Studies such as these are no longer technically challenging and have the advantage of generating large amounts of data at relatively low cost.

However, while there is no doubt that transcriptomic studies have been very informative in identifying genes that respond transcriptionally to low CO_2, the drawback is that the correspondence between mRNA and protein abundance can be variable (Laurent et al., 2010; Vogel and Marcotte, 2012). Although the relationship is generally considered to be positive, protein abundance is influenced not only by mRNA abundance but also by post-transcriptional regulatory processes such as control of translation, post-translational modification or other factors affecting mRNA and protein stability. In a recent systems analysis of the response of Chlamydomonas to increased irradiance, a reasonably positive relationship was observed overall between mRNA and protein abundance (R^2 = 0.24) but the study also identified changes in the abundance of a number of proteins, including photosynthetic proteins, that were independent of mRNA abundance (Mettler et al., 2014).

Fortunately, advances in quantitative proteomics are now enabling the more routine study of protein expression at the genome level. These approaches have been successfully used to study acclimation in microalgae, identifying coordinated changes to photosynthesis and other metabolic pathways in response to a range of stresses (for example, Pandhal et al., 2007; Höhner et al., 2013; McKew et al., 2013; Valledor et al., 2013). Although coverage of the proteome is roughly one order of magnitude below that of RNA sequencing data, with the detection limit of a single shotgun LC/MS-MS experiment currently around 1,500 proteins, the distinct advantage of this method is the ability to work directly at the protein level.

The aim of the experiment described in this chapter was to use quantitative proteomics to identify changes to the proteome induced in response to low CO_2 in the pyrenoid-positive wild-type and pyrenoid-negative spinach RBCS hybrid strains described in Chapter 4. The results presented in Chapter 4 suggest that, upon exposure to low CO_2, the spinach RBCS hybrid strain fails to fully induce several CO_2-responsive genes at the transcript level. A preliminary proteomics experiment using 2D-DIGE did not identify reduced expression of any CCM-related proteins but did identify several proteins that were present in higher abundance in the spinach RBCS hybrid after adaptation to low CO_2 and that were transcriptionally downregulated in true wild-type cells in response to low CO_2 (Brueggeman et al., 2012). Whole genome proteomic analysis of the spinach

RBCS hybrid mutant was chosen to test the hypothesis that the response of the spinach RBCS hybrid mutant to low CO_2 is generally impaired, perhaps in the sensing or signalling pathways. Protein extraction and shotgun LC-MS/MS analysis was performed by Dr Metodi Metodiev and Dr Gergana Metodieva (Proteomics group, School of Biological Sciences, University of Essex).

In summary, this genome-level study of low CO_2-adapted cells identified approximately 11% of the 1,376 proteins detected as differentially expressed in the spinach RBCS hybrid compared to the wild-type strain. Differentially expressed proteins included some CCM/low CO_2-induced proteins as well as other proteins associated with respiration and energy-requiring pathways such as protein synthesis and primary metabolism. This set of proteins was also compared to a transcriptomic study conducted in parallel with the proteomic analysis (Dr Moritz Meyer, unpublished data) as well as to previously published low CO_2-induced changes to the transcriptome of wild-type and *cia5* mutant cells (Brueggeman et al., 2012; Fang et al., 2012).

Overall, the study identified several candidate CCM proteins including those encoded by previously identified low CO_2-induced genes as well as proteins of unknown function. These initial results are promising and indicate that further investigation of the pyrenoid-negative RBCS mutants is likely to yield additional insight into CCM regulation and the role of pyrenoid formation in Chlamydomonas.

Finally, it is important to note that work on the RBCS mutants was carried out in collaboration with other members of the group. Dr Moritz Meyer performed the qRT-PCR analysis described in Chapter 4 as well as the subsequent transcriptomic analysis described in this chapter, which was designed to complement the proteomics experiments and to build on the initial phenotyping of higher plant RBCS hybrid and helixAB mutants (Genkov et al., 2010; Meyer et al., 2012). Oliver Caspari has been focussing on potential Rubisco-Rubisco interactions as well as the effects of pyrenoid loss on the light harvesting and electron transport capabilities of these strains. Where relevant, the results of these complementary studies will be incorporated into the analysis presented in this chapter.

5.2 Results

5.2.1 Overview of proteomic analysis

Twelve LC-MS/MS runs were performed on total proteins extracted from low CO_2-adapted wild-type and spinach RBCS hybrid strains, consisting of three biological replicates per strain and two technical replicates per biological replicate. This resulted in approximately 71,000 mass spectra confidently assigned to proteins present in wild-type (70,822 spectra) or spinach RBCS hybrid (71,001 spectra) strains. Excluding Chlamydomonas and spinach RBCS, the spectra were distributed across 1,376 predicted proteins (see Appendix 1 for a complete list of the proteins identified). Spectral counts per run varied from 11,456 to 12,158.

The mean spectral count (relative abundance) across all six samples was determined for each protein detected and showed the Chlamydomonas total protein samples to be dominated by a few highly abundant proteins (Fig. 5.1). When ranked by mean spectral count and excluding RBCS, 56.4% of counts were associated with the top 50 proteins, 70.6% with the top 100 proteins, 94.1% with the top 500 proteins and 5.9% with the remaining 884 proteins. Many of the most abundant proteins (22 of the top 50) were associated with the photosystems/light-harvesting complexes. Ribosomal proteins also accounted for 100 of the top 500 proteins. A second analysis of the mass spectrometric data was performed to search specifically for phosphorylation of proteins but no significant phosphopeptides were identified.

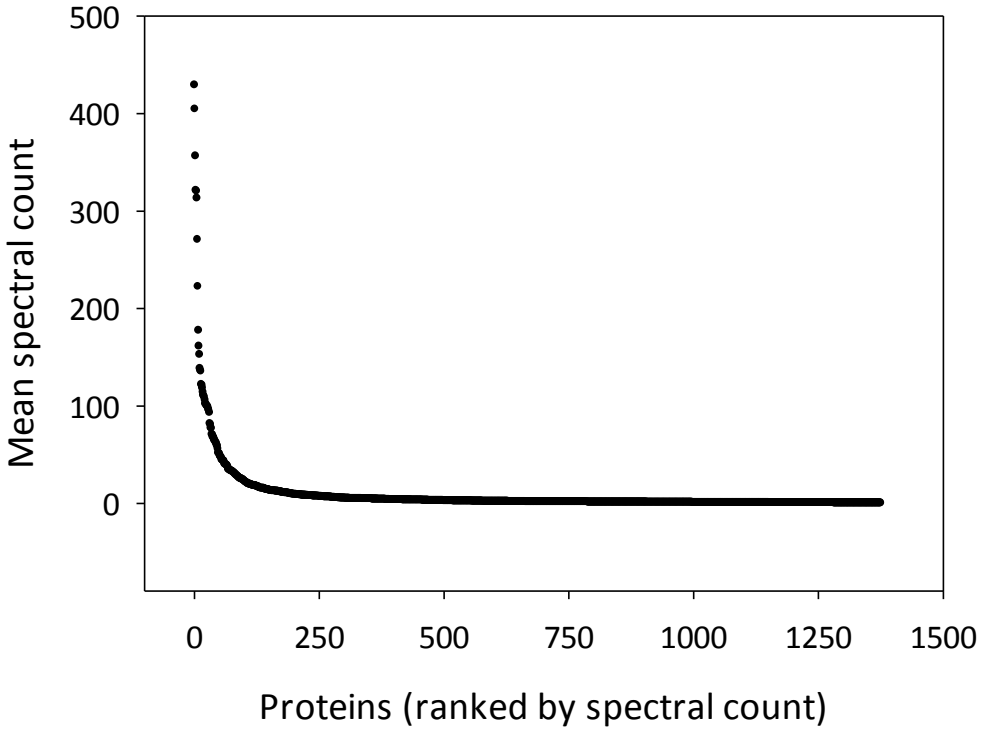


Figure 5.1 Rank abundance of total identified proteins from low CO_2-adapted wild-type and spinach RBCS hybrid strains. Mean spectral count is across two technical replicates of three wild-type and three spinach RBCS hybrid replicate cultures.

To investigate the similarity of wild-type and spinach RBCS hybrid samples, the mean spectral count for each protein detected in the spinach RBCS hybrid strain was plotted against the mean spectral count for the same protein in wild-type cells (Fig. 5.2). A line of best fit was plotted for all 1,376 proteins and showed a strong linear relationship ($R^2 = 0.97$, $y = 1.02x – 1.25$) between the relative abundance of proteins in the wild-type and spinach RBCS hybrid samples. Only 84 proteins were detected in significantly different abundance (G-test, $p<0.05$; closed circles) in the two strains while the remaining 1,292 proteins were not differentially expressed (open circles). Of the differentially expressed proteins, only eight had a mean spectral count in wild-type greater than 100, 19 had mean spectral counts between 25 and 99 and the remaining 57 proteins had mean spectral counts of less than 25. Together with the relatively small number of differentially expressed proteins (6-7% of the total proteins identified), this suggests that the proteomes of wild-type and spinach RBCS hybrid strains differ by only a discrete number of significantly differentially expressed proteins after short-term adaptation to low CO_2.

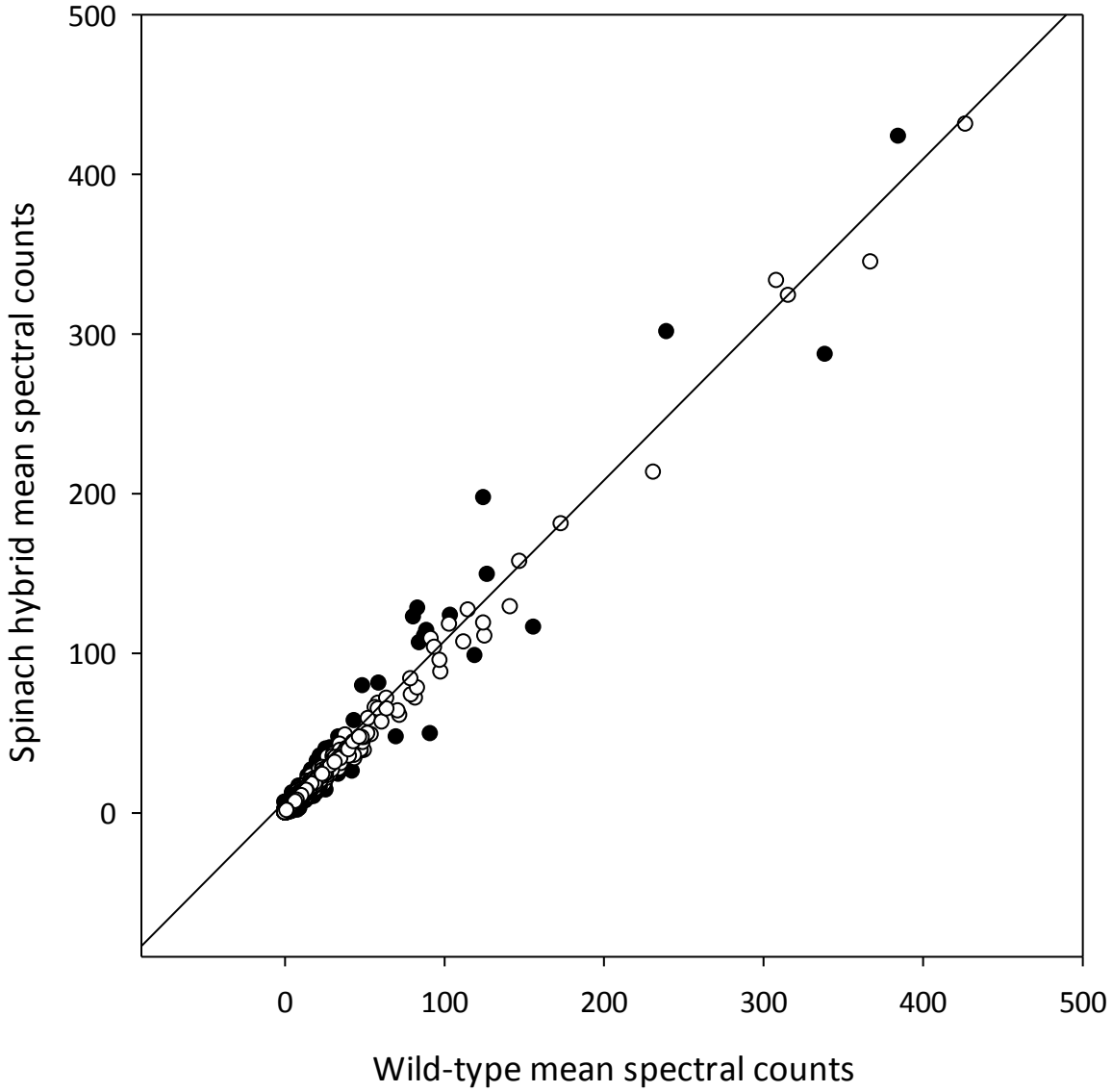
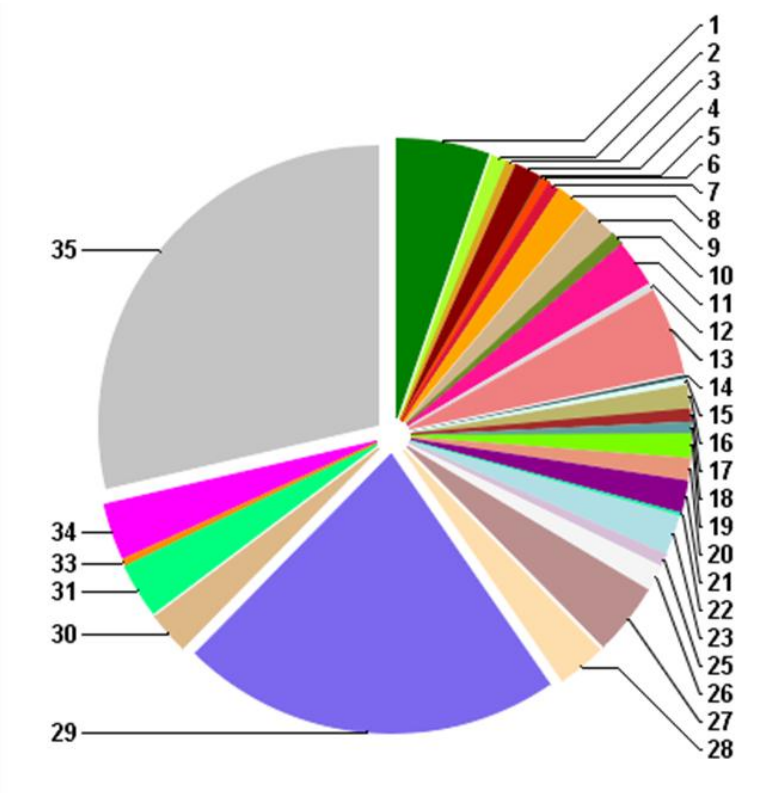


Figure 5.2 Comparison of 1,376 identified proteins (mean spectral counts) in low CO_2-adapted wild-type and spinach RBCS hybrid strains. Differentially expressed proteins ($p<0.05$; closed circles) and non-differentially expressed proteins (open circles) were identified by G-test (excluding Chlamydomonas and spinach RBCS). The equation for the line of best fit for the whole dataset is $y = 1.02x - 1.25$, $R^2 = 0.97$.

The total proteins detected in wild-type and spinach RBCS hybrid strains were also assigned to functional categories using Mercator, an automated online tool for annotation of plant nucleotide and protein sequence data (Lohse et al., 2014; Fig. 5.3). Over a quarter (29%) of the protein sequences were not assigned a functional group and the next largest group was protein assembly, modification and degradation (22%). Proteins involved in RNA binding/transcription, DNA repair and chromatin structure, photosynthesis, transport and primary metabolism (for example, amino acid, lipid and nucleotide metabolism) were also moderately abundant, each accounting for between 2.2 and 5.3% of the total proteins identified. Secondary metabolism and other metabolic pathways accounted for most of the remaining proteins.



1. photosynthesis 5.32%
2. major carbohydrate metabolism (e.g. starch) 0.92%
3. minor carbohydrate metabolism (e.g. sugar kinase) 0.46%
4. glycolysis 1.58%
5. fermentation 0.07%
6. gluconeogenesis/glyoxylate cycle 0.46%
7. pentose phosphate pathway 0.59%
8. TCA/organic transformation 1.77%
9. mitochondrial electron transport/ATP synthesis 1.97%
10. cell wall 0.72%
11. lipid metabolism 2.63%
12. N-metabolism 0.33%
13. amino acid metabolism 4.93%
14. S-assimilation 0.20%
15. metal handling 0.26%
16. secondary metabolism 1.38%
17. hormone metabolism 0.66%
18. co-factor and vitamin metabolism 0.66%
19. tetrapyrrole synthesis 1.31%
20. stress 1.18%
21. redox 1.77%
22. polyamine metabolism 0.13%
23. nucleotide metabolism 2.23%
25. C1-metabolism 0.66%
26. miscellaneous 1.38%
27. RNA 4.07%
28. DNA 2.76%
29. protein 21.94%
30. signalling 2.37%
31. cell 3.09%
33. development 0.33%
34. transport 3.22%
35. not assigned 28.65%.

Figure 5.3 Summary of functional groups (MapMan bins) assigned to all proteins detected using LC-MS/MS analysis of low CO_2-adapted wild-type and spinach RBCS hybrid proteomes.

5.2.2 Analysis of differentially expressed proteins

Of the 1,376 proteins detected, 158 (11.5%) were significantly overrepresented in low CO_2-adapted wild-type or spinach RBCS hybrid strains. More specifically, excluding both Chlamydomonas and spinach RBCS, 94 proteins (6.8%) were identified as differentially expressed ($p<0.05$) by performing a t-test on label-free intensity values (non-linearly proportional to the absolute abundance of the protein). A G-test was also performed on spectral counts (linearly proportional to

the relative protein abundance in the sample), excluding RBCS, and this identified 86 (6.1%) differentially expressed proteins ($p<0.05$). A total of 20 proteins were identified by both t- and G-tests and over twice as many proteins were more abundant in wild-type than spinach RBCS hybrid cells (108 and 47 proteins, respectively).

5.2.2.1 Combined analysis of LC-MS/MS (t-test and G-test) with 2D-DIGE

In total, 163 differentially expressed proteins were identified across both the 2D-DIGE and shotgun LC-MS/MS experiments (Fig. 5.4) and 23 proteins were identified as differentially expressed in at least two analyses (Table 5.1). In general, the direction of change (towards wild-type or spinach RBCS hybrid) was consistent across the analyses; however, some differences were identified. For example, ribosomal protein S3 (RPS3) was more abundant in the wild-type according to both t- and G-tests of the shotgun LC-MS/MS experiment but was more abundant in the spinach RBCS hybrid in the 2D-DIGE experiment. S-adenosylmethionine synthetase (METM) was more abundant in the wild-type according to the t-test and was also more abundant in wild-type in the larger but not the smaller of the two 2D-DIGE spots (915 but not 870). Chlorophyll a-b binding protein of PSII (LHCB4) was more abundant in the spinach RBCS hybrid strain according to both G-test and 2D-DIGE results. Oxygen-evolving enhancer protein 2 of PSII (PSBP) was more abundant in wild-type cells according to the G-test and more abundant in the spinach RBCS hybrid strain in three spots of the 2D-DIGE.

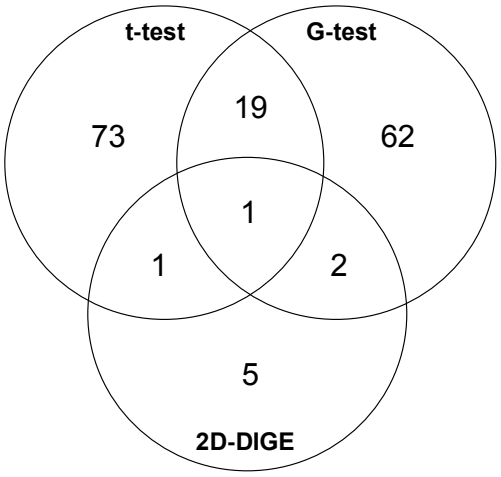


Figure 5.4 Summary of differentially expressed proteins identified by LC-MS/MS analysis (t-test and G-test) and 2D-DIGE experiments. 2D-DIGE includes unique proteins (top candidate for all 12 spots, 9 proteins in total, Table 4.4).

Of the 19 proteins identified in both t-test and G-test analyses (Table 5.1), eleven proteins were consistently more abundant in wild-type and five proteins were consistently more abundant in spinach RBCS hybrid strains. Proteins that were more abundant in the wild-type include those associated with the CCM (CAH5, LCIB), protein synthesis and folding (RPS12, CYN20-2), photosynthesis and respiration (PSB28, COX13), oxidative stress/signalling (PRX1) and the circadian clock (CRB1) as well as three proteins of unknown function. The proteins that were more abundant in the spinach RBCS hybrid mutant were associated with PSII repair (REP27) and primary metabolism (HSD1, PHOA) as well as two proteins of unknown function. The remaining three proteins (40S ribosomal protein S3a, ALDCHL, MDH4) were more abundant in the wild-type according to label-free intensity values and more abundant in the spinach RBCS hybrid strain according to spectral counts.

Overall, of the 23 proteins in Table 5.1, 17 were identified as consistently more abundant in either the wild-type or spinach RBCS hybrid strain by two different analyses. Inconsistencies between 2D-DIGE and LC-MS/MS experiments may be due to methodological differences. 2D-DIGE is able to differentiate between differently modified forms of the same protein on the gel whereas shotgun LC-MS/MS does not include this initial separation step. Differences between t-test and G-test results are due to the differences between absolute and relative abundance of the proteins. In this experiment, both tests were considered useful for providing an initial list of interesting candidates. As a further test, samples were normalised to the control proteins PSBO and GBLP but this did not significantly alter the results.

Table 5.1 Summary of proteins identified as differentially expressed in at least two analyses (LC-MS/MS t-test, LC-MS/MS G-test and 2D-DIGE). For each test the relative abundance of the protein (Up in W or S), UniProt identifier, description of the protein including gene name and the broad functional group are given. Proteins that were either not detected or not found to be differentially expressed are marked as '-'. Proteins that were consistently more abundant in either strain across two analyses are shaded grey. Note that A8HYU5/METM was identified in two spots on the 2D gels, one more abundant in wild-type, the other more abundant in the spinach RBCS hybrid mutant.

Strain in which DE proteins are up			**UniProt ID**	**Description (Gene name)**	**Functional group**
T	**G**	**2D**			
W	W	S	A8I4P5	Ribosomal protein S3 (RPS3)	protein synthesis
W	-	W S	A8HYU5	S-adenosylmethionine synthetase (METM)	primary metabolism
-	S	S	A8J6D1	Chlorophyll a-b binding protein of photosystem II (LHCB4)	photosynthesis
-	W	S	P11471	Oxygen-evolving enhancer protein 2 of photosystem II (PSBP)	photosynthesis
W	W	-	Q0ZAI6	Low CO_2-inducible protein B (LCIB)	CCM
W	W	-	Q39590	Beta carbonic anhydrase (CAH5)	CCM
W	W	-	A8IGE2	Predicted protein	unknown
W	W	-	A8I2V3	2-cys peroxiredoxin (PRX2)	stress
W	W	-	A8IY43	Peptidyl-prolyl *cis-trans*isomerase (CYN20-2)	protein synthesis
W	W	-	A8III5	Photosystem II reaction centre Psb28 protein (PSB28)	photosynthesis
W	W	-	A8IVF6	Cytochrome c oxidase 12 kDa subunit (COX13)	respiration/energy
W	W	-	A8IGM2	Aspartyl aminopeptidase-like protein (AAP1)	protein synthesis
W	W	-	A8J9T0	40S ribosomal protein S12 (RPS12)	protein synthesis
W	W	-	A8J3W1	Predicted protein	unknown
W	W	-	Q6RBZ1	Circadian RNA-binding protein CHLAMY 1 subunit C1 (CRB1)	other
W	S	-	Q42690	Fructose-bisphosphate aldolase 1, chloroplastic (ALDCHL)	respiration/energy
W	S	-	A8JHU0	Malate dehydrogenase (MDH4)	respiration/energy
W	S	-	A8HS48	40S ribosomal protein S3a	protein synthesis
S	S	-	A8I3J7	Homoserine dehydrogenase (HSD1)	primary metabolism
S	S	-	Q2VA40	Phosphorylase (PHOA)	primary metabolism
S	S	-	A8IWP6	Predicted protein	unknown
S	S	-	A1YSB4	Photosystem II repair protein (REP27)	photosynthesis
S	S	-	A8J7X3	Predicted protein	unknown

5.2.2.2 Analysis by protein function

5.2.2.2.1 CCM and low CO_2-inducible proteins

Seven proteins with confirmed roles in the CCM (CAH1, CAH3, CAH5, CCP1, LCIA, LCIB, LCIC) as well as four low CO_2-induced proteins of unknown function (LCI2, LCI5, LCI23 and LCI33) were identified using shotgun LC-MS/MS (Fig. 5.5). Notably, the Ci transporters CCP2, HLA3 and LCI1 were not detected and LCIA was present in only very low abundance (1-2 and 0-1 spectra per sample in wild-type and spinach RBCS hybrid, respectively), although CCP1 was detected at slightly higher levels (4-15 and 0-3 spectra per sample in wild-type and spinach RBCS hybrid, respectively). Using spectral counts as an estimation of the relative abundance of CCM proteins in both strains, LCIB and LCIC were much more abundant than any of the other CCM-related proteins (spectral counts for wild-type and spinach RBCS hybrid strains ranged from 22-46 for LCIB and 38-54 for LCIC). The Rubisco large subunit (rbcL) was present in high abundance (spectral counts of 48-155) and was found to be significantly more abundant in the spinach RBCS hybrid strain based on spectral counts but not label-free intensity. The gene regulatory proteins CIA5 and LCR1 were not detected in either strain. The control proteins used for immunoblots, histone H3 and PSBO, were also detected but were not identified as differentially expressed (Fig. 5.5).

Although two proteins of unknown function, LCI5 and LCI23, were only detected in relatively low abundance, the differences in abundance between wild-type and spinach RBCS hybrid strains were significant. As mentioned in the Introduction, LCI5 is a small (32 kDa), chloroplast-localised protein that is expressed under both high and low CO_2 but is specifically phosphorylated in a low CO_2- and redox-dependent manner (Turkina et al., 2006). LCI5 has three tandem repeats (60 amino acid residues each) and no annotated structural domains. Phyre^2 predicts 100% disorder in the structure of LCI5 (Kelley and Sternberg, 2009), which suggests that LCI5 may function in signalling or as a molecular chaperone (Sickmeier et al., 2007). LCI5 is 74.8% similar to a protein of unknown function from *Volvox carteri* (Vocar20008598m.g) and has no other known homologues. Even less is known about the septin-like protein, LCI23, which has putative GTPase activity and a possible role in signal transduction, the cell cycle or cytoskeleton. LCI23 is predicted to be localised to the endoplasmic reticulum or

plasma membrane with one transmembrane helix at the C-terminus. Only two genes are assigned to this family in the current Chlamydomonas genome (v.5.5).

LCI2 and LCI33, the other two low CO_2-induced proteins of unknown function, were detected at even lower levels (mean spectral counts were <1). There appeared to be no difference in the abundance of LCI2 between wild-type and spinach hybrid strains while LCI33 appeared to be more abundant in the spinach RBCS hybrid based on label-free intensity values. LCI2 is predicted to be a chloroplast-localised ubiquitin-conjugating protein with homologues in other species of microalgae as well as higher plants (>50% similarity) although an alternative functional prediction (Argot^2, Falda et al., 2012) is that LCI2 is involved in unsaturated fatty acid biosynthesis in the chloroplast and may be an integral membrane component. Unlike LCI2, LCI33 has predicted homologues of only low similarity (<10%). LCI33 is predicted to be targeted to the nucleus and has a C2H2 zinc finger (DNA binding) motif, which indicates a potential role in transcription.

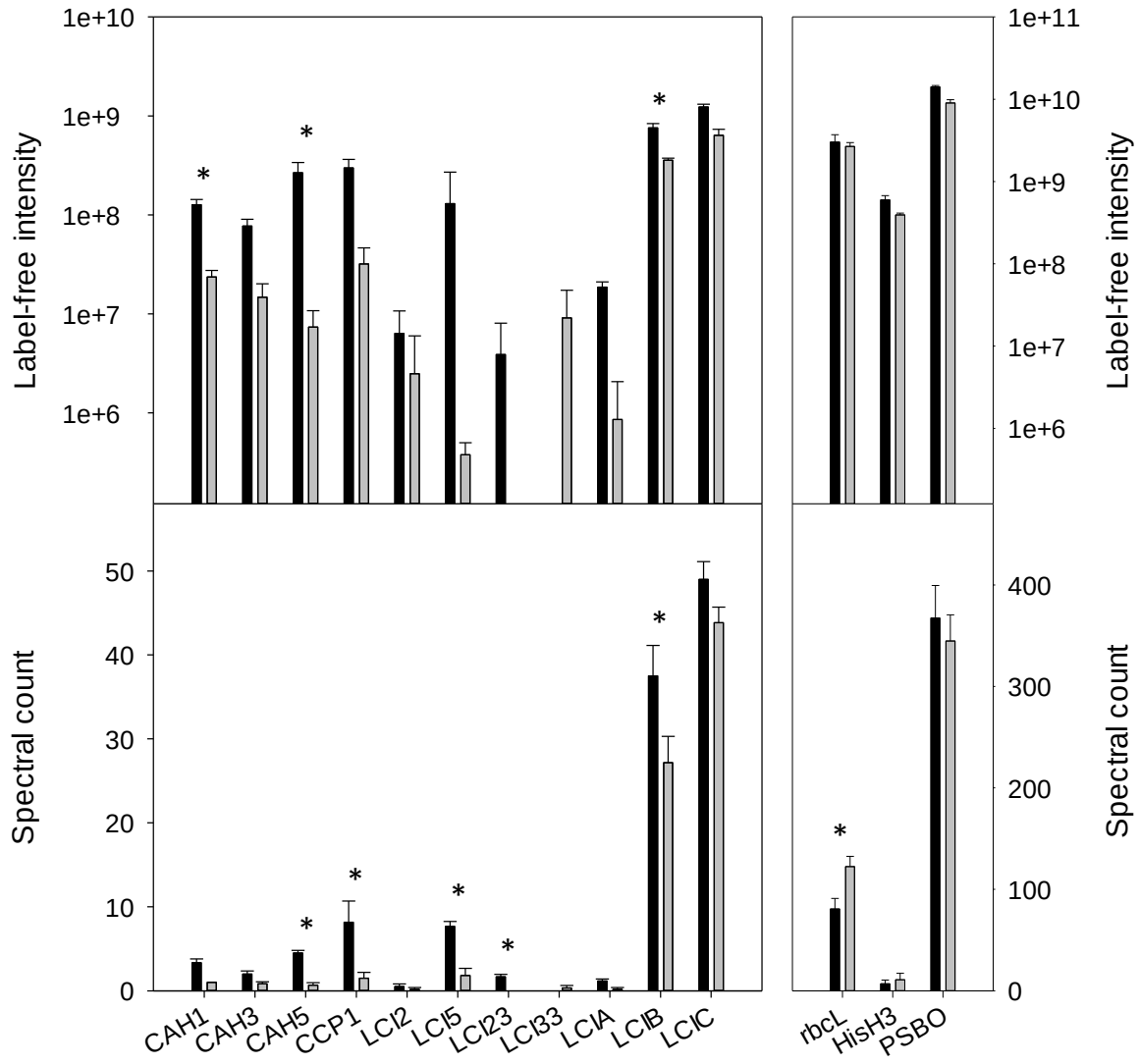


Figure 5.5 Relative quantification of CCM proteins detected in low CO_2-adapted wild-type (black) and spinach RBCS hybrid (grey) strains using LC-MS/MS. CCM and CO_2-responsive proteins are shown on the left and non-CO_2-responsive (rbcL) and control proteins (His H3 and PSBO) are shown on the right. Note the different y-scales. Two measures of protein abundance were used: label-free intensity which is proportional to absolute abundance (top) and spectral count which is a measure of the relative abundance in the sample (bottom). Values shown are the mean and standard error of six measurements (two technical replicates for three replicate cultures). Asterisks represent statistically significant differences ($p<0.05$) in protein abundance between the two strains according to a t-test on label-free intensity values or a G-test on spectral counts.

In total, six CCM proteins (CAH1, CAH5, CCP1, LCI5, LCI23 and LCIB) were significantly less abundant in low CO_2-adapted spinach RBCS hybrid cells than in the wild-type. Except for LCI33, all the other CCM proteins detected appear to present in somewhat lower abundance in the spinach RBCS hybrid strain, although the difference was not statistically significant. Where data for immunoblots and qRT-PCR were available, they were also compared to LC-MS/MS results (Table 5.2). CAH1, LCIB and LCIC mRNA and protein were generally more abundant in the wild-type strain across all three experiments. The

spinach RBCS hybrid also showed reduced expression of CCP1 and LCIA at the mRNA and protein level. No large differences were detected in CAH3 expression at the mRNA or protein level, although wild-type expression may have been slightly higher than the spinach RBCS hybrid. *LCI5* mRNA was more abundant in the spinach RBCS hybrid but LCI5 protein was less abundant in this strain. CAH5, CCP2, CIA5, LCI1, LCI23 and LCR1 were only detected in a single experiment (LC-MS/MS or qRT-PCR alone) and CAH6 and HLA3 were not detected using LC-MS/MS, immunoblots or qRT-PCR.

Table 5.2 Comparison of low CO_2-adapted wild-type and spinach RBCS hybrid strains using LC-MS/MS analysis, immunoblotting and qRT-PCR. p-values are shown for proteins identified as differentially expressed by t-test or G-test. Proteins were either more abundant in the wild-type (W) or spinach RBCS hybrid strain (S). Letters in parentheses indicate that proteins appeared to be somewhat more abundant in this strain but p>0.05. Other differences were clearly not significant (o) or the protein/gene transcript was not detected (n.d.). Measurements of control proteins PSBO and histone H3 (His H3) are also included. CAH6 and HLA3 were not detected or probed in any experiment.

Protein	**LC-MS/MS**			**Blot W/S**	**qRT-PCR W/S**
	t-test p-value	**G-test p-value**	**Up in W/S**		
CAH1	0.045	0.06	W	W	W
CAH3	0.13	0.33	(W)	n.d.	o
CAH5	0.015	8.2×10^{-4}	W	n.d.	n.d.
CCP1	0.54	4.3×10^{-6}	W	n.d.	W
CCP2	n.d.	n.d.	n.d.	n.d.	W
CIA5	n.d.	n.d.	n.d.	n.d.	S
LCI1	n.d.	n.d.	n.d.	n.d.	(W)
LCI5	0.13	1.3×10^{-4}	W	n.d.	S
LCI23	0.067	7.7×10^{-3}	W	n.d.	n.d.
LCIA	0.36	0.17	(W)	n.d.	W
LCIB	0.013	0.029	W	W	W
LCIC	0.16	0.48	(W)	W	(W)
LCR1	n.d.	n.d.	n.d.	n.d.	(W)
rbcL	0.94	1.2×10^{-10}	S	S	o
PSBO	0.11	0.19	o	o	o
His H3	0.17	0.18	o	o	n.d.

5.2.2.2.2 Other proteins of known function

Differentially expressed proteins were manually classified according to broad functional groups and whether they were more abundant in wild-type or spinach RBCS hybrid cells (Fig. 5.6A). As discussed above, all five of the differentially expressed CCM proteins were more abundant in wild-type cells. Fifteen proteins associated with the photosystems were also differentially expressed between the

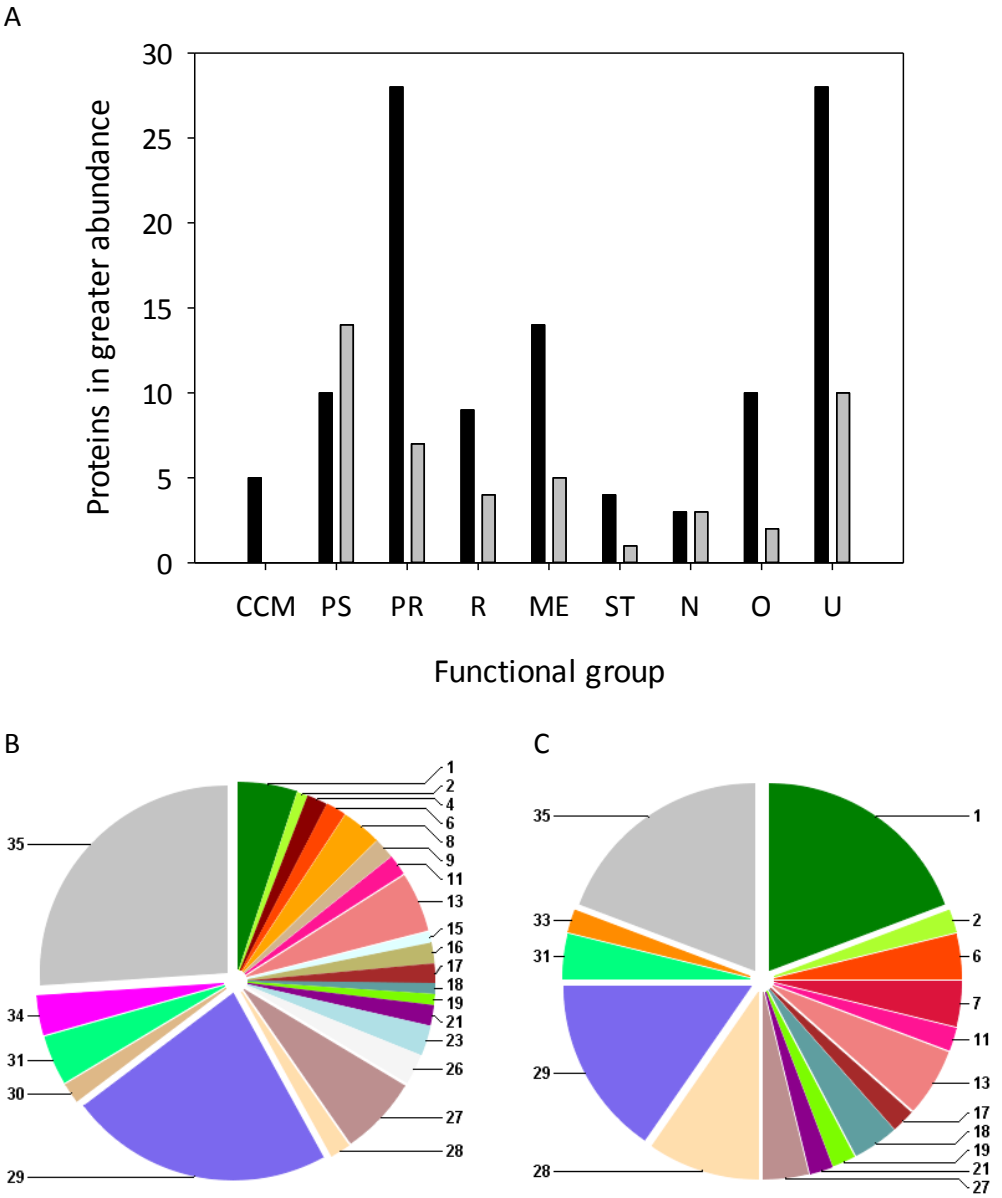
two strains. Three proteins associated with PSI were more abundant in wild-type (psaC, psaD, PSAN) and five were more abundant in spinach RBCS hybrid cells (psaB, LHCI-3, LHCI-5, FNR1, psaA). Three proteins associated with PSII were more abundant in wild-type (PSB28, PSBQ, PSBP) and four were more abundant in spinach RBCS hybrid (REP27, psbD, LHCB4, psbC). Of these, the chlorophyll binding proteins were more abundant in the spinach RBCS hybrid while reaction centre and oxygen evolving enhancer proteins were more abundant in the wild-type. Apocytochrome f (petA) was more abundant in the spinach RBCS hybrid and cytochrome b6f subunit V (PETO) was more abundant in the wild-type. Two ATP synthase subunits (atpA, atpB) were more abundant in the spinach RBCS hybrid. Two fructose-bisphosphate aldolases (ALDCHL, A8JCY4) were more abundant in the wild-type and may be involved in the CBB cycle. Apart from Rubisco (rbcL), no other enzymes from the CBB cycle itself were identified as differentially expressed. However, the CBB regulatory protein, CP12, was found to be more abundant in spinach RBCS hybrid cells.

Protein synthesis appeared to be generally upregulated in the wild-type strain in response to low CO_2. Four times as many proteins were more abundant in wild-type than in spinach RBCS hybrid (28 and 7 proteins, respectively). These proteins included ribosomal proteins, elongation factors, RNA processing factors and enzymes involved in protein folding. Other proteins involved in gene expression such as transcription factors and chromatin remodelling proteins were not generally found to be differentially expressed. The exception was one basal transcription factor (BTF3/A8JBX6) which was more abundant in wild-type cells. This transcription factor appears highly conserved with greater than 60% similarity to transcription factors from diverse plant species (for example, *Fragaria vesca*, *Zea mays* and *Populus tricocarpa*), which is in contrast to the relatively unique CCM master regulator, CIA5, and CO_2-responsive transcription factor, LCR1. CIA5 has homologues only in *Volvox carteri* (31.2% similarity) and *Coccomyxa subelleipsoidea* C-169 (8.6% similarity) while the closest homologues to LCR1 in Chlamydomonas and *V. carteri* are less than 10% similar. Unlike *LCR1*, *BTF3* was not found to be CO_2-responsive in the transcriptome datasets of Brueggeman et al. (2012) and Fang et al. (2012).

As with protein synthesis, the glycolysis and respiration pathways also appeared to be somewhat more highly represented in wild-type cells. Nine enzymes involved in glycolysis, the respiratory electron transport chain and the tricarboxylic acid (TCA) cycle were more abundant in the wild-type while only four enzymes were more abundant in the spinach RBCS hybrid strain. Finally, enzymes mainly involved in primary metabolism, including the amino acid biosynthetic pathway, were preferentially upregulated in wild-type cells (14 proteins) compared to spinach RBCS hybrid cells (five proteins).

Other proteins of interest that were more abundant in the wild-type include: cytoskeletal proteins (α- and β-tubulin), signalling proteins (calcium sensing receptor, Ran-like small GTPase) and a regulator of gene expression (subunit C1 of the circadian RNA-binding protein CHLAMY 1). There was no difference in the number of nuclear-localised proteins (nucleosome and nucleolar proteins) differentially expressed in wild-type or spinach RBCS hybrid cells (three each).

More detailed functional analysis of the differentially expressed proteins was also performed using Mercator (Fig. 5.6B and 5.6C). Broadly supporting the manual analysis, proteins involved in glycolysis, TCA cycle and mitochondrial electron transport/ATP synthesis were preferentially expressed in the wild-type while photosynthetic proteins tended to be in greater abundance in the spinach RBCS hybrid. In addition, metal handling, secondary metabolism, nucleotide metabolism, signalling and transport processes all appeared to be preferentially upregulated in the wild-type while proteins involved in DNA binding/chromatin structure, oxidative pentose phosphate pathway and development were more abundant in the spinach RBCS hybrid.



1. photosynthesis
2. major carbohydrate metabolism
4. glycolysis
6. gluconeogenesis/glyoxylate cycle
7. pentose phosphate pathway
8. TCA/org transformation
9. mitochondrial electron transport/ ATP synthesis
11. lipid metabolism
13. amino acid metabolism
15. metal handling
16. secondary metabolism
17. hormone metabolism
18. co-factor and vitamin metabolism
19. tetrapyrrole synthesis
21. redox
23. nucleotide metabolism
26. miscellaneous
27. RNA
28. DNA
29. protein
30. signalling
31. cell
33. development
34. transport
35. not assigned

Figure 5.6 Functional classification of differentially expressed proteins in low CO_2-adapted wild-type and spinach RBCS hybrid cells. (A) Proteins identified as more abundant in wild-type (black) or spinach RBCS hybrid (grey) strains. Proteins were manually classified into broad functional groups: CCM, low CO_2-induced; PS, photosynthesis; PR, protein synthesis; R, energy and respiration; ME, other metabolic processes; ST, stress/redox; N, nuclear-localised; O, other; U, unknown. More detailed functional groups were also assigned using Mercator to proteins that were more abundant in (B) wild-type or (C) spinach RBCS hybrid cells.

5.2.2.2.3 Proteins of unknown function

Nearly one quarter of the differentially expressed proteins identified in this study were predicted proteins of unknown function. Of these, 28 were more abundant in the wild-type while only 10 were more abundant in the spinach RBCS hybrid. Given the reduced accumulation of known CCM proteins in response to low CO_2 in the spinach RBCS hybrid, the 28 proteins that are more abundant in the wild-type are of particular interest as potential novel CCM candidates. In order to narrow down the list of candidates, the localisation, homologies and functional/structural annotation of each differentially expressed protein were analysed (Table 5.3).

Overall, most of the predicted proteins were present in low abundance in both wild-type and spinach RBCS hybrid cells. The maximum mean spectral count was below five for 32 of the 38 proteins, which places them in a similar range to other low CO_2-induced proteins (CAH1, CAH3, CAH5, LCI2, LCI23, LCI33 and LCIA). The mean spectral counts for the remaining six proteins varied from 5.2 to 35.5, which was similar to the abundance of other CCM-related proteins (CCP1, LCI5, LCIB and LCIC) and the control protein histone H3. All the predicted proteins were present in much lower abundance than cytoskeletal proteins such as β-tubulin (mean spectral count of 70) and photosynthetic proteins rbcL and PSBO (mean spectral counts of 120 and 370, respectively). Fifteen of the proteins of unknown function were predicted to be targeted to the chloroplast, eight to the cytosol, five to the mitochondria, five to the nucleus and the remaining five to other cellular compartments. Only three proteins were predicted to have transmembrane domains, which may implicate them in transport processes. The gene transcripts encoding 22 of these proteins are CO_2-responsive (Brueggeman et al., 2012; Fang et al., 2012).

Computational predictions of protein structure (Phyre^2) yielded similar results to predictions of protein function (Argot^2) for fifteen proteins and no or very low confidence results for a further five proteins. The remaining proteins either had conflicting or lower confidence predictions. Three proteins with predicted nuclear localisation were all more abundant in wild-type cells (A8HP50, A8IY40 and A8JEA7) and were identified as binding DNA or RNA, suggesting they may have a role in low CO_2-responsive gene expression. Other proteins that were preferentially accumulated in the wild-type have putative roles in amino acid

synthesis (A8J3W1 and A8J3Y6), tRNA synthesis (A8IBN3) and proteolysis (A8JAW4, A8IGM2, A8IUN8).

Two proteins that may be involved in photosynthesis were also identified as more abundant in wild-type cells. One protein (A8HNG8) was predicted to be localised to the thylakoid lumen based on a manual analysis of the transit peptide and is likely to be associated with the oxygen evolving enhancer complex of PSII. The second protein (A8J995) is predicted to be a pentapeptide repeat protein with structural homology to a thylakoid lumenal protein from Arabidopsis as well as proteins from (CCM-positive) cyanobacteria *Cyanothece* and *Nostoc* species. Interestingly, one protein that was more abundant in the spinach RBCS hybrid (A8IQG4) had homology to an acetate transporter, although this was not a high confidence prediction.

Table 5.3 Predicted localisation, function and structure of the 38 predicted proteins identified as more abundant in low CO_2-adapted wild-type (W) or spinach RBCS hybrid (S) cells. Mean spectral counts (S.C.) for each strain are also given. Localisation was performed using WoLF PSORT (Horton et al., 2007) and predicted chloroplast-localised proteins were manually checked for dual partite thylakoid signal peptides (thyl) as per Karlsson et al. (1998). Other predicted localisations were abbreviated as follows: chlo, chloroplast; cyto, cytosol/cytoplasm; ER, endoplasmic reticulum; extr, extracellular; mito, mitochondrion; nucl, nucleus; plas, plasma membrane; vacu, vacuole. Transmembrane domains (TMD) were predicted using TMHMM v2.0 (Krogh et al., 2001). Phytozome function was determined using the PANTHER and GO annotations available for the latest genome (v5.5). Argot^2 (Falda et al., 2012) was used to predict protein function and results with a score ≥200 are listed from highest to lowest (annotations in parentheses indicate a low score or no hit found). Phyre^2 (Kelley and Sternberg, 2009) was used to predict protein structure and the top hits (high % confidence, high % coverage) are listed in each case. Proteins with transcripts identified as CO_2-responsive are listed under the heading 'CO_2 DE': B, Brueggeman et al., (2012);F, Fang et al., (2012); U, upregulated in response to low CO_2; D, downregulated in response to low CO_2. Proteins with consistent annotations across the three prediction programs are shaded in grey. Protein are listed alphabetically by UniProt ID.

UniProt ID	Up in W/S	WT S.C.	Sp S.C.	WoLF PSORT Local.	TMD	Phytozome function (PANTHER; KEGGORTH/KEG/GO)	Argot^2 (molecular function; biological process; cellular component)	Phyre^2 structure (%confidence; %coverage)	CO_2 DE
A8HMM7	S	0.0	0.8	chlo	1	no functional annotation	(not found); (not found); (not found)	no high confidence hits (<9%)	B (D)
A8HNG8	W	5.2	1.7	chlo (thyl)	1	oxygen evolving enhancer protein 3 (PSBQ) (uncharacterised thylakoid lumenal polypeptide)	calcium ion binding; photosynthesis; photosystem II oxygen evolving complex	no high confidence hits (<12%)	B (D) F(D)
A8HP50	W	0.5	0.0	nucl	0	RNA-binding protein LUC7-related; spliceosome subunit	mRNA binding; mRNA splice site selection; U1 snRNP	rubrerythrin (88%; 28%), C2H2, DNA/RNA binding (82%; 8%)	B (D) F(D)
A8HSJ6	W	0.8	0.0	extr	0	dolichyl-phosphate-mannose-protein mannosyltransferase, stromal cell-derived factor 2-related; membrane	mannosyltransferase/calcium channel; PAMP/defense response to fungus or bacterium; membrane/ER	stromal cell-derived factor 2-like protein (100%; 82%)	B (D)
A8I0A3	W	0.8	0.0	cyto	0	sialic acid synthase-related; phospholipid:diacylglycerol acyltransferase	acetyl transferase; metabolic process; cytoplasm	Galactoside acetyltransferase-like (100%; 95%), one TMD spanning PM	-
A8I3W8	W	0.7	0.0	chlo/ extr	0	no functional annotation	(not found); (not found); (not found)	no high confidence hits (<41%)	F (U)
A8I9T2	S	0.0	0.5	mito	0	Kinesin motor protein, kinesin-9 (Kif9) family	microtubule motor/ATP binding; microtubule-based movement; kinesin complex/microtubule/cytoplasm	chimera of maltose-binding periplasmic protein and kinesin (100%; 40%), P-loop containing nucleoside triphosphate hydrolases	-

Table 5.3 continued

UniProt ID	Up in W/S	WT S.C.	Sp S.C.	WoLF PSORT Local.	TMD	Phytozome function (PANTHER; KEGGORTH/KEG/GO)	Argot[2] (molecular function; biological process; cellular component)	Phyre[2] structure (%confidence; %coverage)	CO_2 DE
A8IBN3	W	5.2	3.3	cyto	0	prolyl tRNA synthetase	proline -tRNA ligase/ amino acyl-tRNA ligase; tRNA aminoacylation/ translation; cytoplasm	bifunctional glutamate/proline-tRNA ligase (100%; 98%)	F (D)
A8ICU7	S	0.0	0.8	mito	0	serine protease	serine-type peptidase; proteolysis; thylakoid lumen/mitrochondrion	photosystem II d1 protease (100%; 97%)	B (D)
A8IGE2	W	1.2	0.0	ER/ plas	0	survival protein surE, 5'-nucleotidase	hydrolase; dephosphorylation/metabolic process; cytoplasm	5'-nucleotidase surE (100%; 79%)	B (D) F(D)
A8IGM2	W	1.5	0.0	chlo	0	aspartyl aminopeptidase, zinc metalloprotease	zinc ion binding/hydrolase/ aminopeptidase; proteolysis; cystosol/vacuolar membrane/ vacuole	aspartyl aminopeptidase (100%; 80%)	-
A8IGV4	W	1.0	0.0	cyto	0	phosphoenolpyruvate dikinase-related	kinase/ATP binding/catalytic; phosphorylation/starch catabolic process; chloroplast envelope	pyruvate phosphate dikinase, N-terminal domain (100%; 82%)	B (U) F (U)
A8IHW6	W	1.0	0.0	nucl/ plas	0	no functional annotation	(ATP binding); (vesicle-mediated transport); mitochondrion	no high confidence hits (<47%)	B (D)
A8IQG4	S	22.3	35.5	cyto	6	GPR1/FUN34/yaaH family, membrane, mannosyl-3-phosphoglycerate phosphatase	(oxidoreductase); (plasma membrane acetate transport); membrane	no high confidence hits (<33%)	-
A8IRW1	W	1.2	0.0	vacu	0	no functional annotation	(catalytic activity); (metbolic process); (not found)	DNA repair protein Rad51, N-terminal domain (94%; 26%)	-
A8ITZ2	W	0.7	0.0	chlo	0	no functional annotation	Ran GTPase activator; positive regulation of Ran GTPase activity; (kinetochore; cytoplasm)	No high confidence hits (<40%)	-
A8IUN8	W	0.5	0.0	chlo/ cyto	0	26s proteasome non-ATPase regulatory subunit	(not found); multicellular organismal development; proteasome complex/ nucleus/cytoplasm	26s proteasome regulatory subunit rpn5 (100%; 83%)	-

Table 5.3 continued

UniProt ID	Up in W/S	WT S.C.	Sp S.C.	WoLF PSORT Local.	TMD	Phytozome function (PANTHER; KEGGORTH/KEG/GO)	Argot[2] (molecular function; biological process; cellular component)	Phyre[2] structure (%confidence; %coverage)	CO_2 DE
A8IV51	W	0.7	0.0	mito	0	prolyl 4-hyrdoxylase alpha subunit; oxidoreductase activity	itron ion binding/oxidoreductase; oxidation-reduction process; endosome/Golgi	prolyl-4 hydroxylase (100%; 73%)	F (D)
A8IWP6	S	0.0	1.3	nucl	0	calcium independent phospholipase A2 (IPLA2)-related	(not found); lipid metabolic process; (not found)	VipD (phospholipase) (100%; 50%)	F (U)
A8IY40	W	0.7	0.0	chlo/ cyto	0	RNA recognition motif	nucleic adic binding; (not found); nucleus	RNA binding protein (100%; 95%), splicing factor (100%; 94%), ribonucleoprotein (100%; 91%)	B (D) F(D)
A8J3W1	W	1.5	0.0	cyto	0	asparagine synthetase	(transferase activity); asparagine biosynthetic process/cellular amino acid biosynthetic process; cytosol/nucleus	asparagine synthetase b (100%; 86%)	-
A8J3Y6	W	6.7	1.8	mito	0	AIR synthase domain; Phosphoribosylformylglycinamidine synthase	phosphoribosylformylglycinamidine synthase/catalytic; 'de novo' IMP biosynthetic process; cytoplasm/chloroplast/ mitochondrion	phosphoribosylformylglycinamidine synthase (100%; 89%)	-
A8J462	S	0.0	1.0	chlo	0	alanyl tRNA synthetase	zinc ion binding/alanine-tRNA ligase; tRNA aminoacylation/translation; cytoplasm/chloroplast/ mitochondrion	alanyl-tRNA synthetase (100%; 70%)	B (D)
A8J4M0	W	1.8	0.2	nucl	0	no functional annotation	(signal transducer); (signal transduction); (membrane)	keratin, type II cytoskeletal 5 (94%; 47%)	B (D)
A8J5B8	S	20.3	32.2	cyto	0	AMP-activated protein kinase, gamma subunit	adenyl nucleotide binding; (metabolic process); (membrane/cytoplasm)	nuclear protein snf4 (100%; 72%), AMP binding, CBS domain pair	B (D) F(D)

Table 5.3 continued

UniProt ID	Up in W/S	WT S.C.	Sp S.C.	WoLF PSORT Local.	TMD	Phytozome function (PANTHER; KEGGORTH/KEG/GO)	Argot[2] (molecular function; biological process; cellular component)	Phyre[2] structure (%confidence; %coverage)	CO_2 DE
A8J6M8	W	0.5	0.0	vacu	0	N-linked oligosaccharide processing/ glucosidase II beta subunit-like protein	(metal ion binding); N-glycan processing/defense response to bacterium; ER	low-density lipoprotein receptor-related protein (100%; 80%)	B (D)
A8J7P4	W	2.0	1.2	cyto	0	no functional annotation	(ATP binding); (ATP catabolic process); (cytoplasm)	no high confidence hits (<37%)	B (D) F(D)
A8J7X3	S	0.0	1.3	plas	0	n/a	metal ion binding/(hydrolase); (signal transduction); (membrane)	ferredoxin reductase FAD-binding domain-like (95%, 4%)	F (U)
A8J995	W	0.7	0.0	chlo	0	pentapeptide repeats (8 copies)	(not found); (not found); (chloroplast thylakoid lumen/membrane)	secreted effector protein (100%; 64%), A.t. pentapeptide repeat thylakoid lumenal protein (100%; 54%), cyanobacterial homologues (Cyanothece and Nostoc sp.)	-
A8JAW4	W	14.0	10.0	cyto	0	protease family M24, proliferation-associated protein 2G24; metallopeptidase	metalloexopeptidase/aminopeptidase; proteolysis; nucleolus/plasma membrane	putative curved dna-binding protein (100%; 89%), ribosome (100%; 91%), aminopeptidase (100%; 87%)	B (D)
A8JC15	W	0.5	0.0	chlo	0	no functional annotation	hydrolyzing O-glycosyl compounds; carbohydrate metabolic process; (not found)	no high confidence hits (<32%)	F (U)
A8JCS8	W	0.5	0.0	chlo (thyl)	0	ACT domain (metabolism); amino acid binding	amino acid binding; metabolic process; chloroplast stroma	glycine cleavage system transcriptional repressor (100%; 59%)	-
A8JD42	W	0.8	0.0	chlo (thyl)	0	domain of unknown function 1350 (in cyanobacteria and plants, may be involved in photosynthesis)	hydrolase; (metabolic process); chloroplast	esterase (98%; 52%), giberellin receptor (98%; 59%), hydrolase	-
A8JEA7	W	4.5	1.0	nucl	0	regulator of chromatin condensation	DNA binding/metal ion binding; (negative regulation of GTPase activity); nucleolus	regulator of chromosome condensation (100%; 82%)	-

Table 5.3 continued

UniProt ID	Up in W/S	WT S.C.	Sp S.C.	WoLF PSORT Local.	TMD	Phytozome function (PANTHER; KEGGORTH/KEG/GO)	Argot[2] (molecular function; biological process; cellular component)	Phyre[2] structure (%confidence; %coverage)	CO_2 DE
A8JEP0	S	0.0	1.0	chlo	0	no functional annotation	(not found); (regulation of transcription, DNA-templated); (nucleus)	no high confidence hits (<18%)	-
A8JEQ7	W	0.8	0.0	chlo	0	rhodanese-like domain	(tranferase activity); (not found); chloroplast envelope	thiosulfate sulfurtransferase (100%; 87%)	B (D)
A8JFF0	W	0.7	0.0	chlo	2	no functional annotation	(not found); (not found); (integral component of membrane)	no high confidence hits (<21%)	-
Q84X71	S	0.8	4.0	mito	0	ankyrin repeat and protein kinase domain-containing protein	(not found); (not found); plasma membrane	burrh DNA-binding protein (100%; 95%), ankyrin repeat	-

5.2.3 Comparison of transcriptome and proteome datasets

5.2.3.1 Low CO_2-adapted wild-type and higher plant RBCS hybrid strains

The transcriptomes of pyrenoid-positive wild-type and pyrenoid-negative higher plant RBCS hybrid strains (spinach, Arabidopsis and sunflower) were recently sequenced to investigate the transcriptional response of cells to pyrenoid loss and to low CO_2 (Dr Moritz Meyer, unpublished data). To complement the proteomic analysis, strains for transcriptomic analysis were also adapted to low CO_2 for three hours before harvest. The initial analysis consisted of determining differences in gene expression levels between pyrenoid-positive (wild-type) and pyrenoid-negative (higher plant RBCS hybrid) strains adapted to low CO_2 for three hours. These results will be discussed solely as part of the analysis of LC-MS/MS data presented in this chapter rather than as a full description of the transcriptional response (for example, assigning biological process and molecular functions) of the high plant hybrid mutants to low CO_2.

5.2.3.1.1 Overview of RNA sequencing experiment

Overall, six pyrenoid-positive and six pyrenoid-negative samples were sequenced, consisting of two technical replicates and three biological replicates each. A total of 19,143 unique transcripts were detected compared to the nearly 1,400 proteins identified using LC-MS/MS and the 2,500 individual proteins resolved on the 2D gels. The fold-change between gene transcripts in the wild-type compared to higher plant RBCS hybrid strains was calculated for 16,469 transcripts and varied from 0.005 (log_2.-7.2) to 2000 (log_2.11) (Fig. 5.7). Approximately one fifth (21%) of the total transcripts detected showed a greater than two-fold difference in abundance between pyrenoid-positive and pyrenoid-negative strains. Statistical analysis of expression levels across the twelve samples identified 2,466 gene transcripts that were more abundant in the wild-type, 1,949 transcripts that were more abundant in the higher plant RBCS hybrid strains, 399 genes that were expressed equally across the strains and 14,329 genes that showed no significant pattern of expression. Overall, 1.3 times more gene transcripts were upregulated in the wild-type compared to 2.3 times more proteins, which may be related to differences in coverage as well as actual expression levels.

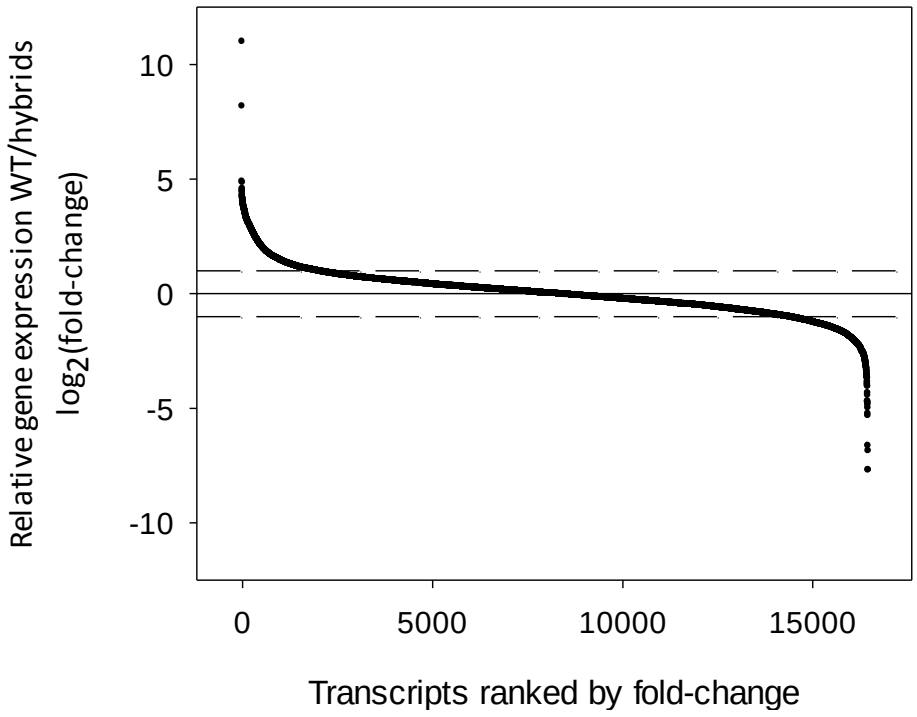


Figure 5.7 Abundance of gene transcripts in wild-type relative to the higher plant RBCS hybrid strains.

5.2.3.1.2 Expression of low CO_2-induced genes

Fifty-six transcripts annotated as low CO_2-inducible (LCI) were detected but only ten were identified as significantly more abundant in pyrenoid-positive (three genes) or pyrenoid-negative (seven genes) strains. *LCI2* and *LCI5* gene transcripts were both identified as being less abundant in wild-type cells whereas LCI5 was more abundant at the protein level and no difference was detected in the abundance of LCI2. *LCI14* was significantly more abundant in the wild-type but was not differentially expressed at the protein level. None of the other differentially expressed LCI gene transcripts overlapped with LC-MS/MS data.

However, significant patterns were assigned according to stringent criteria (>99% confidence) so the fold-change in gene expression in wild-type cells relative to the pyrenoid-less higher plant RBCS hybrids was also analysed to determine whether there was a general pattern in the expression of LCI genes. Fold-change in expression was determined for 55 CCM-related or low CO_2-induced transcripts (Fig. 5.8). In general, the differences in transcript abundance observed across these genes were small and only five gene transcripts were at least two-fold more abundant in the wild-type (*CAH4*, *LCI14*, *LCIA*, *LCIE*, and DNA mismatch repair protein/MutS homologue *MSH6*) while twelve gene transcripts were at least two-

fold more abundant in the higher plant RBCS hybrids (*LCI3*, *LCI5*, *LCI6*, *LCI11*, *LCI18*, *LCI24*, *LCI33*, *LCI34*, *LCI35*, stress-related chlorophyll a/b binding protein *LHCSR1*, low CO_2 and stress-induced one-helix protein *OHP2*, and serine hydroxymethyltransferase *SHMT1*). Transcripts encoding the regulatory factors, CIA5 and LCR1, were present in similar abundance in wild-type and higher plant RBCS hybrid cells.

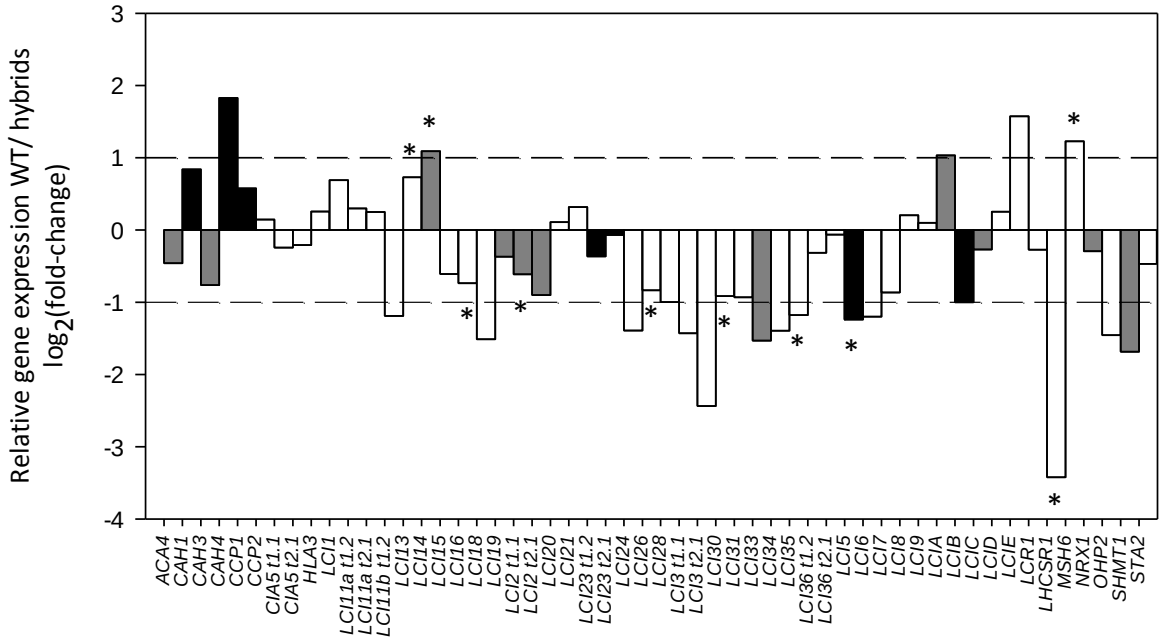


Figure 5.8 Abundance of CCM and low CO_2-induced gene transcripts in wild-type cells relative to higher plant RBCS hybrids. Positive values indicate transcripts that are more abundant in low CO_2-adapted wild-type. Statistically significant differences (>99% confidence) are indicated with an asterisk. Black bars indicate proteins identified by LC-MS/MS that were significantly more abundant in low CO_2-adapted wild-type cells.Grey bars indicate proteins that were identified by LC-MS/MS but that were not differentially expressed between wild-type and spinach RBCS hybrid strains.

5.2.3.1.3 Overlap between RNAseq and LC-MS/MS results

Except for chloroplast-encoded genes, the proteins identified using LC-MS/MS were mapped back to their respective gene transcripts. In total, 175 transcripts were identified that encoded 144 differentially expressed proteins and 1,350 transcripts were mapped to 1,218 non-differentially expressed proteins. The differential expression of only 27 transcripts (15% of the 175 transcripts mapped) was matched by differential expression at the protein level (Table 5.4). A similar proportion (20%) of the total transcripts detected were also differentially expressed, which suggests that differential expression at the transcript level was not strongly correlated with differential expression at the protein level (see Appendix 2 for the full comparison).

Table 5.4 Comparison of relative protein abundance (wild-type vs spinach RBCS hybrid) with relative mRNA abundance (wild-type vs higher plant RBCS hybrids). The number of transcripts in each category is listed as well as the percentage of the total number transcripts identified.

		Relative protein abundance		
		Up in WT	**Up in Sp**	**Non-DE**
Relative mRNA abundance	**Up in WT**	21 17%	6 12%	249 19%
	Up in hybrids	9 7%	7 14%	82 6%
	Equal/ no pattern	90 73%	30 59%	964 71%
	NA	4 3%	8 16%	47 4%

However, as discussed above, stringent criteria were used to assign patterns of differential expression to the transcriptome data so the relative abundance of transcripts encoding differentially and non-differentially expressed proteins was analysed in more detail (Fig. 5.9). A correlation between mRNA and protein abundance would be seen in greater relative mRNA abundance in the wild-type compared to spinach RBCS hybrid and non-differentially expressed samples; however, this was not observed. The median relative expression (WT/hybrids) was 1.06 for proteins that were more abundant in wild-type cells, 1.00 for proteins that were more abundant in spinach RBCS hybrid cells and 1.14 for proteins that were not identified as differentially expressed. This suggests that differential accumulation of proteins was not related to a general increase in the accumulation of gene transcripts.

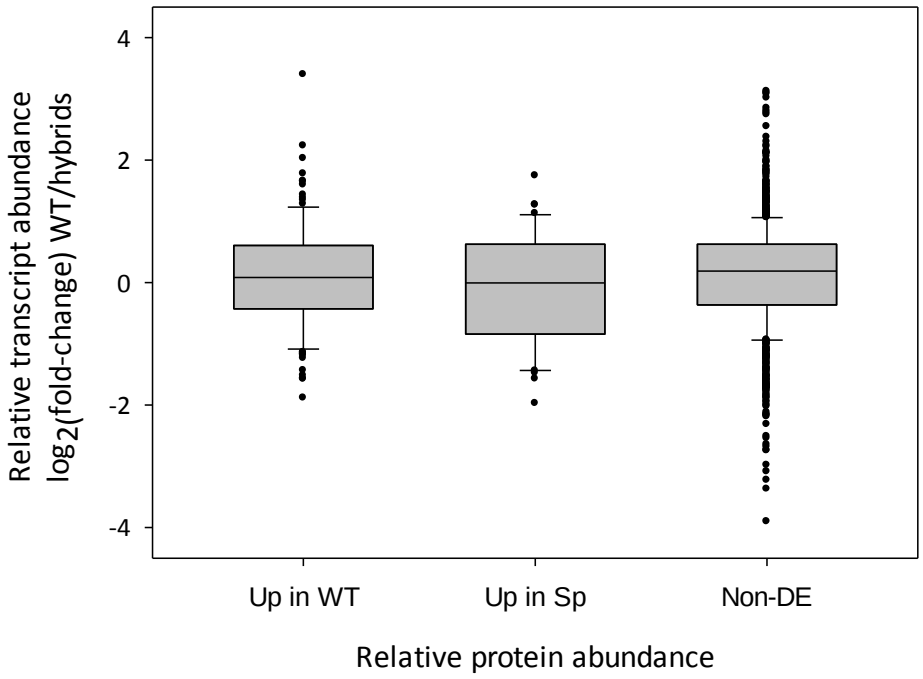


Figure 5.9 Relative abundance of transcripts (WT/hybrids) encoding proteins identified as more abundant in wild-type (Up in WT), more abundant in spinach RBCS hybrid (Up in Sp) or equally abundant (non-DE). No significant difference between the medians was found (Kruskal-Wallis ANOVA, p=0.547).

The relatively small overlap between differentially expressed proteins and differentially expressed gene transcripts may be due to the limitations of protein detection using LC-MS/MS or variability in the induction of the CCM in these strains. The overall correlation between relative protein abundance (mean spectral count) and relative mRNA abundance (mean reads per kilobase per million or RPKM) was very low (R^2=0.06) suggesting that protein accumulation in response to low CO_2 was mostly independent of changes in mRNA accumulation.

5.2.3.2 Low CO_2-induced gene expression in wild-type and *cia5* mutant strains

The 158 proteins identified as differentially expressed in the CCM-positive wild-type compared to the CCM-negative spinach RBCS hybrid strain were compared to gene transcripts that were differentially regulated in response to CO_2 in true wild-type and *cia5* mutant strains (Brueggeman et al., 2012; Fang et al., 2012) (Table 5.5; Appendix 2). Transcripts encoding 64 differentially expressed proteins were identified as CO_2-responsive by Brueggeman et al. (2012) (Supplemental datasets 2 and 3), the majority of which are downregulated in response to low CO_2 (52 gene transcripts). In addition, transcripts encoding approximately one quarter of the proteins (39) were identified as differentially expressed in a different wild-

type strain by Fang et al., (2012) (Supplemental dataset 1). The expression of 30 of these genes showed a significant response to CO_2 concentrations (C-effect) and expression of 31 genes was affected by the *cia5* mutation (strain/S-effect). Overall, nine proteins that were more abundant in the wild-type were also induced at the transcript level in response to low CO_2 in both transcriptome datasets. This included the CCM/low CO_2-induced proteins CAH1, CAH5, CCP1, LCI5, LCI23 and LCIB. The other proteins found in all three datasets were malate dehydrogenase (MDH2; A8ICG9), β-ureidopropionase (A8HPY4) and a predicted protein (A8IGV4). All three proteins were present in relatively low abundance (mean spectral counts of 1.0-1.3 in the wild-type and 0.0 in spinach RBCS hybrid). Malate dehydrogenase is involved in the TCA cycle, although this was the least abundant isozyme detected, suggesting that it may have an alternative role. β-ureidopropionase is involved in uracil degradation and the predicted protein (A8IGV4) has sequence similarity to an α-glucan water dikinase/pyruvate phosphate dikinase (Table 5.3).

As a comparison, 158 proteins that were not identified as differentially expressed in the LC-MS/MS dataset were randomly chosen and also compared to the transcriptome datasets (Table 5.5). Similar numbers of differentially expressed gene transcripts were found to overlap with this set of non-differentially expressed proteins as overlapped with the differentially expressed proteins (73 and 40 transcripts in the Brueggeman and Fang datasets, respectively). Overall, this suggests that CO_2-responsive genes were not preferentially associated with differential protein accumulation in wild-type and spinach RBCS hybrid cells.

Table 5.5 Comparison of proteins identified in wild-type versus spinach RBCS hybrid (LC-MS/MS) with genes differentially expressed in wild-type and *cia5* mutant strains in response to low CO_2 (Brueggeman et al., 2012; Fang et al., 2012). The number of proteins identified in two or more datasets is listed for 158 differentially expressed (Total DE, including proteins up in WT or up in Sp) or non-differentially expressed (non-DE) proteins. Differential expression of gene transcripts was significant if q-value<0.05 for either CO_2 concentration (C-effect) or for the *cia5* mutant compared to the wild-type (S-effect). Genes were either upregulated (positive C-effect) or downregulated (negative C-effect) in response to low CO_2.

	Total DE	Up in WT	Up in Sp	**Non-DE**
Bruegg. Total DE	64	50	14	73
C-effect, positive	12	10	2	9
C-effect, negative	52	40	12	64
Fang total DE	39	32	6	40
C-effect, positive	13	11	2	16
C-effect, negative	17	15	2	17
S-effect	31	24	4	32
Bruegg. and Fang C-effect	23	21	2	23
C-effect, positive	9	9	0	5
C-effect, negative	14	12	2	16

5.3 Discussion

5.3.1 Shotgun LC-MS/MS analysis as a tool to investigate CCM induction and metabolism

While LC-MS/MS is a very sensitive technique, the dynamic range is reduced for complex mixtures of proteins, especially those that are dominated by a few highly abundant proteins (Schulze and Usadel, 2010). For example, LC-MS/MS analysis of the *E. huxleyi* proteome, using the same method and instruments, yielded a total of 1,835 proteins (18 runs) compared to the 1,376 proteins (12 runs) identified in this study (McKew et al., 2013). This is likely due the observation that, in Chlamydomonas, the most abundant proteins accounted for even greater proportion of the proteome than in *E. huxleyi* (56% and 44% of the total spectral counts were associated with the top 50 proteins, respectively). These limits to detection distinguish proteomics approaches from microarray and RNA sequencing experiments, which are able to detect up to 16,000 individual gene transcripts in Chlamydomonas (Miura et al., 2004; Brueggeman et al., 2012; Fang et al., 2012).

The lower coverage of a proteomics experiment can be overcome by cellular fractionation and enrichment prior to mass spectrometric analysis and, in the

future, large-scale proteomics approaches will be required to complement the substantial microarray and transcriptome datasets that have been recently generated (Miura et al., 2004; Yamano et al., 2008; Brueggeman et al., 2012; Fang et al., 2012). Better quantification of protein abundance should enable a distinction to be made between what is primarily a transcriptional response and what is occurring at the functional (protein) level. This may narrow down the number of CCM candidates further than the current list of 1,141 published genes that are transcriptionally upregulated in response to low CO_2 (Miura et al., 2004; Wang et al., 2005; Yamano et al., 2008; Yamano and Fukuzawa, 2009; Brueggeman et al., 2012; Fang et al., 2012; see Supplemental Fig. 7 of Brueggeman et al., 2012). Indeed, the results of this LC-MS/MS experiment demonstrate that a genome-scale proteomics approach can complement existing studies, as well as build on current knowledge and understanding of CCM regulation.

5.3.2 The pyrenoid-less spinach RBCS hybrid mutant does not accumulate CCM proteins to wild-type levels

Eleven low CO_2-induced proteins were detected in this study, of which six failed to accumulate to wild-type levels in the spinach RBCS hybrid mutant. With the exception of LCI33, the other CCM-related proteins also appeared to be somewhat more abundant in the wild-type, although this was not supported by statistical tests. This is consistent with the qRT-PCR and immunoblot data presented in Chapter 4 which show lower than wild-type accumulation of CCM mRNA and proteins in all three higher plant RBCS hybrid strains. This reduced induction of CCM proteins could be responsible for the observed impairment of CCM activity in the spinach RBCS hybrid strain. However, the possibility cannot be ruled out that other proteins, also required for CCM activity, were not induced in the spinach RBCS hybrid mutant.

In addition, how loss of the pyrenoid leads to impaired CCM induction in the spinach RBCS hybrid strain remains an open question. Although qRT-PCR and RNA sequencing data are somewhat conflicting, reduced CCM protein accumulation could be due to reduced mRNA accumulation, in which case, the sensing and signalling pathways affecting the synthesis and stability of CCM gene transcripts will require further investigation. For example, while electron transport

appears necessary for carbon accumulation (Spalding and Ogren, 1982; Spalding et al., 1984), probably due to the demand for high energy compounds to energise Ci transport, there is some evidence to suggest that electron transport is also required for expression of CCM components. Inhibiting electron transport using DCMU blocks accumulation of CAH1 and mitochondrial CA mRNA and protein in cells switched from high to low CO_2 (Dionisio-Sese et al., 1990; Eriksson et al., 1998). In PSI and PSII mutants, some high light-induced CCM gene transcripts (*CAH1*, *CAH5* and *HLA3*) are not induced, suggesting that electron transport may also be required for expression (Im and Grossman, 2002). In a green alga of the same family as Chlamydomonas, *Dunaliella tertiolecta*, Rubisco aggregation in the pyrenoid was also shown to be affected by light intensity and was inhibited by DCMU (Lin and Carpenter, 1997b). This suggests that when CO_2 concentration drops, both wild-type and spinach RBCS hybrid strains may initially respond by inducing CCM-related genes. In the wild-type, this increases affinity for Ci and allows cells to continue to photosynthesise. In contrast, the spinach RBCS hybrid may be unable to maintain normal photosynthetic rates without pyrenoid-facilitated CCM activity. This might lead to the downregulation of light harvesting components and electron transport, which could then feed back to inhibit further CCM induction. Time courses investigating the response of electron transport and CO_2 fixation rates during acclimation to low CO_2 in wild-type and pyrenoid-less strains could help to untangle the relationship between these factors.

5.3.3 Pyrenoid loss leads to altered primary metabolism as well impaired CCM induction

The 158 differentially expressed proteins identified using LC-MS/MS accounted for approximately 37% of the total spectral counts and reflected the altered expression of a number of biological pathways that were not directly related to the CCM. Similar to the reduced accumulation of CCM proteins, enzymes involved in primary metabolism also tended to be present in lower abundance in spinach RBCS hybrid cells. In the interpretation of these results, the metabolism of the spinach RBCS hybrid strain is assumed to be equivalent to that of the wild-type at high CO_2 based on wild-type-like growth rates and gene expression under conditions when the CCM and presumably, the pyrenoid, are not required. If this is the case, spinach RBCS hybrid cells must either fail to induce or actively degrade these enzymes during adaptation to low CO_2 given the short (3 h) adaptation and

relatively long estimated half-lives of these proteins (for example: citrate synthase, A8J2S0, 35 h; 40S ribosomal protein S3a, A8HS48, 42 h; and peptidyl-prolyl *cis-trans* isomerase, A8IY43, 87 h; Supplemental Table 4 of Mastrobuoni et al., 2012).

In either case, low CO_2-adapted wild-type cells appear to be more metabolically active overall with a greater number of proteins upregulated in the TCA cycle, respiration, fatty acid, amino acid and protein synthesis pathways. This may be due to higher rates of carbon fixation in wild-type cells but it could also be due to the need for increased protein synthesis during CCM induction. Greater metabolic activity in wild-type cells would also be consistent with the inhibited growth of the spinach RBCS hybrid at low CO_2 (Chapter 4). If spinach RBCS hybrid cells are unable to adapt to low CO_2 and continue through the cell cycle, a general reduction in the abundance of primary metabolic enzymes after three hours at low CO_2 might be indicative of the arrested growth of this strain.

In true wild-type cells, a general decrease in transcription associated with these energy-requiring pathways is observed one hour after the switch to low CO_2 but transcript levels have generally recovered by three hours (Brueggeman et al., 2012). A transient decline such as this may not have had a large effect on overall protein abundance in the wild-type cells in this study. If, however, transcriptional recovery were dependent on activation of the CCM, transcript levels may have remained low in the spinach RBCS hybrid mutant, leading to a subsequent reduction in protein amount. Initial analysis indicates that photosynthesis and protein synthesis-related pathways may indeed be transcriptionally upregulated in the wild-type but more detailed analysis of biological pathways differentially regulated between the wild-type and higher plant RBCS hybrids is needed.

5.3.4 Are pyrenoid-less mutants also more sensitive to light?

Qualitative evidence suggests that growth of the low CO_2-adapted higher plant RBCS hybrid mutants is similar to wild-type at very low light intensities (10 µmol photons m^{-2} s^{-1}) but is inhibited at higher light intensities (> 50 µmol photons m^{-2} s^{-1}; Oliver Caspari, unpublished data). Low CO_2-adapted pyrenoid-less mutants also exhibit thylakoid hyperstacking and more plastoglobuli than pyrenoid-positive strains (Meyer et al., 2012). This suggests that, without a functioning

CCM, higher plant RBCS hybrid mutants are light-sensitive during growth at low CO_2.

Plastoglobuli are globular compartments containing lipids and proteins that are associated with stress and senescence in plastids and which, in chloroplasts, may work to protect the photosynthetic machinery from oxidative stress damage (Nacir and Brehelin, 2013). In an *Arabidopsis thaliana* mutant plant, impaired primary metabolism leads to thylakoid hyperstacking and increased numbers of plastogolubli which indicate both light and oxidative stress (Häusler et al., 2009). A similar phenotype was observed in a second light-sensitive Arabidopsis mutant lacking plastoglobule kinases (Lundquist et al., 2013). Thylakoid hyperstacking mimics adaptation to low light in higher plants and algae (for example, Berner et al., 1989; Kirchhoff, 2014) and may be associated with altered stoichiometry of the photosystems.

In addition to differences in growth and chloroplast ultrastructure associated with pyrenoid loss, reduced quantum yield and increased non-photochemical quenching are observed in low CO_2-adapted higher plant RBCS hybrid strains compared to the wild-type. The amount of non-photochemical quenching in pyrenoid-less mutants is dependent on both growth and experimental light intensities and can be returned to wild-type levels by high CO_2. Pyrenoid-less RBCS mutants also show reduced linear and cyclic electron flow when maintained at low CO_2, and the linear electron flow can be rapidly restored by the addition of high levels of inorganic carbon (Oliver Caspari, unpublished data). Overall, these gross phenotypic and photosynthetic differences in pyrenoid-negative compared to pyrenoid-positive strains suggest that loss of the pyrenoid and CCM results in cells suffering photodamage at low CO_2, leading to the activation of photoprotective mechanisms such as increased non-photochemical quenching and plastoglobuli production and reorganisation of the thylakoid membranes.

However, analysis of the proteomes of wild-type and spinach RBCS hybrid mutants after three hours at low CO_2 identified differential expression of only a small number of proteins that might be expected to be induced in response to light stress. Twelve molecular chaperones/stress-induced proteins (HSP/ClpP) and three isozymes of the thylakoid metalloprotease FtsH were detected but none of these was differentially expressed. Similarly, the stress-related chlorophyll binding

protein LHCSR3 was not differentially expressed between the spinach RBCS hybrid and wild-type strains even though *LHCSR3* transcripts are known to accumulate under high light stress (Maruyama et al., 2014). In *E. huxleyi*, other stress-related chlorophyll-binding LI818 proteins were more abundant in high-light adapted cells (McKew et al., 2013), which is consistent with the observation that five chlorophyll binding proteins were more abundant in the spinach RBCS hybrid strain.

In addition, proteins involved in protection from oxidative stress might be expected to be more abundant in light-sensitive spinach RBCS hybrid cells (for example, Eshdat et al., 1997; Asada, 2006); however, the three redox-related proteins (PRX2, peroxiredoxin, NTR1) that were differentially expressed were all more abundant in the wild-type. Other redoxin-domain containing proteins (Reiske ferredoxin, TRXH, TRXh2, TRXm, PRX5) were also detected in slightly higher abundance in the wild-type (spectral count ratio > 1.2) although these differences were not statistically significant. No differences were detected in the abundance of glutaredoxins (GRX1 and GRX2), glutathione peroxidases (GPX, gpxh), ascorbate peroxidases (A8JG56, apx1), monodehydroascorbate reductase (MDAR1) or superoxide dismutases (FSD1, MSD2, MSD1). Catalase (CAT) was significantly more abundant in the spinach RBCS hybrid. The differences in the expected abundances of these proteins may be due to a role for these proteins in signalling as well as detoxification (Michelet et al., 2006) or perhaps cells were harvested at the wrong time to capture the major changes in redox-related protein abundance.

Finally, few of the previously identified high-light induced genes were identified in the list of differentially expressed proteins. *HLA6*, encoding the serine hydroxymethyltransferase (SHMT1) was detected in a ratio of 1.6 (Sp/WT) but this difference was not statistically significant and the high light-induced nuclease (HLA8) was present in equal abundance (Im and Grossman, 2002). Only six genes from a list of 64 high light-induced genes were positively identified as differentially expressed at the protein level and these were split evenly between wild-type and spinach RBCS hybrid strains (Table 1 of Im et al., 2003). Comparison with the partial proteome of high light- and very high light-adapted cells did not identify significant overlap, with only CAH1 and CAH4 differentially expressed in both datasets (detected at high but not very high light) (Förster et al.,

2006). Lastly, while activities of the photorespiratory enzymes glycolate dehydrogenase and malate synthase are both highly responsive to high light stress (Davis et al., 2013), malate synthase (MS1) was more abundant in the wild-type and glycolate dehydrogenase (GDH1) was not differentially expressed.

In general, while there is physiological evidence to suggest that pyrenoid-less mutants are light-sensitive, evidence for this at the molecular level remains unconvincing. As mentioned above, differences in the timing of harvest and growth conditions could be responsible for the observed lack of induction of stress-related proteins in the spinach RBCS hybrid mutant. More detailed physiological time courses to optimise the timing of harvest for molecular studies might help to resolve these differences.

5.3.5 RBCS mutants as a system to investigate regulation of the CCM

The RBCS (higher plant RBCS hybrid) mutants represent a unique system to investigate CCM regulation because they are the only genetically characterised strains that accumulate normal levels of catalytically competent Rubisco while entirely lacking a pyrenoid. This makes them excellent candidates for studies of the mechanism of Rubisco aggregation as well as studies investigating the relationship between pyrenoid presence or absence and CCM induction and function.

5.3.5.1 Identification of novel CCM candidates

The reduced CCM induction at the protein level identified in this study indicates that the higher plant RBCS hybrid mutants could be used to identify new components of the CCM. Novel CCM candidates are likely to be those proteins that are present in greater abundance in low CO_2-adapted wild-type cells. Proteins encoded by genes previously identified as low CO_2-induced (LCI2, LCI5, LCI23 and LCI33) are obvious candidates.

LCI2 is perhaps the least promising candidate because transcripts are more abundant in the *cia5* mutant and this gene is not found to be CO_2-responsive in either of the latest transcriptome datasets (Brueggeman et al., 2012; Fang et al., 2012). Functional predictions implicate LCI2 in protein degradation or fatty acid biosynthesis, which suggest that it may only have a peripheral role in acclimation

to low CO_2. *LCI33* is a stronger candidate because it is upregulated in response to low CO_2 (Brueggeman et al., 2012) and structural and localisation predictions suggest it may be a transcription factor, although, unexpectedly, LCI33 was slightly (but not significantly) more abundant in spinach RBCS hybrid rather than wild-type cells. Expression of *LCI23*, encoding a septin-like protein, was strongly induced by low CO_2 in the wild-type but not the *cia5* mutant strain and annotation suggests it encodes a transmembrane protein (Fang et al., 2012). Localisation of LCI23 to the plasma membrane, chloroplast envelope or thylakoid membrane would give added support to a putative role in the CCM. Finally, LCI5 is also a strong candidate for involvement in the CCM. The transcript is highly abundant, moderately CO_2-responsive and somewhat repressed in the *cia5* mutant (Brueggeman et al., 2012; Fang et al., 2012). LCI5 is predicted to be targeted to the chloroplast and preliminary immunofluorescence data indicate the LCI5 is localised to the pyrenoid under low CO_2 (Dr Moritz Meyer, unpublished data). Localisation of LCI5 in high CO_2- and dark-adapted cells will also be necessary, particularly given that the low CO_2-dependent phosphorylation of LCI5 could alter its localisation or function during CCM induction (Turkina et al., 2006). For all of these candidates, generation of insertional mutants or amiRNA-mediated knockdown of gene expression could help determine the importance of these putative low CO_2-induced proteins for CCM activity.

In addition, the 28 proteins of unknown function that are more abundant in the wild-type may also include novel CCM components. Due to the large amount of work involved in manual validation, computational predictions of sequence similarity, structure and function were used as an initial screen for proteins of particular interest. While there have been recent advances in this field (reviewed by Nemoto et al., 2013; Radivojac et al., 2013), consistent predictions across the tools were made for less than half the unknown proteins and for several proteins there was still no annotation at all. Comparison of these proteins with the sequences of 1,000 plant transcriptomes (1KP project) may give additional clues into which ones are most likely to be CCM-specific.

5.4 Conclusions and future work

Analysis of the proteomes of low CO_2-adapted wild-type and spinach RBCS hybrid strains has provided further insight into CCM and metabolic regulation in a

pyrenoid-less mutant as well as demonstrating that a shotgun LC-MS/MS approach can be effectively adapted to study low CO_2 acclimation in Chlamydomonas. Despite differing from the wild-type in the sequence of only a single gene, the spinach RBCS hybrid mutant shows large differences in cellular ultrastructure (loss of pyrenoid) and gene expression. The spinach RBCS hybrid mutant fails to fully induce CCM components in response to low CO_2 and this, along with the complete lack of Rubisco aggregation, is likely to be responsible for the absence of any detectable increase in photosynthetic affinity for Ci in low CO_2-adapted cells. The apparent downregulation of primary metabolism also observed in the pyrenoid-negative and CCM-negative spinach RBCS hybrid mutant may simply be a downstream result of this impaired carbon fixation but further evidence is needed.

Follow up experiments should include an analysis of both strains at high CO_2 as well as analysis of membrane-enriched fractions for detection of Ci transporters. Specific analysis of the nuclear proteomes, via nucleus-tagging and pull-downs, will be useful for identifying signalling cascades involved in CCM gene expression. Careful analysis of post-translational modification of CCM components may uncover factors determining the final localisation, activity and abundance of CCM proteins.

6 GENERAL DISCUSSION

The overarching aim of this thesis was to investigate the regulation of the carbon-concentrating mechanism in both wild-type and mutant Chlamydomonas strains. While advances have been made in the identification of genes/proteins required for adaptation to low CO_2, relatively little is known about the mechanisms and pathways governing non-transcriptional responses to low CO_2. Even less is known of the integration of other signals, for example light, the cell cycle and cellular energy and nutrient status, into overall CCM expression. A better understanding of the relative functions and interactions of individual components, as well signals and mechanisms governing their activity, will lead to a more complete picture of the CCM in this model eukaryote.

Eukaryotic CCMs are estimated to be responsible for approximately 25% of global net primary production with prokaryotic CCMs accounting for a further 25% (Field et al., 1998; Falkowski and Raven, 2007). Algal CCMs are thus vitally important for world-wide carbon fixation, helping to overcome the kinetic constraints of the primary photosynthetic carboxylase, Rubisco, and enabling growth in environments where external concentrations of Ci are limiting. An understanding of algal growth in different environments is important for predicting potential shifts in phytoplankton communities due to future ocean acidification (Raven et al., 2005; Shi et al., 2010; Lohbeck et al., 2014).

In addition, the identification of factors limiting algal productivity is significant for attempts to generate sustainable biofuels as well as on a fundamental level. Furthermore, the gains in photosynthetic efficiency and primary productivity

associated with an active CCM are increasingly the target of collaborations aimed at reengineering C3 photosynthesis in terrestrial plants, which relies solely on the diffusive supply of CO_2. Hence, improved knowledge of the Chlamydomonas CCM will not only advance our understanding of this important aspect of basic photosynthetic biology, it will also facilitate current and future attempts to introduce elements of the algal CCM into higher plants.

Firstly, in order to identify potential non-transcriptional levels of CCM regulation, in-depth time course experiments were used to investigate the extent of CCM induction in true wild-type cells under contrasting conditions of light and CO_2 concentration (Chapter 3; Mitchell et al., 2014). A comparison was made between the relatively poorly understood induction of the CCM during the dark-to-light transition in synchronised cells grown continuously at low CO_2 and the better-defined induction of the CCM in asynchronous cultures adapting to low CO_2. Detailed investigation of changes at the whole cell and molecular level during the first six hours of adaptation to low CO_2 expanded on the results of previous studies, even in this relatively well characterised system. This study indicated that de novo induction of the CCM in asynchronous cultures is a coordinated response that relies on the transcriptional response of numerous low CO_2-induced genes, which in turn drives protein accumulation and increased whole cell affinity for Ci. In contrast, CCM activity in synchronous cultures was uncoupled from mRNA and protein accumulation. CCM activity was repressed two hours before dawn even though CCM proteins were present in relatively high abundance. The CCM was then fully inducible one hour before dawn, without concomitant increases in CCM proteins, but coincident with preferential localisation of Rubisco and CAH3 to the pyrenoid. This relocalisation of Rubisco and CAH3 prior to dawn suggested that an endogenous signal such as the circadian clock might be at work, possibly by directing the post-translational modification and mobilisation/targeting of these proteins. This work also highlighted the potential role of protein aggregation in the pyrenoid in a fully functioning CCM.

Hence, the aim of the second study, described in Chapter 4, was to investigate further the importance of aggregation of Rubisco in the pyrenoid to the expression of the CCM using detailed physiological and molecular characterisation of pyrenoid-negative higher plant RBCS hybrid mutants. Low CO_2-adapted mutants

did not show increased affinity for Ci compared to high CO_2-adapted cells and appeared to have differential accumulation of low CO_2-responsive mRNA and proteins encoded by CO_2-responsive genes. Accumulation of Ci transporter mRNA was reduced in the pyrenoid-negative higher plant RBCS hybrids while several proteins that are transcriptionally downregulated in wild-type cells in response to low CO_2 were more abundant in low CO_2-adapted spinach RBCS hybrid cells. This suggested that the dispersal of Rubisco throughout the stroma was not solely responsible for the CCM-negative phenotype of RBCS mutants and that the pyrenoid itself might affect CCM induction.

To investigate this hypothesis in more detail, differential expression of proteins in low CO_2-adapted wild-type and spinach RBCS hybrid strains was further investigated at the genome scale, as described in Chapter 5. Approximately 10% of all the proteins detected were differentially expressed between the strains and functional analyses identified greater accumulation of CCM proteins and primary metabolic enzymes in wild-type cells. However, the differential expression of these proteins was not correlated with the differential expression of gene transcripts in the wild-type compared to the higher plant RBCS hybrids or with previously identified CO_2-responsive genes in true wild-type strains. Overall, this indicated that while loss of the pyrenoid was accompanied by changes in both CCM expression and cellular metabolism, this was not likely to be due to a specific impairment in the perception of, or response to, low CO_2. A model of CCM regulation was thus proposed that included: cross-talk between Rubisco aggregation/carbon fixation; light harvesting reactions/electron transport; and cellular metabolism and CCM expression.

This final chapter consists of a discussion of the results of the above studies in the broader context of CCM regulation, that is, in light of what is known or hypothesised about the signalling and other pathways that modulate CCM activity in response to a range of endogenous and exogenous cues (as summarised in Fig. 6.1). In particular, the response of cells to the external stimuli of CO_2 concentrations and light as well as the internal cues of the circadian clock will be discussed. The role of Rubisco aggregation/pyrenoid state will be discussed as it relates to CCM activity and regulation as well as in the context of current work towards engineering a pyrenoid-based carbon-concentrating mechanism into

higher plants (Meyer and Griffiths, 2013). The need for integration of these different internal and external signals will then be discussed in the context of transcriptional regulation and the possible existence of multiple *cis*-regulatory elements in the promoters of CCM-related genes. Finally, possible mechanisms of pyrenoid biogenesis, protein relocalisation and post-transcriptional or post-translational regulation of gene expression and protein function will be evaluated. Where pertinent, examples of the above processes occurring in other organisms, especially the better characterised cyanobacterial system, will be briefly presented but this is not intended to be an exhaustive survey of the literature. Possible follow-up experiments will be identified as and when they arise, preceding a later section which will give a fuller account of current and future studies aimed at answering the key questions raised in this discussion. The chapter will conclude with a brief account of how advances described in this dissertation, as well as other recent advances in the field, may open up new possibilities and avenues of research.

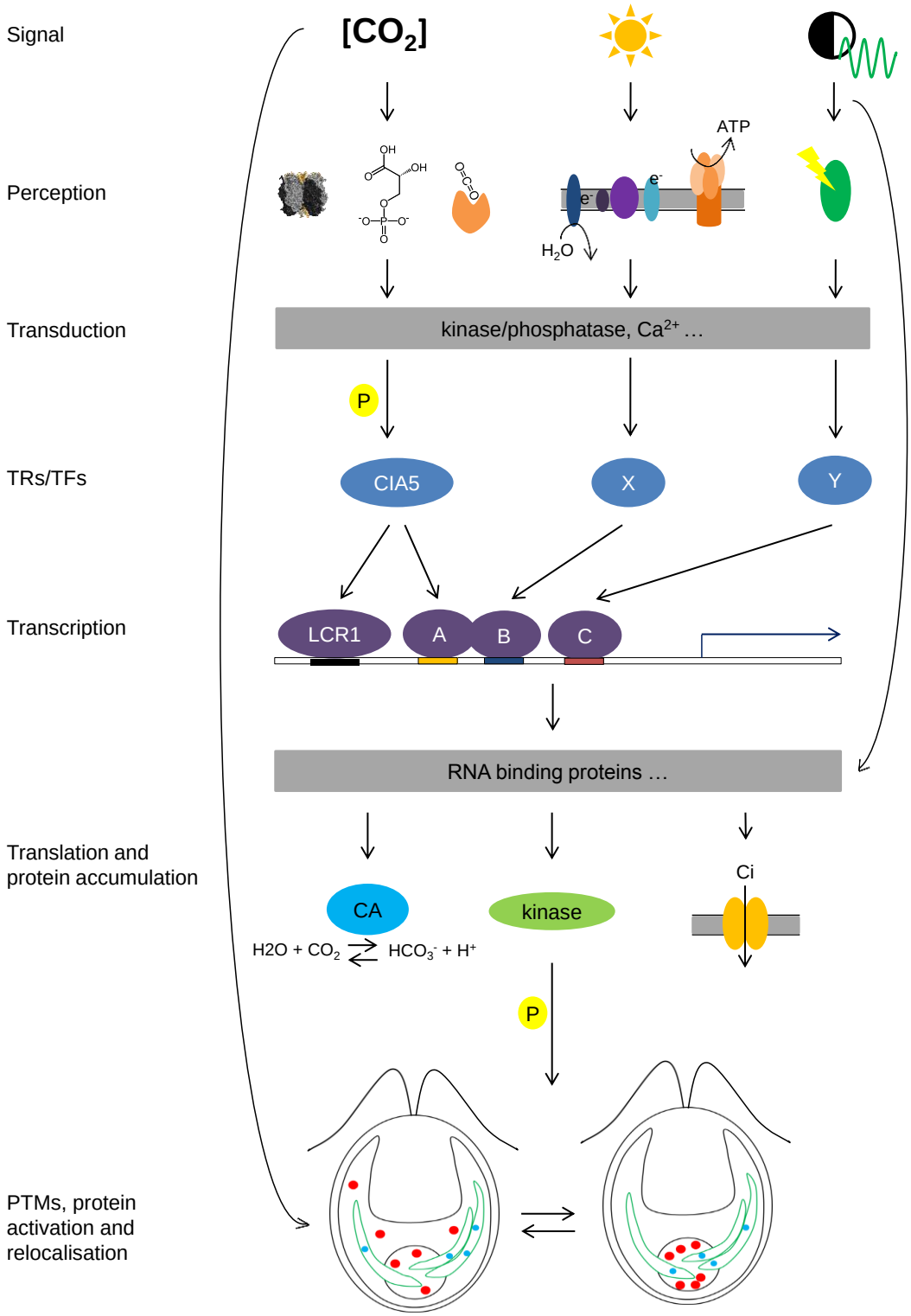


Figure 6.1 Summary of general understanding and hypotheses regarding CCM regulation in *Chlamydomonas reinhardtii*. Cues affecting CCM induction include CO_2 concentration, light and circadian or other endogenous signals. Perception of CO_2 concentration may be via metabolic enzymes, metabolites or direct sensing; light availability is likely to be indirectly detected via the redox state of the photosynthetic electron transport chain and phytochromes may be involved in the circadian sensing of light/dark cycles. CCM gene expression is controlled by at least two tiers of transcription regulators/factors including CIA5 and LCR1. Activation and relocalisation of proteins may occur through allosteric interactions or post-translational modifications mediated by existing or newly synthesised proteins. Little is known about the mechanism of signal transduction or specific factors affecting CCM protein accumulation (grey boxes).

6.1 CCM activity is modulated by endogenous and exogenous signals

CCM activity must be regulated in response to cellular energy and nutrient status, which in turn depend on external factors such as light, Ci and mineral availability. Three major drivers of CCM expression and their putative signalling pathways are summarised in Fig. 6.1. This section will first focus on the response of cells to CO_2 levels and light as these are the two major drivers of CCM expression. Second, the potential role of the circadian clock in controlling CCM activity in synchronised cells will be discussed. Although this element of CCM regulation has been largely ignored in the context of laboratory-based experiments, it is likely to be of high biological relevance. Finally, evidence will be presented for the role of the pyrenoid both as a necessary component of a high affinity CCM and as an essential element governing expression of other CCM components.

6.1.1 CO_2 concentration

Time course experiments investigating acclimation of wild-type cells to low CO_2 showed a rapid, coordinated response, likely to be at least partly driven at the transcriptional level (Chapter 3; Mitchell et al., 2014). All nine CO_2-responsive genes measured in this study are under the control of CIA5 (Fang et al., 2012), which appears to be the primary coordinator of altered CCM and metabolic gene expression. However, while CIA5-dependent induction of the CCM appears to be specifically linked to low CO_2 concentrations (Fukuzawa et al., 2001), the mechanism by which this signal is detected and transduced to the nucleus is unknown. Proposed signals include the concentration of either total inorganic carbon or dissolved CO_2 or the ratio of CO_2 to O_2, which may be sensed via internal Ci pools, photorespiratory or CBB cycle intermediates or the redox state of the electron transport chain (reviewed by Kaplan and Reinhold, 1999).

In the freshwater cyanobacterium *Synechococcus elongatus* PCC 7492, induction of a high affinity CCM depends mainly on the size of the internal inorganic carbon pool (Woodger et al., 2005). Additionally, although an as yet unidentified oxygen-dependent process is required for maximum CCM expression and activity, the evidence does not support a role for photorespiratory activity in CCM induction (Woodger et al., 2005). A study using another freshwater cyanobacterium

Synechocystis sp. PCC 6803 has identified binding of metabolites by a transcriptional regulator (Daley et al., 2012), thereby providing a mechanism linking the internal Ci pool with expression of CCM-related genes. In these cells, high Ci concentrations tend to lead to high levels of α-ketoglutarate and $NADP^+$ which bind to the LysR-type transcriptional regulator CcmR and maintain the active repression of two inducible Ci transporters, NDH-I_3 and SbtA. In contrast, low Ci concentrations lead to reduced carboxylation and increased oxygenation by Rubisco as well as an accumulation of RuBP and 2-phosphoglycolate. Binding of these metabolites activates a different LysR-type regulator, CmpR, which then activates the expression of a third Ci transporter, the ABC-type bicarbonate transporter BCT1. The activation of BCT1 by CmpR has been found in both *Synechocystis* and *Synechococcus* species (Nishimura et al., 2008; Daley et al., 2012).

However, in eukaryotic algae, the mechanism is more complicated. While most CCM genes appear to be nucleus-encoded and so are under the control of nuclear-localised transcription factors, alterations in metabolite concentrations will likely be sensed in other cellular compartments. Changes in metabolite concentrations might affect signal transduction pathways and gene expression via the post-translational modification of proteins or by these metabolites binding and altering the activity of enzymes. The observation that some enzymes (unrelated to the CCM) also have RNA binding properties raises the question whether similar mechanisms might link the metabolic and regulatory processes involved in CO_2 sensing and CCM expression (Bohne et al., 2013; see also reviews by Bhardwaj and Wilkinson, 2005; Ciesla, 2006).

In the best understood example of CO_2-responsive gene regulation in a eukaryotic algal system, cAMP-responsive promoter elements are involved in the repression of chloroplastic carbonic anhydrase gene *ptca1* under high CO_2/cAMP in the marine diatom *Phaeodactylum tricornutum* (Harada et al., 2006; Ohno et al., 2012), but the exact mechanism of signal transduction remains unknown. In contrast, chemical modification of cell surface proteins in Chlamydomonas inhibits the accumulation of high CO_2 periplasmic marker protein H43 (Hanawa et al., 2007), which suggests that a putative CO_2 receptor exists in that space while the mechanisms of gene regulation are unknown. It would be interesting to observe

the effects of these cell surface chemical modifiers on the expression of known low CO_2-induced genes to determine whether the same sensor is involved in both the high and low CO_2 responses. To add further difficulty to the identification of a putative CO_2 sensor in Chlamydomonas, no CO_2-binding receptor proteins have yet been identified in higher plants; however, studies of Drosophila and mice suggest that CO_2 sensing may occur via G-protein coupled-receptors or carbonic anhydrases (Kim et al., 2010). Future work to elucidate the pathways of CO_2-sensing and CIA5-mediated induction of the CCM should include identification of the interaction partners of CIA5 via yeast-two-hybrid assays or nuclear pull-down assays from high and low CO_2-adapted cells.

6.1.2 Light in asynchronous and synchronous cultures

Light is an important regulator of CCM activity, for example, by driving photosynthesis to produce the high energy compounds that are presumably used to power Ci uptake and accumulation (Spalding and Ogren, 1982; Spalding et al., 1984). Light also appears to be required for expression of CCM genes (Dionisio-Sese et al., 1990; Kucho et al., 1999; Im et al., 2003) and can affect the localisation of key proteins (Yamano et al., 2010; Mitchell et al., 2014; Tirumani et al., 2014). Effects of light on CCM activity are observed in asynchronous as well as synchronous cultures, indicating that light can influence CCM expression independently of the role it plays in entraining the circadian clock. This section will cover the influence of irradiance itself on CCM expression while the specific role of the circadian clock will be discussed in the following section.

In synchronised Chlamydomonas, light was shown to have a strong effect on the transcriptional regulation of CCM-related genes (Mitchell et al., 2014; Tirumani et al., 2014). In addition, a strong transcriptional response to light pre-treatment before dawn was observed for most genes and the Ci transporters in particular (Chapter 3; Mitchell et al., 2014). Despite the relatively high abundance of CCM proteins just prior to dawn, the degree of upregulation of the CO_2-responsive genes was similar to that observed during the de novo induction of the CCM in response to low CO_2. This suggests that similar processes may regulate both transcriptional responses and that protein accumulation is not solely dependent on mRNA abundance. Indeed, while Tirumani et al., (2014) showed increased accumulation of *CAH3*, transcripts in the dark, CAH3 protein did not peak until the

light period, which also indicates a likely role for post-transcriptional regulation of CCM components.

A similar light-dependence of CCM expression has been identified in cyanobacteria. In *Synechocystis* sp. PCC 6803, microarray experiments show that genes encoding proteins involved in inducible Ci uptake are strongly induced by light and low CO_2 but only minimally (and possibly more transiently) induced by light in the presence of high CO_2 (Burnap et al., 2013). In *Synechococcus elongatus*, light is also required for CCM gene induction although very low light levels, that is, levels that do not lead to a detectable increase in electron transport/oxygen evolution, are sufficient (McGinn et al., 2005). Finally, high irradiance increases the expression of CCM genes under limiting Ci (Woodger et al., 2003; McGinn et al., 2004), although this could be due to temporary depletion of Ci from increased photosynthesis, as has been suggested to be the case for high-light induction of CCM genes in Chlamydomonas (Im and Grossman, 2002).

A likely candidate for linking light intensity and light-related signal transduction with CCM expression is the redox state of the chloroplast. Redox regulation of metabolism is particularly important in photosynthetic organisms and comprises the electron transport chain, ferredoxin-NADP oxidoreductase, thioredoxins and reactive oxygen species (Scheibe and Dietz, 2012). Inhibitors of linear electron transport also inhibit the induction of two CO_2-responsive carbonic anhydrases in asynchronous Chlamydomonas (Dionisio et al., 1989b; Eriksson et al., 1998), although the signalling pathway behind this is not known. The gene encoding thioredoxin f1, which is involved in CP12-mediated regulation of the CBB cycle (Marri et al., 2014), is also CO_2-responsive and has been postulated to have a regulatory role in adaptation to low CO_2 (Miura et al., 2004), although there is no direct evidence for redox-related modifications affecting the activity of CCM components in Chlamydomonas. In contrast, in *P. tricornutum*, activity of a pyrenoidal carbonic anhydrase is regulated by thioredoxin, thereby linking the photosynthetic electron transport chain with the pyrenoid and CCM (Kikutani et al., 2012). Finally, light quality (blue light, specifically) has also been shown to affect the expression of CAH1 (Dionisio et al., 1989a), which suggests that specific photoreceptors must also be involved in acclimation to low CO_2, although the effect of different wavelengths of light was not investigated in this study.

6.1.3 Circadian clock or other endogenous signals

CAH1 is the only CO_2-responsive gene thus far to be shown to be under circadian control (Fujiwara et al., 1996) and, unlike the majority of CCM-related genes, did not show a strong transcriptional response to light exposure before dawn. In addition, relocalisation of Rubisco and CAH3 to the pyrenoid one hour before dawn and without changes in the external environment suggested that these CCM components might also be influenced by circadian or other endogenous signals (Chapter 3; Mitchell et al., 2014). The oscillation of a number of other CCM gene transcripts during dark/light cycles suggests a more general influence of the circadian clock on transcription of CCM-related genes (Tirumani et al., 2014), although this remains to be properly tested, for example, by switching synchronised cells to continuous light or continuous dark over extended subjective day/night periods.

Given the interdependence of the CCM and photosynthesis as well as natural growth conditions (alternating periods of light and dark), circadian regulation of the CCM would help to time maximum CCM expression with periods of maximum photosynthetic activity. In this respect, relocalisation of key proteins prior to dawn appears to almost anticipate the light period and thereby maximise the length of time for efficient carbon fixation. That the oscillation in $K_{0.5}$ observed by Marcus et al. (1986) in synchronised Chlamydomonas was correlated with changes in the phosphorylation state of the thylakoid membranes and with maximum photosynthetic rates supports this idea. Similarly, in the dinoflagellate *Gonyaulax*, Rubisco localisation is under circadian control and Rubisco aggregation into pyrenoids during the day coincides with periods of peak carbon fixation (Nassoury et al., 2001).

In *Synechocystis*, microarray experiments identified circadian oscillations in numerous CCM-related genes including those involved in CO_2 uptake, bicarbonate uptake, carboxysome formation, genes of the ABC-type bicarbonate transporter *cmp* operon and a gene encoding a *cmp* operon regulator (Layana and Diambra, 2011). Similarly, in *Synechococcus*, genes in the carboxysome shell *ccmKLMNO-rbcLS* cluster show rhythmic expression, peaking at subjective dawn in wild-type cells (Ito et al., 2009). These experiments provide strong evidence that expression

of the CCM is under circadian control in cyanobacteria but evidence for global circadian regulation of the CCM in Chlamydomonas is still lacking.

6.1.4 The pyrenoid as both functional component and regulator of the CCM

Physiological and molecular characterisation of higher plant RBCS hybrid mutants showed that strains that were unable to form a pyrenoid also had very low photosynthetic affinity for Ci and a high CO_2-requiring phenotype (Chapter 4). While the pyrenoid itself is likely to play an important role in concentrating CO_2 around Rubisco, proteomic studies indicated that pyrenoid loss was also associated with reduced accumulation of other CCM components (Chapter 5).

In the high CO_2-requiring *cia6* mutant, the normally conspicuous single pyrenoid is fragmented into small, irregular-shaped pyrenoids, containing only 35% of the cell's Rubisco and this mutation is associated with low affinity for Ci (Ma et al., 2011). The mutation occurs in a gene encoding a SET domain methyltransferase, suggesting that abnormal pyrenoid formation may be the result of incorrect post-translational modifications; however, no methyltransferase activity was shown in vitro. Under high CO_2, the affinity of the *cia6* mutant for Ci was around two times lower than the respective wild-type and, similarly, our study found that pyrenoid-less spinach RBCS hybrid cells had a three times lower affinity for Ci than wild-type cells. Furthermore, this reduced affinity for Ci was even more pronounced in low CO_2-adapted *cia6* and spinach RBCS hybrid cells (Ma et al., 2011; Chapter 4). These data suggest that the presence of a single, spheroidal pyrenoid may increase photosynthetic affinity for Ci at high as well as low CO_2, although loss of the pyrenoid only appears to affect growth at low CO_2.

Another similarity between *cia6* and true pyrenoid-less mutants is the reduced accumulation of certain CCM mRNA and protein (CAH4, CCP1 and LCIB) in response to low CO_2 (Ma et al., 2011). This is consistent with qRT-PCR and LC-MS/MS data obtained for the low CO_2-adapted pyrenoid-less spinach RBCS hybrid mutant identifying weaker induction of Ci transporter gene transcripts and CCM-related proteins (Chapters 4 and 5). Together, these results imply that some form of regulatory feedback exists between pyrenoid formation and CCM induction.

Cross talk between Rubisco packaging – perhaps also in proximity to a carbonic anhydrase – and CCM expression also appears to occur in cyanobacteria, despite the lack of molecular homology between carboxysomes and pyrenoids. For example, carboxysome-less *ΔccmM* mutants of two different genera (*Synechocystis* and *Synechococcus*) fail to accumulate CCM-related gene transcripts to wild-type levels in response to low CO_2 (Woodger et al., 2005; Hackenberg et al., 2012). In *Synechococcus*, the CCM is induced primarily in response to a decreased internal Ci pool so the higher internal Ci accumulation in the *ΔccmM* mutant is likely to be responsible for the partial suppression of Ci transporter genes (Woodger et al., 2005). On the other hand, in *Synechocystis*, the binding of photosynthetic and photorespiratory metabolites to CCM-related transcription factors (Nishimura et al., 2008; Daley et al., 2012) provides a different means of sensing and responding to changes in carbon metabolism, including changes in Rubisco packaging.

Further work is required to clarify whether similar mechanisms are occurring in Chlamydomonas pyrenoid mutants. The *cia6* mutant, for example, has a reduced ability to accumulate Ci suggesting that a high internal Ci pool is not responsible for the observed suppression of low CO_2-inducible genes (Ma et al., 2011). Whether altered Rubisco packaging interferes with other sensing or signalling pathways or whether loss of carbon fixation inhibits CCM gene expression via downregulation of the electron transport chain remains to be determined. Future studies should include: measurements of internal Ci accumulation in pyrenoid-less higher plant RBCS hybrid mutants; determination of metabolite pools in the *cia6* and pyrenoid-less higher plant RBCS hybrid mutants; and the detailed time courses investigating the interplay of carbon fixation, light harvesting and CCM induction suggested in Chapter 5. The development or identification of mutants with alterable levels of Rubisco accumulation/packaging would also enable more detailed investigation of possible cross talk between the pyrenoid and CCM. Further discussion of pyrenoid biogenesis and the correct localisation of Rubisco to the pyrenoid will follow in section 6.3.1.

6.1.4.1 Implications for the engineering of a pyrenoid-based CCM in higher plant chloroplasts

Several collaborations are underway to re-engineer photosynthesis by introducing carbon concentrating mechanisms from higher plants and algae into terrestrial C3

plants. For example, efforts are being made to transform rice into a C4 crop (von Caemmerer et al., 2012, C4 rice project, http://c4rice.irri.org/) as well as to engineer carboxysomes and cyanobacterial Ci transporters into higher plants (Price et al., 2012, RIPE project). Work in our group, in collaboration with the John Innes Centre (Norwich), the University of Edinburgh and the Carnegie Institution for Science (Stanford), is focussed on engineering elements of the Chlamydomonas CCM, including the pyrenoid, into higher plants (Meyer and Griffiths, 2013, CAPP project, http://cambridgecapp.wordpress.com/). Modelling of both pyrenoid- and carboxysome-based CCMs indicates that aggregation of Rubisco in these structures is an important element of the native CCM as well as intracellular Ci accumulation in higher plants (Badger et al., 1998; Price et al., 2011; Price et al., 2012; McGrath and Long, 2014). Work on the pyrenoid-negative RBCS mutants has provided further experimental evidence of the essential role of the pyrenoid in CCM activity (Chapters 4 and 5; Meyer et al., 2012).

Following on from this, site-directed mutagenesis experiments are underway to identify the specific Rubisco amino acid residues required for pyrenoid formation. Previous work in our group identified two surface-exposed RBCS α-helices as necessary for pyrenoid formation (Meyer et al., 2012). Substitution of the amino acids that differ between pyrenoid-positive algae and pyrenoid-negative higher plants should determine the minimal sequence requirements for Rubisco aggregation in Chlamydomonas. Given the potential for alterations to RBCS primary structure to alter Rubisco holoenzyme kinetics (Spreitzer, 2003), as well as the relatively poor kinetics observed in mutants with only the RBCS α-helices substituted (Meyer et al., 2012), this highly targeted approach is worth pursuing in parallel with the experiments using native Chlamydomonas RBCS sequences described below.

In the meantime, experiments with Arabidopsis are also being undertaken to test the feasibility of hybrid Rubisco holoenzyme and pyrenoid assembly (Dr Alistair McCormick, University of Edinburgh). Stable as well as transient protoplast expression of Chlamydomonas *RBCS1* and *RBCS2* genes in an Arabidopsis *RBCS* knockdown mutant, expressing only 15% of wild-type Rubisco levels (Izumi et al., 2012) will first test whether the Chlamydomonas RBCS is able to assemble with the Arabidopsis rbcL into a functional holoenzyme. Secondly, these

experiments will test whether aggregation of Rubisco occurs via Chlamydomonas RBCSs, without the presence of other Chlamydomonas-specific factors, and whether any Rubisco-Rubisco interactions are sufficient to improve the photosynthetic characteristics of these transgenic plants. Encouragingly, the transient expression of β-carboxysomes in *Nicotiana* has demonstrated the feasibility of assembling novel highly organised protein structures in higher plant chloroplasts (Lin et al., 2014). Finally, using synthetic approaches rather than biological mimicry to re-engineer C3 photosynthesis might also be successful, for example, novel molecular scaffolds may be introduced to tether Rubisco to carbonic anhydrases (MAGIC project, http://magic.psrg.org.uk/).

6.2 Integration of signals via multiple transcription/regulatory factors

The nuclear-localised protein CIA5 appears to control the induction of a large number of genes required for adaptation to low CO_2, including carbonic anhydrases and Ci transporters, but the mechanism of CIA5-related gene transcription is unknown (Fig. 6.1). While substitution of a single histidine residue (H54) in the first putative zinc finger motif abolishes CIA5 activity (Xiang et al., 2001), there is no experimental evidence suggesting that CIA5 directly binds DNA. CIA5 does, however, appear to act upstream of the CO_2-responsive transcription factor LCR1, which in turn regulates the expression of at least three low CO_2-induced genes and has been shown to bind *cis*-regulatory elements in the *CAH1* promoter (Yoshioka et al., 2004). By combining these empirical data with a FAIREseq experiment, Winck et al. (2013) developed a model of low CO_2-induced gene regulation and identified a complex network of regulatory factors with LCR1 and LCR2 (a second Myb-domain transcription factor) regulating the transcription of a total of five additional genes. The existence of such a network may allow the rapid activation of large numbers of genes as well as the integration of multiple signals through the action of different regulatory factors. Coordinated regulation of CCM genes might also be facilitated by the clustering of several CCM-related genes into a 100 kb region of the genome as well as the head-to-head confirmation of *LCID*/*CCP2* and *LCIE*/*CCP1* gene pairs (Supplemental Fig. S20 of Merchant et al., 2007; Supplemental Fig. S6 of Brueggeman et al., 2012). Nevertheless, coordination of gene regulation in the prokaryotic system is even

simpler with most of the CCM-related genes organised into several operons that can then be transcribed as one. Another advantage of clustering CCM genes in this way is that it appears to be associated with lower protein variability and horizontal gene transfer (Shi and Falkowski, 2008), which might be important to maintaining functional interactions between carboxysome shell proteins or the subunits of multimeric Ci transporters.

6.2.1 *Cis*-regulatory elements in CO_2- and light-responsive CCM gene promoters

The presence of binding sites for several transcription factors in CCM gene promoters, each activated by a different signal or pathway, may facilitate the regulation of gene expression in a changing environment. For example, *CAH1* expression appears to be controlled by at least three sets of regulatory factors: first, proteins that bind four enhancer element consensus motifs (EECs) in the upstream region (-651 to -231 bp 5' of the start codon) and that are present independent of CO_2 concentration (Kucho et al., 2003); second, *LCR1*, which is only present under low CO_2 and is required for amplification and maintenance of expression (Yoshioka et al., 2004); and third, as yet unidentified factors belonging to the circadian clock.

Of nine putative *cis*-regulatory elements or transcription factor binding sites identified in the upstream sequence of the *CAH1* gene (Brueggeman et al., 2012, Kucho et al., 1999, Kucho et al., 2003), two are present in most of the CO_2-responsive genes investigated in Chapter 3. At least one copy of the CO_2-responsive motif CGCGCC (Kucho et al., 1999) and the LCR1 binding motif GANTTNC are present in the 5' region of *LCI1*, *CCP1*, *CCP2*, *HLA3*, *LCR1*, *LCIB* and *RBCS2* as well as 5' of *CAH1*. *LCIC* also has one copy of CGCGCC while *RBCS1* and *RBCS2* have 4 to 5 copies of GANTTNC. A 39-mer repeat of the putative transcription factor binding motif AGGGGCAGGGG is present in both *CAH1* and *LCI1* upstream sequences, which may be involved the strong low CO_2-inducibility of these genes. However, the motif was only identified computationally by Brueggeman et al. (2012) and the relevant transcription factor remains unknown. The partial overlap of putative *cis*-regulatory elements present in *CAH1* and other CCM-related gene promoters may account for similarities in the CO_2 response while the regulatory elements unique to *CAH1* may be recognised by

factors involved in different signalling pathways, for example, the circadian clock. Work is in progress to identify *cis*-regulatory elements present in the promoters of highly CO_2-responsive genes and bioinformatic analyses are also underway to identify other putative regulatory elements associated with the transcriptional response to low CO_2 (see also section 6.4.2).

6.2.2 Evidence for a multi-level response

Expression patterns of genes in response to light and CO_2 concentrations sometimes differ according to growth conditions, for example, whether expression is measured in synchronised or unsynchronised cells. These effects have been identified in this as well as previous studies and may be indicative of a complex, possibly multi-tiered, effect of different regulatory cues. For example, exposure to light prior to dawn is sufficient to drive high levels of Ci transporter gene expression but does not affect the expression of *CAH1* (Chapter 3; Mitchell et al., 2014), suggesting that the circadian control of *CAH1* expression in synchronised cells is stronger than any specific response to light. Synthesis of *CAH1* transcripts also begins in the dark, just prior to dawn, which lends further support to the relative importance of the circadian rather than the light response in synchronised cells (Rawat and Moroney, 1995). In asynchronous cells, while light itself is an important signal, it is not an absolute requirement for *CAH1* induction. In asynchronous cells switched to low CO_2 in the dark, carbonic anhydrase activity is induced but to a lesser extent and with a lag of several hours compared to cells switched to low CO_2 in the light (Bozzo and Colman, 2000), suggesting that light speeds up the response to low CO_2 as well as increasing the magnitude of the eventual response. Finally, while cells adapting to low CO_2 in the dark also appear to induce active Ci uptake, this induction occurs before that of carbonic anhydrase activity (Bozzo and Colman, 2000), indicating the possible existence of at least two signalling pathways that respond to the same (low CO_2) signal.

Two distinct regulatory mechanisms have been observed during the sulphur deprivation response in Chlamydomonas. When sulphur levels are low, induction of one set of genes (SO_4^{2-} transporters) can occur without protein synthesis whereas induction of the second set (SO_4^{2-} assimilation and cellular reorganisation) is blocked by the addition of the cytoplasmic ribosome inhibitor cycloheximide (Aksoy et al., 2013). Elements in the first tier of the sulphur

deprivation response appear to be already present under sulphur-replete conditions and are rapidly activated via protein modification or ligand binding when external sulphur concentrations drop whereas other unidentified elements involved in second tier gene regulation are presumably synthesised de novo.

It is tempting to speculate that similar tiers of regulation may operate in cells responding to low CO_2. As a general stress response, cells deprived of sulphur or CO_2 exhibit downregulation of cell division, metabolism and photosynthesis, at least transiently, followed by the specific induction of the relevant nutrient uptake and assimilation systems. The rapid relocalisation of CAH3 in response to light before dawn suggests that regulation occurs by way of existing protein modifiers rather than transcription. In contrast, Rubisco localisation does not change in response to the same pre-dawn light treatment, indicating that relocalisation is driven by other cues, possibly circadian-controlled transcriptional processes (Chapter 3; Mitchell et al., 2014). Employing fast and slow, or general and specific, regulatory mechanisms may better enable cells to cope with unexpected changes or fluctuations in the environment, for example, by balancing the dual roles of CAH3 in PSII stability and CCM function, as well as preparing cells for diurnal changes affecting photosynthetic capacity.

6.3 Post-transcriptional and post-translational regulation

Post-transcriptional processes such as factors affecting mRNA stability and translation are important and relatively well characterised for the chloroplast-encoded photosynthetic genes whereas understanding of translational control in the cytoplasm remains limited (reviewed in Wobbe et al., 2008). These processes are important for the coordinated expression of nuclear and organellar genes as well as the assembly and maintenance of multi-subunit complexes. Regulatory mechanisms at the post-transcriptional level might facilitate CCM regulation over different time scales and with greater flexibility than a purely transcriptional response; however, specific examples of post-transcriptional regulation of CCM activity in Chlamydomonas are limited.

Thus far, a functional effect of post-translational modification of a CCM protein has been determined only for CAH3, although post-translational activation of CIA5 is likely. The mobility of proteins such as LCIB and LCIC may also be controlled at

the post-translational level with several modifications identified, including phosphorylation (Yamano et al., 2010), S-nitrosylation (Morisse et al., 2014) and glutathionylation (Zaffagnini et al., 2012a). Future experiments using inhibitors of protein modification as well as immunoprecipitation of mobile CCM components combined with mass spectrometry could identify potentially important functional modifications. Complementation of the *CAH3* and *LCIB* mutants using native and modified genes could also identify the specific amino acid residues required for low CO_2-induced phosphorylation, activation and relocalisation.

6.3.1 Pyrenoid biogenesis and protein localisation

Several key questions remain regarding the molecular composition and biogenesis of the pyrenoid, some of which may be resolved by studies of post-transcriptional regulation. The speed of Rubisco and CAH3 relocalisation implicates post-translational modification of the proteins themselves or perhaps the modification of other molecular linkers or chaperones (Fig. 6.1). The presence of mobile and fixed fractions of these proteins is a possibility yet to be explored; nucleation of the pyrenoid around a central thylakoid 'knot' might account both for correct positioning in the chloroplast and for a fraction of Rubisco that remains inaccessible to modifying enzymes. However, there is as yet no experimental evidence supporting these hypotheses.

Genome-wide analysis of the proteomes of low CO_2-adapted wild-type and spinach RBCS hybrid strains did not find strong evidence of a linker protein and or post-translational modifications. A protein with a role in tethering Rubisco holoenzymes would presumably exist in relatively high abundance, that is, in roughly stoichiometric amounts to Rubisco. In low CO_2-adapted *Synechococcus*, Rubisco subunits and Rubisco-binding/carboxysome shell proteins (CcmM) are present in the stoichiometric ratio of 8.0 (rbcL) to 5.3 (RbcS) to 3.2 (M58, CcmM long form) to 4.5 (M35, CcmM short form) (Long et al., 2011). In wild-type cells, mean spectral counts for Rubisco were 80.7 for rbcL and 30.7 for RBCS. Only one predicted protein (A8HPF4) was detected with a mean spectral count greater than 24 (approximately 80% of RBCS spectral count). This protein was not differentially expressed between wild-type and spinach RBCS hybrid strains and is predicted to be involved in anion transmembrane transport in the mitochondrial outer membrane (UniProt, Argot[2] and Phyre[2]). The other proteins detected in this

abundance are unlikely candidate linker proteins because annotations suggest roles in protein synthesis (ribosomes), metabolism or photosynthesis, although dual roles are possible even for well characterised proteins.

The small, unstructured, low CO_2-induced protein, LCI5, is perhaps the most likely candidate for a protein involved in Rubisco aggregation detected using an LC-MS/MS approach. LCI5 has been preliminarily localised to the pyrenoid under low CO_2, had a mean spectral count of 7.7 in the wild-type and was more abundant in these cells than in the spinach RBCS hybrid. Further investigation of LCI5 localisation and function will be needed to confirm this hypothesis. More targeted proteomic approaches will also be required to identify low CO_2-induced post-translational modifications of Rubisco as well as any putative molecular chaperones of pyrenoid assembly. Experimental confirmation of CIA6 methyltransferase activity would also make this protein a strong candidate for involvement in post-translational modification-mediated Rubisco aggregation.

Nevertheless, the self-assembly of pyrenoidal proteins, perhaps facilitated by post-translational modification, remains an intriguing possibility. In *P. tricornutum*, two stromal β-carbonic anhydrases are de-repressed at low CO_2 and form clusters in vivo and in vitro mediated by amphipathic C-terminal helices. Although the identity of other interacting partners in vivo is not known, the redox-regulation, self-assembly and likely pyrenoidal localisation of these carbonic anhydrases neatly link pyrenoid formation and CCM function with irradiance (reviewed in Matsuda et al., 2011). The evolutionary distance between α- and β-carbonic anhydrases, as well as the thylakoid lumenal localisation of α-CAH3 compared to the stromal localisation of β-PtCA1 and β-PtCA2, do not allow a direct comparison to be drawn between carbonic anhydrases and pyrenoid formation. Nevertheless, the ability of CCM proteins to form self-organising structures in response to changes in the environment is worth further investigation. For example, interaction of the peri-pyrenoidal Ci recapture proteins, LCIB and LCIC, with CAH6 or other pyrenoidal proteins could be tested using in vivo and in vitro methods.

Finally, molecular factors involved in pyrenoid biogenesis and protein relocalisation might include components of a plastoskeleton, although evidence for the general existence of plastoskeletons in photosynthetic eukaryotes is lacking (reviewed in Osteryoung and Pyke, 2014). Limited evidence in moss suggests that

the protein FtsZ forms plastoskeleton-like filaments in chloroplasts (Kiessling et al., 2000; Martin et al., 2009). Tubule-like structures have been observed in green algal chloroplasts but their molecular composition is not known (reviewed in Reski, 2009). Chlamydomonas has two *ftsZ* genes (Wang et al., 2003) and, while no FtsZ proteins were identified in the LC-MS/MS dataset, both *FTSZ1* and *FTSZ2* gene transcripts were significantly upregulated in the pyrenoid-positive wild-type compared to pyrenoid-negative higher plant RBCS hybrid cells. Expression of *ftsZ* genes is also circadian, peaking during the dark period (Hu et al., 2008), which would be consistent with a role in mobilisation of Rubisco at the end of the dark period.

6.3.2 A role for protein phosphorylation in CCM regulation

There is mounting evidence for the importance of kinase/phosphatase systems in at least three different subcellular compartments during CCM induction (Fig. 6.1). First, in the nucleus, changes to the phosphorylation state of CIA5 may alter its activity. CIA5 has fourteen predicted phosphorylation sites based on the primary amino acid sequence (Fukuzawa et al., 2001) and four phosphorylated residues have been identified with high confidence in mixotrophically-grown cells (Wang et al., 2014a). Notably, a truncated *CIA5* gene is able to rescue the growth of the *cia5* mutant but the protein, which is missing the C-terminal domain, is no longer CO_2-responsive and the rescued mutant constitutively expresses low CO_2-induced genes (Xiang et al., 2001). This suggests that post-translational modification of the C-terminal sequence of CIA5 might be required to switch off activity under high CO_2; however, only one of the experimentally identified phosphorylated residues occurs within this domain.

Second, in the stroma, phosphorylation of the chloroplastic (possibly pyrenoidal) protein LCI5 occurs in a redox- and low CO_2-dependent manner (Turkina et al., 2006) and up to 16 phosphorylation sites have been identified on this protein (Wang et al., 2014a). This suggests that post-translational modification may alter LCI5 activity during induction of the CCM. Preliminary immunofluorescence localisation to the pyrenoid under low CO_2 (Dr Moritz Meyer, unpublished data), as well as increased accumulation of LCI5 in the pyrenoid-positive wild-type compared to the pyrenoid-negative spinach RBCS hybrid (Chapter 5), indicate a potential role in pyrenoid formation, possibly mediated by phosphorylation. Third,

in the thylakoid lumen, the low CO_2-induced phosphorylation of CAH3 coincides with increased carbonic anhydrase activity and relocalisation to the pyrenoid (Blanco-Rivero et al., 2012). Thus, identification of low CO_2-induced kinases (or phosphatases) in any of these subcellular compartments might help confirm the role of phosphorylation in regulation of CIA5, LCI5 and CAH3 while identification of additional kinase targets might also uncover novel CCM components.

Indeed, phosphorylation of CCM-related proteins appears to be relatively common. In the most recent analysis of the phosphoproteome of mixotrophic Chlamydomonas cells, 16 low CO_2-induced proteins were identified as having at least one high confidence phosphorylation site, namely, CAH6, CAH8, CCM1, HLA3, LCI2, LCI5, LCI9, LCI15, LCI23, LCI24, LCI33, LCI34, LCIC, LCID, OHP1 and a predicted protein (A8J0B3) (Supplemental Table S1 of Wang et al., 2014a). RBCS1/2 and rbcL were also phosphorylated on residues Ser98 and Ser107 and on Thr471, respectively. Based on the crystal structure of Chlamydomonas Rubisco (Taylor et al., 2001), one of the phosphorylation sites on RBCS1 appears not to be solvent exposed in the holoenzyme suggesting that modification of this residue may occur prior to holoenzyme assembly.

These results differ from the LC-MS/MS study described in Chapter 5 in which no phosphopeptides were detected with high confidence. However, the experiment was not specifically designed to capture post-translational modifications and protein coverage was also lower. In this study, 12 LC-MS/MS runs on whole cell extracts identified 1,376 proteins, whereas Wang et al., (2014a) performed 180 LC-MS/MS runs on fractions enriched for phosphopeptides and detected 6,136 proteins. Phosphoproteins were not overrepresented amongst differentially expressed proteins in wild-type and spinach RBCS hybrid strains (42.4% compared to 43.5% of non-differentially expressed proteins.

Phosphorylation of proteins also appears to be important in cyanobacteria, with 4 % of phosphoproteins identified in *Synechococcus* sp. PCC 7002 involved in photosynthesis (Yang et al., 2013). However, of the CCM proteins, only the carboxysomal facet protein ccmK4 was identified as having a low confidence phosphorylation site. The other CCM proteins involved in Ci transport (ndhB, sbtA, bicA) and the carboxysome shell (ccmK1, ccmK2, ccmK3, ccmL, ccmM, ccmN) were detected but were not found to be phosphorylated, although this could be

due to the growth conditions (1% CO_2) used. Hence, while activation of Ci transport is proposed to occur by redox regulation and/or phosphorylation, experimental evidence of this is still limited (Price, 2011). Investigation of CCM induction, protein activity and localisation in the presence of kinase inhibitors may further elucidate the role of protein phosphorylation in these processes.

6.3.3 Other post-translational modification of CCM components

While phosphorylation is a key signalling process that appears to have a role in CCM induction in Chlamydomonas, other post-translational modifications of CCM proteins are likely. CIA5, for example, has an apparent molecular mass of 100 kDa but a predicted molecular mass of 70 kDa (Wang et al., 2005), indicating the possible presence of constitutive post-translational modifications. CIA5 also has putative glycosylation and myristoylation sites as well as phosphorylation sites (Fukuzawa et al., 2001). S-nitrosylation of proteins functions in signalling mechanisms and both LCIC and LCIE are S-nitrosylated under conditions of nitrosative stress (Morisse et al., 2014). LCIB is glutathionylated in mixotrophic cells and glutathionylation is implicated in redox regulation and oxidative stress conditions, including the regulation of CBB enzymes (Zaffagnini et al., 2012a, 2012b). Such redox-sensitive protein modifications would enable coordinated regulation of the CCM and carbon fixation in accordance with light and oxidative stress conditions. The reversible nature of most of these modifications (glycosylation, phosphorylation, S-nitrosylation, glutathionylation) may enable rapid changes to be made in protein activity.

6.4 Future directions

During the production of this dissertation, several lines of research were developed that were either unsuccessful or not able to be brought to conclusion. The following section summarises the progress made to date and suggests future directions that this research could take.

6.4.1 Providing a molecular definition of the pyrenoid

Recent technical advances have allowed the detailed analysis of subcellular and organellar proteomes (for example, Terashima et al., 2011; Winck et al., 2012) but even the most recent analysis of isolated pyrenoids is now decades old (del Río et

al., 1996). One of the original aims of this study was to understand the molecular composition of the pyrenoid by isolating intact pyrenoids for mass spectrometric analysis. Attempts were made to replicate the original density centrifugation protocol for pyrenoid isolation Kuchitsu et al. (1988), but this was difficult due to the fact that pyrenoids do not appear to be membrane-bound or to have an ordered protein shell like carboxysomes and the interactions holding the proteins together are unknown. So two alternative approaches were designed to isolate Rubisco plus any interacting proteins.

Site-directed mutagenesis was used to introduce a 6xHistidine tag at the C-terminus of either the Chlamydomonas or spinach (negative control) *RBCS* gene and the constructs were successfully used to rescue a *ΔRBCS* mutant. The transformants, despite being selected for based on restored autotrophic growth, failed to accumulate Rubisco to detectable levels. Analysis of gene expression revealed that *His-RBCS* mRNA was present in very low abundance whereas *rbcL* expression was similar to the wild-type.

Rather than introduce tags, which might hinder Rubisco holoenzyme assembly, an anti-Rubisco antibody column was also trialled to isolate native Rubisco. Rubisco was successfully isolated using this approach and was detected in multiple bands using immunoblots. Further work will be needed to compare wild-type, pyrenoid negative and *ΔRBCS* mutants to determine whether this technique is useful for isolating Rubisco-protein complexes. Use of membrane permeable (reversible) chemical crosslinkers prior to disruption of cells, would be an additional way of capturing Rubisco-protein interactions.

6.4.2 Identifying *cis*-regulatory elements in CCM promoters

While the transcriptional response to low CO_2 is relatively well characterised, the *cis*- and *trans*-factors involved in CCM-related gene expression are largely unknown. Work is in progress to identify *cis*-regulatory elements present in the promoters of the highly CO_2-responsive genes, *LCI1* and *LCIA*, using 5' nested deletion series fused to a reporter gene. Analysis of the mRNA levels of the reporter gene under different conditions of light and CO_2 concentration should enable identification of regions of the promoter that contain sequences recognised by one or more transcription factors. Five nested truncated regions of the *LCIA*

promoter (from -1,500 to -100) were fused to a green fluorescent protein (*GFP*) gene optimised for expression in Chlamydomonas (Fuhrmann et al., 1999). Ten transformants for each of the five constructs were screened for insertion and expression of the reporter gene. Genomic PCR yielded inconclusive results and RT-PCR and qRT-PCR identified only very low levels of *GFP* mRNA. Future work should focus on screening additional transformants for reporter gene expression as well as on generating transformants using the *LCI1* constructs. Use of an alternative reporter gene might also facilitate screening by PCR-based methods, for example, the arylsulfatase (*ARS*) gene has been successfully used in 5' nested deletion studies in Chlamydomonas (Kucho et al., 1999).

The low-throughput, targeted approach described above can be complemented by large-scale bioinformatic analyses. A wealth of data now exist on (CCM) gene regulation from the two recent transcriptomic studies (Brueggeman et al., 2012; Fang et al., 2012) as well as predictions of genome-wide gene regulatory networks (Winck et al., 2013) and predicted *cis*-regulatory elements (Ding et al., 2012). A pilot study was undertaken with Billy Aldridge (Part III Systems Biology, University of Cambridge) with the aim of identifying putative *cis*-regulatory elements in the 5' region of CO_2-responsive genes using de novo motif prediction programs. Several motifs were identified that were enriched in the CO_2-responsive gene datasets and a more comprehensive analysis might yield further candidates (Aldridge, 2014). Future work should include mutation of these putative transcription factor binding sites combined with reporter gene assays to determine the functional importance of these DNA motifs to low CO_2-responsive gene expression. Gel-shift assays could then be used to identify the factors that bind these regulatory elements and protein microarrays might also be developed to determine the protein interacting partners of CIA5, LCR1 and other regulatory factors (Gong et al., 2008).

6.4.3 Functional characterisation of CCM components

Careful physiological and biochemical analyses are still required to elucidate the exact roles of a number of known CCM components as well as any novel candidates identified in this and other studies. Knockdown of gene expression using artificial microRNAs (Molnar et al., 2009) was used as a more targeted approach to assess gene function than the generation of gene knockout strains through random insertional mutagenesis. Mutants were generated that

accumulated less than 10% of wild-type mRNA levels of putative Ci transporters *HLA3* or *LCIA*. No strong growth or CCM phenotype was observed in these mutants but further characterisation is required. *HLA3* knockdown mutants appeared to have a slight growth defect at high pH and low CO_2, similar to previously published *HLA3* mutants (Duanmu et al., 2009). The presence of multiple, redundant Ci transport systems might explain the lack of growth defect observed in artificial microRNA mutants. Determination of the expression levels of other Ci transporter genes might indicate whether increased expression of these components is able to compensate for loss of a single transporter. Use of artificial microRNAs will also be useful to determine whether novel CCM candidates do indeed play a role in CCM activity and growth at low CO_2.

A parallel screen is also underway to assess the phenotype of insertional mutants from the group of Dr Martin Jonikas (Carnegie Institution for Science; Zhang et al., 2014). Candidate genes were selected from a pilot study covering approximately 10% of the Chlamydomonas genome and include: putative chloroplast-localised protein (Ser/Thr) kinases, Rubisco chaperone protein RBCX, CO_2-responsive transcription factors/regulators identified by (Winck et al., 2013) and putative transcription factor, LCI33.

6.5 Conclusions

Unlike previous studies aimed at identifying transcriptional responses to low CO_2 and novel CCM components via mutagenic screens, the experiments described in this dissertation have instead focused mainly on non-transcriptional mechanisms of regulation and the physiological and molecular effects of pyrenoid-loss. Relocalisation of pyrenoid proteins prior to dawn as well as the lack of correlation between mRNA and protein abundance in pyrenoid-positive and pyrenoid-negative cells highlight the potential importance of post-transcriptional and post-translational mechanisms of regulating gene expression and protein function. Future studies of the CCM will need to include the identification of potential regulators of transcription and protein activity. In particular, a better understanding of post-translational regulation of CCM protein activity, especially of the Ci transporters, will be required for the successful engineering of an algal CCM into higher plant chloroplasts.

The search for novel CCM candidates will continue and will be greatly facilitated by the results from the first genome-wide insertional mutagenesis screen as well as continued improvement in genome assembly and annotation. Analysis of the 1KP project transcriptome data should enable a better understanding of the diversity and evolution of CCM-related genes while advances in de novo *cis*-regulatory motif prediction may help identify sequences and *trans* factors common to low CO_2-responsive genes in Chlamydomonas. Finally, as more and more details of CCM activity, function and regulation are clarified, computational modelling of the Chlamydomonas CCM will be useful to clarify the relative importance and interaction of components as well as generating testable hypotheses for future experimental validation. In short, the forthcoming availability of new data and tools brings the promise of answering longstanding questions in the field.

7 REFERENCES

Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image processing with ImageJ. Biophotonics International **11:** 36-42

Aizawa K, Miyachi S (1986) Carbonic anhydrase and CO_2 concentrating mechanisms in microalgae and cyanobacteria. FEMS Microbiology Letters **39:** 215-233

Aksoy M, Pootakham W, Pollock SV, Moseley JL, Gonzalez-Ballester D, Grossman AR (2013) Tiered regulation of sulfur deprivation responses in *Chlamydomonas reinhardtii* and identification of an associated regulatory factor. Plant Physiology **162:** 195-211

Aldridge B (2014) Parallel experimental and bioinformatic searches for DNA motifs associated with low CO_2 induced genes in *Chlamydomonas reinhardtii*. University of Cambridge, Cambridge

Arvidsson S, Kwasniewski M, Riaño-Pachón DM, Mueller-Roeber B (2008) QuantPrime - a flexible tool for reliable high-throughput primer design for quantitative PCR. BMC Bioinformatics **9:** 465

Asada K (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. Plant Physiology **141:** 391-396

Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. Canadian Journal of Botany **76:** 1052-1071

Badger MR, Kaplan A, Berry JA (1980) Internal inorganic carbon pool of *Chlamydomonas reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. Plant Physiology **66:** 407-413

Badger MR, Price GD (1994) The role of carbonic anhydrase in photosynthesis. Annual Review of Plant Physiology and Plant Molecular Biology **45:** 369-392

Bailly J, Coleman JR (1988) Effect of CO_2 concentration on protein biosynthesis and carbonic anhydrase expression in *Chlamydomonas reinhardtii*. Plant Physiology **87:** 833-840

Beardall J, Giordano M (2002) Ecological implications of microalgal and cyanobacterial CO_2 concentrating mechanisms, and their regulation. Functional Plant Biology **29:** 335-347

Berner T, Dubinsky Z, Wyman K, Falkowski P (1989) Photoacclimation and the "package" effect in *Dunaliella tertiolecta* (Chlorophyceae). Journal of Phycology **25:** 70-78

Bhardwaj A, Wilkinson MF (2005) A metabolic enzyme doing double duty as a transcription factor. BioEssays **27:** 467-471

Bienvenut WV, Espagne C, Martinez A, Majeran W, Valot B, Zivy M, Vallon O, Adam Z, Meinnel T, Giglione C (2011) Dynamics of post-translational modifications and protein stability in the stroma of *Chlamydomonas reinhardtii* chloroplasts. Proteomics **11:** 1734-1750

Blanco-Rivero A, Shutova T, Roman MJ, Villarejo A, Martinez F (2012) Phosphorylation controls the localization and activation of the lumenal carbonic anhydrase in *Chlamydomonas reinhardtii*. PLoS One **7:** e49063

Bohne AV, Schwarz C, Schottkowski M, Lidschreiber M, Piotrowski M, Zerges W, Nickelsen J (2013) Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biology **11:** e1001482

Borkhsenious ON, Mason CB, Moroney JV (1998) The intracellular localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. Plant Physiology **116:** 1585-1591

Bozzo GG, Colman B (2000) The induction of inorganic carbon transport and external carbonic anhydrase in *Chlamydomonas reinhardtii* is regulated by external CO_2 concentration. Plant Cell and Environment **23:** 1137-1144

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry **72:** 248-254

Brueggeman AJ, Gangadharaiah DS, Cserhati MF, Casero D, Weeks DP, Ladunga I (2012) Activation of the carbon concentrating mechanism by CO_2 deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii*. Plant Cell **24:** 1860-1875

Burnap RL, Nambudiri R, Holland S (2013) Regulation of the carbon-concentrating mechanism in the cyanobacterium *Synechocystis* sp. PCC6803 in response to changing light intensity and inorganic carbon availability. Photosynthesis Research **118:** 115-124

Burow MD, Chen ZY, Mouton TM, Moroney JV (1996) Isolation of cDNA clones of genes induced upon transfer of *Chlamydomonas reinhardtii* cells to low CO_2. Plant Molecular Biology **31:** 443-448

Chang IF, Chen PJ, Shen CH, Hsieh TJ, Hsu YW, Huang BL, Kuo CI, Chen YT, Chu HA, Yeh KW, Huang LC (2010) Proteomic profiling of proteins associated with the rejuvenation of *Sequoia sempervirens* (D. Don) Endl. Proteome Science **8:** 64

Chen ZX, Green D, Westhoff C, Spreitzer RJ (1990) Nuclear mutation restores the reduced CO_2/O_2 specificity of ribulosebisphosphate carboxylase/oxygenase in a temperature-conditional chloroplast mutant of *Chlamydomonas reinhardtii*. Archives of Biochemistry and Biophysics **283:** 60-67

Chen ZY, Lavigne LL, Mason CB, Moroney JV (1997) Cloning and overexpression of two cDNAs encoding the low-CO_2-inducible chloroplast envelope protein LIP-36 from *Chlamydomonas reinhardtii*. Plant Physiology **114:** 265-273

Ciesla J (2006) Metabolic enzymes that bind RNA: yet another level of cellular regulatory network? Acta Biochimica Polonica **53:** 11-32

Coleman JR, Grossman AR (1984) Biosynthesis of carbonic anhydrase in *Chlamydomonas reinhardtii* during adaptation to low CO_2. Proceedings of the National Academy of Sciences of the United States of America **81:** 6049-6053

Colombo SL, Pollock SV, Eger KA, Godfrey AC, Adams JE, Mason CB, Moroney JV (2002) Use of the bleomycin resistance gene to generate tagged insertional mutants of *Chlamydomonas reinhardtii* that require elevated CO_2 for optimal growth. Functional Plant Biology **29:** 231-241

Daley SM, Kappell AD, Carrick MJ, Burnap RL (2012) Regulation of the cyanobacterial CO_2-concentrating mechanism involves internal sensing of $NADP^+$ and α-ketogutarate levels by transcription factor CcmR. PLoS One **7:** e41286

Davis MC, Fiehn O, Durnford DG (2013) Metabolic acclimation to excess light intensity in *Chlamydomonas reinhardtii*. Plant Cell and Environment **36:** 1391 – 1405

del Río MJ, García-Reina G, Ramazanov Z (1996) The ultrastructure and polypeptide composition of the pyrenoid from *Dunaliella tertiolecta*. Scientia Marina **60:** 155-160

Ding J, Li X, Hu H (2012) Systematic prediction of *cis*-regulatory elements in the *Chlamydomonas reinhardtii* genome using comparative genomics. Plant Physiology **160:** 613-623

Dionisio-Sese ML, Fukuzawa H, Miyachi S (1990) Light-induced carbonic anhydrase expression in *Chlamydomonas reinhardtii*. Plant Physiology **94:** 1103-1110

Dionisio ML, Tsuzuki M, Miyachi S (1989a) Blue light induction of carbonic anhydrase activity in *Chlamydomonas reinhardtii*. Plant and Cell Physiology **30:** 215-219

Dionisio ML, Tsuzuki M, Miyachi S (1989b) Light requirement for carbonic anhydrase induction in *Chlamydomonas reinhardtii*. Plant and Cell Physiology **30:** 207-213

Duanmu D, Miller AR, Horken KM, Weeks DP, Spalding MH (2009) Knockdown of limiting-CO_2-induced gene *HLA3* decreases HCO_3^- transport and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences of the United States of America **106:** 5990-5995

Duanmu D, Spalding MH (2011) Insertional suppressors of *Chlamydomonas reinhardtii* that restore growth of air-dier *lcib* mutants in low CO_2. Photosynthesis Research **109:** 123-132

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

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

Eriksson M, Karlsson J, Ramazanov Z, Gardestrom P, Samuelsson G (1996) Discovery of an algal mitochondrial carbonic anhydrase: molecular cloning and characterization of a low-CO_2-induced polypeptide in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences of the United States of America **93:** 12031-12034

Eriksson M, Villand P, Gardestrom P, Samuelsson G (1998) Induction and regulation of expression of a low-CO_2-induced mitochondrial carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Physiology **116:** 637-641

Eshdat Y, Holland D, Faltin Z, BenHayyim G (1997) Plant glutathione peroxidases. Physiologia Plantarum **100:** 234-240

Espie GS, Kimber MS (2011) Carboxysomes: cyanobacterial RubisCO comes in small packages. Photosynthesis Research **109:** 7-20

Falda M, Toppo S, Pescarolo A, Lavezzo E, Di Camillo B, Facchinetti A, Cilia E, Velasco R, Fontana P (2012) Argot2: a large scale function prediction tool relying on semantic similarity of weighted Gene Ontology terms. BMC Bioinformatics **13 Suppl 4:** S14

Falkowski P, Raven J (2007) Aquatic photosynthesis, Ed 2nd. Princeton University Press, Princeton

Fang W, Si Y, Douglass S, Casero D, Merchant SS, Pellegrini M, Ladunga I, Liu P, Spalding MH (2012) Transcriptome-wide changes in *Chlamydomonas reinhardtii* gene expression regulated by carbon dioxide and the CO_2-concentrating mechanism regulator CIA5/CCM1. Plant Cell **24:** 1876-1893

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science **281:** 237-240

Förster B, Mathesius U, Pogson BJ (2006) Comparative proteomics of high light stress in the model alga *Chlamydomonas reinhardtii*. Proteomics **6:** 4309-4320

Fuhrmann M, Oertel W, Hegemann P (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. Plant Journal **19:** 353-361

Fujiwara S, Fukuzawa H, Tachiki A, Miyachi S (1990) Structure and differential expression of two genes encoding carbonic anhydrase in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences of the United States of America **87:** 9779-9783

Fujiwara S, Ishida N, Tsuzuki M (1996) Circadian expression of the carbonic anhydrase gene, *Cah1*, in *Chlamydomonas reinhardtii*. Plant Molecular Biology **32:** 745-749

Fukuzawa H, Fujiwara S, Yamamoto Y, Dionisiosese ML, Miyachi S (1990) cDNA cloning, sequence, and expression of carbonic anhydrase in *Chlamydomonas reinhardtii* - Regulation by environmental CO_2 concentration. Proceedings of the National Academy of Sciences of the United States of America **87:** 4383-4387

Fukuzawa H, Ishizaki K, Miura K, Matsueda S, Ino-ue T, Kucho K, Ohyama K (1998) Isolation and characterization of high CO_2-requiring mutants from *Chlamydomonas reinhardtii* by gene tagging. Canadian Journal of Botany **76:** 1092-1097

Fukuzawa H, Miura K, Ishizaki K, Kucho KI, Saito T, Kohinata T, Ohyama K (2001) *Ccm1*, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing CO_2 availability. Proceedings of the National Academy of Sciences of the United States of America **98:** 5347-5352

Funke RP, Kovar JL, Weeks DP (1997) Intracellular carbonic anhydrase is essential to photosynthesis in *Chlamydomonas reinhardtii* at atmospheric levels of CO_2: Demonstration via genomic complementation of the high-CO_2-requiring mutant *ca-1*. Plant Physiology **114:** 237-244

Gao XH, Zaffagnini M, Bedhomme M, Michelet L, Cassier-Chauvat C, Decottignies P, Lemaire SD (2010) Biochemical characterization of glutaredoxins from *Chlamydomonas reinhardtii*: kinetics and specificity in deglutathionylation reactions. FEBS Letters **584:** 2242-2248

Genkov T, Meyer M, Griffiths H, Spreitzer RJ (2010) Functional hybrid rubisco enzymes with plant small subunits and algal large subunits: engineered *rbcS* cDNA for expression in *Chlamydomonas*. Journal of Biological Chemistry **285:** 19833-19841

Geraghty AM, Spalding MH (1996) Molecular and structural changes in *Chlamydomonas* under limiting CO_2 - A possible mitochondrial role in adaptation. Plant Physiology **111:** 1339-1347

Giordano M, Beardall J, Raven JA (2005) CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. Annual Review of Plant Biology **56:** 99-131

Giordano M, Norici A, Forssen M, Eriksson M, Raven JA (2003) An anaplerotic role for mitochondrial carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Physiology **132:** 2126-2134

Gong W, He K, Covington M, Dinesh-Kumar SP, Snyder M, Harmer SL, Zhu YX, Deng XW (2008) The development of protein microarrays and their applications in DNA-protein and protein-protein interaction analyses of *Arabidopsis* transcription factors. Molecular Plant **1:** 27-41

Goodenough UW, Levine RP (1970) Chloroplast structure and function in *ac-20*, a mutant strain of *Chlamydomonas reinhardtii*. III. Chloroplast ribosomes and membrane organization. Journal of Cell Biology **44:** 547-562

Hackenberg C, Huege J, Engelhardt A, Wittink F, Laue M, Matthijs HC, Kopka J, Bauwe H, Hagemann M (2012) Low-carbon acclimation in carboxysome-less and photorespiratory mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. Microbiology **158:** 398-413

Hanawa Y, Watanabe M, Karatsu Y, Fukuzawa H, Shiraiwa Y (2007) Induction of a high-CO_2-inducible, periplasmic protein, H43, and its application as a high-CO_2-responsive marker for study of the high-CO_2-sensing mechanism in *Chlamydomonas reinhardtii*. Plant and Cell Physiology **48:** 299-309

Hanson DT, Franklin LA, Samuelsson G, Badger MR (2003) The *Chlamydomonas reinhardtii cia3* mutant lacking a thylakoid lumen-localized carbonic anhydrase is limited by CO_2 supply to Rubisco and not Photosystem II function in vivo. Plant Physiology **132:** 2267-2275

Harada H, Nakajima K, Sakaue K, Matsuda Y (2006) CO_2 sensing at ocean surface mediated by cAMP in a marine diatom. Plant Physiology **142:** 1318-1328

Häusler RE, Geimer S, Kunz HH, Schmitz J, Dörmann P, Bell K, Hetfeld S, Guballa A, Flügge UI (2009) Chlororespiration and grana hyperstacking: how an Arabidopsis double mutant can survive despite defects in starch biosynthesis and daily carbon export from chloroplasts. Plant Physiology **149:** 515-533

Höhner R, Barth J, Magneschi L, Jaeger D, Niehues A, Bald T, Grossman A, Fufezan C, Hippler M (2013) The metabolic status drives acclimation of iron deficiency responses in *Chlamydomonas reinhardtii* as revealed by proteomics based hierarchical clustering and reverse genetics. Molecular and Cellular Proteomics **12:** 2774-2790

Holdsworth RH (1971) Isolation and partial characterization of pyrenoid protein of *Eremosphaera viridis*. Journal of Cell Biology **51:** 499-513

Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. Nucleic Acids Research **35:** W585-587

Hu Y, Chen ZW, Liu WZ, Liu XL, He YK (2008) Chloroplast division is regulated by the circadian expression of *FTSZ* and *MIN* genes in *Chlamydomonas reinhardtii*. European Journal of Phycology **43:** 207-215

Hwang SB, Herrin DL (1994) Control of *lhc* gene transcription by the circadian clock in *Chlamydomonas reinhardtii*. Plant Molecular Biology **26:** 557-569

Im CS, Grossman AR (2002) Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. Plant Journal **30:** 301-313

Im CS, Zhang Z, Shrager J, Chang CW, Grossman AR (2003) Analysis of light and CO_2 regulation in *Chlamydomonas reinhardtii* using genome-wide approaches. Photosynthesis Research **75:** 111-125

Ito H, Mutsuda M, Murayama Y, Tomita J, Hosokawa N, Terauchi K, Sugita C, Sugita M, Kondo T, Iwasaki H (2009) Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. Proceedings of the National Academy of Sciences of the United States of America **106:** 14168-14173

Izumi M, Tsunoda H, Suzuki Y, Makino A, Ishida H (2012) *RBCS1A* and *RBCS3B*, two major members within the Arabidopsis *RBCS* multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. Journal of Experimental Botany **63:** 2159-2170

Jeffrey SW, Humphrey GF (1975) New spectrophotometric equations for determining chlorophylls *a*, *b*, c_1, and c_2 in higher plants, algae and natural phytoplankton. Biochemie und Physiologie der Pflanze **165:** 191-194

Jungnick N, Ma Y, Mukherjee B, Cronan JC, Speed DJ, Laborde SM, Longstreth DJ, Moroney JV (2014) The carbon concentrating mechanism in *Chlamydomonas reinhardtii*: finding the missing pieces. Photosynthesis Research **121:** 159-173

Kaplan A, Reinhold L (1999) CO_2 concentrating mechanisms in photosynthetic microorganisms. Annual Review of Plant Physiology and Plant Molecular Biology **50:** 539-570

Karlsson J, Clarke AK, Chen ZY, Hugghins SY, Park YI, Husic HD, Moroney JV, Samuelsson G (1998) A novel α-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO_2. EMBO Journal **17:** 1208-1216

Katoh A, Lee KS, Fukuzawa H, Ohyama K, Ogawa T (1996a) *cemA* homologue essential to CO_2 transport in the cyanobacterium *Synechocystis* PCC6803. Proceedings of the National Academy of Sciences of the United States of America **93:** 4006-4010

Katoh A, Sonoda M, Katoh H, Ogawa T (1996b) Absence of light-induced proton extrusion in a *cotA*-less mutant of *Synechocystis* sp. strain PCC6803. Journal of Bacteriology **178:** 5452-5455

Katzman GL, Carlson SJ, Marcus Y, Moroney JV, Togasaki RK (1994) Carbonic anhydrase activity in isolated chloroplasts of wild-type and high CO_2-dependent mutants of *Chlamydomonas reinhardtii* as studied by a new assay. Plant Physiology **105:** 1197-1202

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols **4:** 363-371

Kerby NW, Evans LV (1978) Isolation and partial characterisation of pyrenoids from the brown alga *Pilayella littoralis* (L.) Kjellm. Planta **142:** 91-95

Kiessling J, Kruse S, Rensing SA, Harter K, Decker EL, Reski R (2000) Visualization of a cytoskeleton-like FtsZ network in chloroplasts. Journal of Cell Biology **151:** 945-950

Kikutani S, Tanaka R, Yamazaki Y, Hara S, Hisabori T, Kroth PG, Matsuda Y (2012) Redox regulation of carbonic anhydrases via thioredoxin in chloroplast of the marine diatom *Phaeodactylum tricornutum*. Journal of Biological Chemistry **287:** 20689-20700

Kim TH, Bohmer M, Hu H, Nishimura N, Schroeder JI (2010) Guard cell signal transduction network: advances in understanding abscisic acid, CO_2, and Ca^{2+} signaling. Annual Review of Plant Biology **61:** 561-591

Kirchhoff H (2014) Structural changes of the thylakoid membrane network induced by high light stress in plant chloroplasts. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **369:** 20130225

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

Kropat J, Hong-Hermesdorf A, Casero D, Ent P, Castruita M, Pellegrini M, Merchant SS, Malasarn D (2011) A revised mineral nutrient supplement increases biomass and growth rate in *Chlamydomonas reinhardtii*. Plant Journal **66:** 770-780

Kuchitsu K, Tsuzuki M, Miyachi S (1988) Characterization of the pyrenoid isolated from unicellular green alga *Chlamydomonas reinhardtii*: Particulate form of rubisco protein. Protoplasma **144:** 17-24

Kuchitsu K, Tsuzuki M, Miyachi S (1991) Polypeptide composition and enzyme activities of the pyrenoid and its regulation by CO_2 concentration in unicellular green algae. Canadian Journal of Botany **69:** 1062-1069

Kucho K, Ohyama K, Fukuzawa H (1999) CO_2-responsive transcriptional regulation of *CAH1* encoding carbonic anhydrase is mediated by enhancer and silencer regions in *Chlamydomonas reinhardtii*. Plant Physiology **121:** 1329-1337

Kucho K, Yoshioka S, Taniguchi F, Ohyama K, Fukuzawa H (2003) *Cis*-acting elements and DNA-binding proteins involved in CO_2-responsive transcriptional activation of *Cah1* encoding a periplasmic carbonic anhydrase *Chlamydomonas reinhardtii*. Plant Physiology **133:** 783-793

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:** 680-685

Laurent JM, Vogel C, Kwon T, Craig SA, Boutz DR, Huse HK, Nozue K, Walia H, Whiteley M, Ronald PC, Marcotte EM (2010) Protein abundances are more conserved than mRNA abundances across diverse taxa. Proteomics **10:** 4209-4212

Layana C, Diambra L (2011) Time-course analysis of cyanobacterium transcriptome: detecting oscillatory genes. PLoS ONE **6:** e26291

Lin MT, Occhialini A, Andralojc JP, Devonshire J, Hines KM, Parry MA, Hanson MR (2014) β-carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. Plant Journal **79:** 1-12

Lin S, Carpenter EJ (1997a) Pyrenoid localization of Rubisco in relation to the cell cycle and growth phase of *Dunaliella tertiolecta* (Chlorophyceae). Phycologia **36:** 24-31

Lin S, Carpenter EJ (1997b) Rubisco of *Dunaliella tertiolecta* is redistributed between the pyrenoid and the stroma as a light/shade response. Marine Biology **127:** 521-529

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

Lohbeck KT, Riebesell U, Reusch TB (2014) Gene expression changes in the coccolithophore *Emiliania huxleyi* after 500 generations of selection to ocean acidification. Proceedings of the Royal Society. Series B, Biological Sciences **281:** 20140003

Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell and Environment **37:** 1250-1258

Long BM, Rae BD, Badger MR, Price GD (2011) Over-expression of the β-carboxysomal CcmM protein in *Synechococcus* PCC7942 reveals a tight co-regulation of carboxysomal carbonic anhydrase (CcaA) and M58 content. Photosynthesis Research **109:** 33-45

Lundquist PK, Poliakov A, Giacomelli L, Friso G, Appel M, McQuinn RP, Krasnoff SB, Rowland E, Ponnala L, Sun Q, van Wijk KJ (2013) Loss of plastoglobule kinases ABC1K1 and ABC1K3 causes conditional degreening, modified prenyl-lipids, and recruitment of the jasmonic acid pathway. Plant Cell **25:** 1818-1839

Ma Y, Pollock SV, Xiao Y, Cunnusamy K, Moroney JV (2011) Identification of a novel gene, *CIA6*, required for normal pyrenoid formation in *Chlamydomonas reinhardtii*. Plant Physiology **156:** 884-896

Maberly SC (1996) Diel, episodic and seasonal changes in pH and concentrations of inorganic carbon in a productive lake. Freshwater Biology **35:** 579-598

Manuel LJ, Moroney JV (1988) Inorganic carbon accumulation by *Chlamydomonas reinhardtii*: new proteins are made during adaptation to low CO_2. Plant Physiology **88:** 491-496

Marcus Y, Schuster G, Michaels A, Kaplan A (1986) Adaptation to CO_2 levels and changes in the phosphorylation of thylakoid proteins during the cell cycle of *Chlamydomonas reinhardtii*. Plant Physiology **80:** 604-607

Marek LF, Spalding MH (1991) Changes in photorespiratory enzyme activity in response to limiting CO_2 in *Chlamydomonas reinhardtii*. Plant Physiology **97:** 420-425

Mariscal V, Moulin P, Orsel M, Miller AJ, Fernandez E, Galvan A (2006) Differential regulation of the *Chlamydomonas Nar1* gene family by carbon and nitrogen. Protist **157:** 421-433

Markelova AG, Sinetova MP, Kupriyanova EV, Pronina NA (2009) Distribution and functional role of carbonic anhydrase Cah3 associated with thylakoid membranes in the chloroplast and pyrenoid of *Chlamydomonas reinhardtii*. Russian Journal of Plant Physiology **56:** 761-768

Marri L, Thieulin-Pardo G, Lebrun R, Puppo R, Zaffagnini M, Trost P, Gontero B, Sparla F (2014) CP12-mediated protection of Calvin-Benson cycle enzymes from oxidative stress. Biochimie **97:** 228-237

Martin A, Lang D, Hanke ST, Mueller SJ, Sarnighausen E, Vervliet-Scheebaum M, Reski R (2009) Targeted gene knockouts reveal overlapping functions of the five *Physcomitrella patens* FtsZ isoforms in chloroplast division, chloroplast shaping, cell patterning, plant development, and gravity sensing. Molecular Plant **2:** 1359-1372

Maruyama S, Tokutsu R, Minagawa J (2014) Transcriptional regulation of the stress-responsive light harvesting complex genes in *Chlamydomonas reinhardtii*. Plant and Cell Physiology **55:** 1304-1310

Mastrobuoni G, Irgang S, Pietzke M, Aßmus HE, Wenzel M, Schulze WX, Kempa S (2012) Proteome dynamics and early salt stress response of the photosynthetic organism *Chlamydomonas reinhardtii*. BMC Genomics **13:** 215

Matsuda Y, Nakajima K, Tachibana M (2011) Recent progresses on the genetic basis of the regulation of CO_2 acquisition systems in response to CO_2 concentration. Photosynthesis Research **109:** 191-203

McGinn PJ, Jones MJ, Macdonald AB, Campbell DA (2005) Light is required for low-CO_2-mediated induction of transcripts encoding components of the CO_2-concentrating mechanism in the cyanobacterium *Synechococcus elongatus*: analysis by quantitative reverse transcription-polymerase chain reaction. Canadian Journal of Botany **83:** 711-720

McGinn PJ, Price GD, Badger MR (2004) High light enhances the expression of low-CO_2-inducible transcripts involved in the CO_2-concentrating mechanism in *Synechocystis* sp PCC6803. Plant Cell and Environment **27:** 615-626

McGrath JM, Long SP (2014) Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. Plant Physiology **164:** 2247-2261

McKay RML, Gibbs SP (1991) Composition and function of pyrenoids: Cytochemical and immunocytochemical approaches. Canadian Journal of Botany **69:** 1040-1052

McKay RML, Gibbs SP, Vaughn KC (1991) Rubisco activase is present in the pyrenoid of green algae. Protoplasma **162:** 38-45

McKew BA, Lefebvre SC, Achterberg EP, Metodieva G, Raines CA, Metodiev MV, Geider RJ (2013) Plasticity in the proteome of *Emiliania huxleyi* CCMP 1516 to extremes of light is highly targeted. New Phytologist **200:** 61-73

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Marechal-Drouard L, Marshall WF, Qu LH, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov VV, Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, Cardol P, Cerutti H, Chanfreau G, Chen CL, Cognat V, Croft MT, Dent R, Dutcher S, Fernandez E, Fukuzawa H, Gonzalez-Ballester D, Gonzalez-Halphen D, Hallmann A, Hanikenne M, Hippler M, Inwood W, Jabbari K, Kalanon M, Kuras R, Lefebvre PA, Lemaire SD, Lobanov AV, Lohr M, Manuell A, Meier I, Mets L, Mittag M, Mittelmeier T, Moroney JV, Moseley J, Napoli C, Nedelcu AM, Niyogi K, Novoselov SV, Paulsen IT, Pazour G, Purton S, Ral JP, Riaño-Pachón DM, Riekhof W, Rymarquis L, Schroda M, Stern D, Umen J, Willows R, Wilson N, Zimmer SL, Allmer J, Balk J, Bisova K, Chen CJ, Elias M, Gendler K, Hauser C, Lamb MR, Ledford H, Long JC, Minagawa J, Page MD, Pan J, Pootakham W, Roje S, Rose A, Stahlberg E, Terauchi AM, Yang P, Ball S, Bowler C, Dieckmann CL, Gladyshev VN, Green P, Jorgensen R, Mayfield S, Mueller-Roeber B, Rajamani S, Sayre RT, Brokstein P, Dubchak I, Goodstein D, Hornick L, Huang YW, Jhaveri J, Luo Y, Martinez D, Ngau WC, Otillar B, Poliakov A, Porter A, Szajkowski L, Werner G, Zhou K, Grigoriev IV, Rokhsar DS, Grossman AR (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science **318:** 245-250

Metodieva G, Nogueira-De-Souza NC, Greenwood C, Al-Janabi K, Leng L, Bucala R, Metodiev MV (2013) CD74-dependent deregulation of the tumor suppressor scribble in human epithelial and breast cancer cells. Neoplasia **15:** 660-668

Mettler T, Muhlhaus T, Hemme D, Schottler MA, Rupprecht J, Idoine A, Veyel D, Pal SK, Yaneva-Roder L, Winck FV, Sommer F, Vosloh D, Seiwert B, Erban A, Burgos A, Arvidsson S, Schonfelder S, Arnold A, Gunther M, Krause U, Lohse M, Kopka J, Nikoloski Z, Mueller-Roeber B, Willmitzer L, Bock R, Schroda M, Stitt M (2014) Systems analysis of the response of photosynthesis, metabolism, and growth to an increase in irradiance in the photosynthetic model organism *Chlamydomonas reinhardtii*. Plant Cell **26:** 2310-2350

Meyer M, Griffiths H (2013) Origins and diversity of eukaryotic CO_2-concentrating mechanisms: lessons for the future. Journal of Experimental Botany **64:** 769-786

Meyer MT, Genkov T, Skepper JN, Jouhet J, Mitchell MC, Spreitzer RJ, Griffiths H (2012) Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. Proceedings of the National Academy of the United States of America **109:** 19474-19479

Michelet L, Zaffagnini M, Massot V, Keryer E, Vanacker H, Miginiac-Maslow M, Issakidis-Bourguet E, Lemaire SD (2006) Thioredoxins, glutaredoxins, and glutathionylation: new crosstalks to explore. Photosynthesis Research **89:** 225-245

Mitchell MC, Meyer MT, Griffiths H (2014) Dynamics of carbon concentrating mechanism induction and protein relocalisation during the dark to light transition in synchronised *Chlamydomonas reinhardtii*. Plant Physiology **166:** 1073-1082

Mitra M, Lato SM, Ynalvez RA, Xiao Y, Moroney JV (2004) Identification of a new chloroplast carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Physiology **135:** 173-182

Miura K, Yamano T, Yoshioka S, Kohinata T, Inoue Y, Taniguchi F, Asamizu E, Nakamura Y, Tabata S, Yamato KT, Ohyama K, Fukuzawa H (2004) Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Physiology **135:** 1595-1607

Molnar A, Bassett A, Thuenemann E, Schwach F, Karkare S, Ossowski S, Weigel D, Baulcombe D (2009) Highly specific gene silencing by artificial microRNAs in the unicellular alga *Chlamydomonas reinhardtii*. Plant Journal **58:** 165-174

Monné M, Palmieri F (2014) Antiporters of the mitochondrial carrier family. Current Topics in Membranes **73:** 289-320

Morisse S, Zaffagnini M, Gao XH, Lemaire SD, Marchand CH (2014) Insight into protein S-nitrosylation in *Chlamydomonas reinhardtii*. Antioxidants and Redox Signaling **21:** 1271-1284

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H (1999) Role of pyrenoids in the CO_2-concentrating mechanism: comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of *Chlamydomonas* and *Chloromonas* (Volvocales). Planta **208:** 365-372

Moroney JV, Husic HD, Tolbert NE, Kitayama M, Manuel LJ, Togasaki RK (1989) Isolation and characterization of a mutant of *Chlamydomonas reinhardtii* deficient in the CO_2 concentrating mechanism. Plant Physiology **89:** 897-903

Moroney JV, Ma Y, Frey WD, Fusilier KA, Pham TT, Simms TA, DiMario RJ, Yang J, Mukherjee B (2011) The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. Photosynthesis Research **109:** 133-149

Moroney JV, Tolbert NE (1985) Inorganic carbon uptake by *Chlamydomonas reinhardtii*. Plant Physiology **77:** 253-258

Moroney JV, Tolbert NE, Sears BB (1986) Complementation analysis of the inorganic carbon concentrating mechanism of *Chlamydomonas reinhardtii*. Molecular and General Genetics **204:** 199-203

Moroney JV, Ynalvez RA (2007) Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. Eukaryotic Cell **6:** 1251-1259

Nacir H, Brehelin C (2013) When proteomics reveals unsuspected roles: the plastoglobule example. Frontiers in Plant Science **4:** 114

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

Nemoto W, Saito A, Oikawa H (2013) Recent advances in functional region prediction by using structural and evolutionary information - Remaining problems and future extensions. Computational and Structural Biotechnology Journal **8:** e201308007

Nishimura T, Takahashi Y, Yamaguchi O, Suzuki H, Maeda S, Omata T (2008) Mechanism of low CO_2-induced activation of the *cmp* bicarbonate transporter operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus* strain PCC 7942. Molecular Microbiology **68:** 98-109

Ohnishi N, Mukherjee B, Tsujikawa T, Yanase M, Nakano H, Moroney JV, Fukuzawa H (2010) Expression of a low CO_2-inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. Plant Cell **22:** 3105-3117

Ohno N, Inoue T, Yamashiki R, Nakajima K, Kitahara Y, Ishibashi M, Matsuda Y (2012) CO_2-cAMP-responsive *cis*-elements targeted by a transcription factor with CREB/ATF-like basic zipper domain in the marine diatom *Phaeodactylum tricornutum*. Plant Physiology **158:** 499-513

Okabe Y, Okada M (1988) Isolation of native pyrenoid core matrix having ribulose 1,5-bisphosphate carboxylase activity from the green alga *Bryopsis maxima*. Plant and Cell Physiology **29:** 89-96

Osteryoung KW, Pyke KA (2014) Division and dynamic morphology of plastids. Annual Review of Plant Biology **65:** 443-472

Palmqvist K, Yu JW, Badger MR (1994) Carbonic anhydrase activity and inorganic carbon fluxes in low-Ci and high-Ci cells of *Chlamydomonas reinhardtii* and *Scenedesmus obliquus*. Physiologia Plantarum **90:** 537-547

Pandhal J, Wright PC, Biggs CA (2007) A quantitative proteomic analysis of light adaptation in a globally significant marine cyanobacterium *Prochlorococcus marinus* MED4. Journal of Proteome Research **6:** 996-1005

Park Y-I, Karlsson J, Rojdestvenski I, Pronina NA, Klimov VV, Öquist G, Samuelsson G (1999) Role of a novel photosystem II-associated carbonic anhydrase in photosynthetic carbon assimilation in *Chlamydomonas reinhardtii*. FEBS Letters **444:** 102-105

Pazour GJ, Agrin N, Leszyk J, Witman GB (2005) Proteomic analysis of a eukaryotic cilium. Journal of Cell Biology **170:** 103-113

Pollock SV, Prout DL, Godfrey AC, Lemaire SD, Moroney JV (2004) The *Chlamydomonas reinhardtii* proteins Ccp1 and Ccp2 are required for long-term growth, but are not necessary for efficient photosynthesis, in a low-CO_2 environment. Plant Molecular Biology **56:** 125-132

Pootakham W, Gonzalez-Ballester D, Grossman AR (2010) Identification and regulation of plasma membrane sulfate transporters in Chlamydomonas. Plant Physiology **153:** 1653-1668

Price GD (2011) Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. Photosynthesis Research **109:** 47-57

Price GD, Badger MR (1989) Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO_2-requiring phenotype - evidence for a central role for carboxysomes in the CO_2 concentrating mechanism. Plant Physiology **91:** 505-513

Price GD, Badger MR, von Caemmerer S (2011) The prospect of using cyanobacterial bicarbonate transporters to improve leaf photosynthesis in C3 crop plants. Plant Physiology **155:** 20-26

Price GD, Badger MR, Woodger FJ, Long BM (2008) Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. Journal of Experimental Botany **59:** 1441-1461

Price GD, Pengelly JJ, Förster B, Du J, Whitney SM, von Caemmerer S, Badger MR, Howitt SM, Evans JR (2012) The cyanobacterial CCM as a source of genes for improving photosynthetic CO_2 fixation in crop species. Journal of Experimental Botany **64:** 753-768

Pronina NA, Borodin VV (1993) CO_2 stress and CO_2 concentration mechanism: investigation by means of photosystem-deficient and carbonic anhydrase-deficient mutants of *Chlamydomonas reinhardtii*. Photosynthetica **28:** 131-140

Pronina NA, Semenenko VE (1992) Role of the pyrenoid in concentration, generation, and fixation of CO_2 in the chloroplast of microalgae. Soviet Plant Physiology **39:** 470-476

Radivojac P, Clark WT, Oron TR, Schnoes AM, Wittkop T, Sokolov A, Graim K, Funk C, Verspoor K, Ben-Hur A, Pandey G, Yunes JM, Talwalkar AS, Repo S, Souza ML, Piovesan D, Casadio R, Wang Z, Cheng JL, Fang H, Gough J, Koskinen P, Toronen P, Nokso-Koivisto J, Holm L, Cozzetto D, Buchan DWA, Bryson K, Jones DT, Limaye B, Inamdar H, Datta A, Manjari SK, Joshi R, Chitale M, Kihara D, Lisewski AM, Erdin S, Venner E, Lichtarge O, Rentzsch R, Yang HX, Romero AE, Bhat P, Paccanaro A, Hamp T, Kassner R, Seemayer S, Vicedo E, Schaefer C, Achten D, Auer F, Boehm A, Braun T, Hecht M, Heron M, Honigschmid P, Hopf TA, Kaufmann S, Kiening M, Krompass D, Landerer C, Mahlich Y, Roos M, Bjorne J, Salakoski T, Wong A, Shatkay H, Gatzmann F, Sommer I, Wass MN, Sternberg MJE, Skunca N, Supek F, Bosnjak M, Panov P, Dzeroski S, Smuc T, Kourmpetis YAI, van Dijk ADJ, ter Braak CJF, Zhou YP, Gong QT, Dong XR, Tian WD, Falda M, Fontana P, Lavezzo E, Di Camillo B, Toppo S, Lan L, Djuric N, Guo YH, Vucetic S, Bairoch A, Linial M, Babbitt PC, Brenner SE, Orengo C, Rost B, Mooney SD, Friedberg I (2013) A large-scale evaluation of computational protein function prediction. Nature Methods **10:** 221-227

Rae BD, Long BM, Badger MR, Price GD (2013) Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. Microbiology and Molecular Biology Reviews **77:** 357-379

Ramazanov Z, Mason CB, Geraghty AM, Spalding MH, Moroney JV (1993) The low CO_2-inducible 36-kilodalton protein is localized to the chloroplast envelope of *Chlamydomonas reinhardtii*. Plant Physiology **101:** 1195-1199

Ramazanov Z, Rawat M, Henk MC, Mason CB, Matthews SW, Moroney JV (1994) The induction of the CO_2-concentrating mechanism is correlated with the formation of the starch sheath around the pyrenoid of *Chlamydomonas reinhardtii*. Planta **195:** 210-216

Ravanel S, Gakiere B, Job D, Douce R (1998) The specific features of methionine biosynthesis and metabolism in plants. Proceedings of the National Academy of Sciences of the United States of America **95:** 7805-7812

Raven JA (1997) CO_2-concentrating mechanisms: a direct role for thylakoid lumen acidification? Plant Cell and Environment **20:** 147-154

Raven, JA, Caldiera, K, Elderfield, H, Hoegh-Guldberg, O, Liss, P, Riebesell, U, Shepherd, P, Turley, C, Watson, A (2005) Ocean acidification due to increasing atmospheric carbon dioxide. Royal Society, London, UK.

Raven JA (2001) A role for mitochondrial carbonic anhydrase in limiting CO_2 leakage from low CO_2-grown cells of Chlamydomonas reinhardtii. Plant Cell and Environment **24:** 261-265

Raven JA (2010) Inorganic carbon acquisition by eukayotic algae: four current questions. Photosynthesis Research **106:** 123-134

Raven JA, Beardall J, Giordano M (2014) Energy costs of carbon dioxide concentrating mechanisms in aquatic organisms. Photosynthesis Research **12:** 111-124

Raven JA, Giordano M, Beardall J, Maberly SC (2012) Algal evolution in relation to atmospheric CO_2: carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **367:** 493-507

Rawat M, Henk MC, Lavigne LL, Moroney JV (1996) *Chlamydomonas reinhardtii* mutants without ribulose-1,5-bisphosphate carboxylase-oxygenase lack a detectable pyrenoid. Planta **198:** 263-270

Rawat M, Moroney JV (1995) The regulation of carbonic anhydrase and ribulose-1,5-bisphosphate carboxylase oxygenase activase by light and CO_2 in *Chlamydomonas reinhardtii*. Plant Physiology **109:** 937-944

Renberg L, Johansson AI, Shutova T, Stenlund H, Aksmann A, Raven JA, Gardestrom P, Moritz T, Samuelsson G (2010) A metabolomic approach to study major metabolite changes during acclimation to limiting CO_2 in *Chlamydomonas reinhardtii*. Plant Physiology **154:** 187-196

Reski R (2009) Challenges to our current view on chloroplasts. Biological Chemistry **390:** 731-738

Riaño-Pachón DM, Correa LGG, Trejos-Espinosa R, Mueller-Roeber B (2008) Green transcription factors: a Chlamydomonas overview. Genetics **179:** 31-39

Rolland N, Dorne AJ, Amoroso G, Sültemeyer DF, Joyard J, Rochaix JD (1997) Disruption of the plastid *ycf10* open reading frame affects uptake of inorganic carbon in the chloroplast of *Chlamydomonas*. EMBO Journal **16:** 6713-6726

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology **132:** 365-386

Sasaki Y, Sekiguchi K, Nagano Y, Matsuno R (1993) Chloroplast envelope protein encoded by chloroplast genome. FEBS Letters **316:** 93-98

Scheibe R, Dietz KJ (2012) Reduction-oxidation network for flexible adjustment of cellular metabolism in photoautotrophic cells. Plant Cell and Environment **35:** 202-216

Schloss JA (1990) A *Chlamydomonas* gene encodes a G protein β subunit-like polypeptide. Molecular and General Genetics **221:** 443-452

Schulze WX, Usadel B (2010) Quantitation in mass-spectrometry-based proteomics. Annual Review of Plant Biology **61:** 491-516

Shi D, Xu Y, Hopkinson BM, Morel FM (2010) Effect of ocean acidification on iron availability to marine phytoplankton. Science **327:** 676-679

Shi T, Falkowski PG (2008) Genome evolution in cyanobacteria: the stable core and the variable shell. Proceedings of the National Academy of Sciences of the United States of America **105:** 2510-2515

Shutova T, Kenneweg H, Buchta J, Nikitina J, Terentyev V, Chernyshov S, Andersson B, Allakhverdiev SI, Klimov VV, Dau H, Junge W, Samuelsson G (2008) The photosystem II-associated Cah3 *in Chlamydomonas* enhances the O_2 evolution rate by proton removal. EMBO Journal **27:** 782-791

Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK (2007) DisProt: the database of disordered proteins. Nucleic Acids Research **35:** D786-793

Sinetova MA, Kupriyanova EV, Markelova AG, Allakhverdiev SI, Pronina NA (2012) Identification and functional role of the carbonic anhydrase Cah3 in thylakoid membranes of pyrenoid of *Chlamydomonas reinhardtii*. Biochimica Et Biophysica Acta-Bioenergetics **1817:** 1248-1255

Skepper JN (2000) Immunocytochemical strategies for electron microscopy: choice or compromise. Journal of Microscopy **199:** 1-36

Spalding MH, Critchley C, Govindjee, Orgren WL (1984) Influence of carbon dioxide concentration during growth on fluorescence induction characteristics of the green alga *Chlamydomonas reinhardii*. Photosynthesis Research **5:** 169-176

Spalding MH, Jeffrey M (1989) Membrane-associated polypeptides induced in *Chlamydomonas* by limiting CO_2 concentrations. Plant Physiology **89:** 133-137

Spalding MH, Ogren WL (1982) Photosynthesis is required for induction of the CO_2-concentrating system in *Chlamydomonas reinhardii*. FEBS Letters **145:** 41-44

Spalding MH, Ogren WL (1983) Evidence for a saturable transport component in the inorganic carbon uptake of *Chlamydomonas reinhardii*. FEBS Letters **154:** 335-338

Spalding MH, Spreitzer RJ, Ogren WL (1983a) Carbonic anhydrase-deficient mutant of *Chlamydomonas reinhardii* requires elevated carbon dioxide concentration for photoautotrophic growth. Plant Physiology **73:** 268-272

Spalding MH, Spreitzer RJ, Ogren WL (1983b) Genetic and physiological analysis of the CO_2-concentrating system of *Chlamydomonas reinhardii*. Planta **159:** 261-266

Spalding MH, Spreitzer RJ, Ogren WL (1983c) Reduced inorganic carbon transport in a CO_2-requiring mutant of *Chlamydomonas reinhardii*. Plant Physiology **73:** 273-276

Spalding MH, Winder TL, Anderson JC, Geraghty AM, Marek LF (1991) Changes in protein and gene expression during induction of the CO_2-concentrating mechanism in wild-type and mutant *Chlamydomonas*. Canadian Journal of Botany **69:** 1008-1016

Spreitzer RJ (2003) Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Archives of Biochemistry and Biophysics **414:** 141-149

Spreitzer RJ, Mets L (1981) Photosynthesis-deficient mutants of *Chlamydomonas reinhardii* with associated light-sensitive phenotypes. Plant Physiology **67:** 565-569

Spreitzer RJ, Salvucci ME (2002) Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. Annual Review of Plant Biology **53:** 449-475

Sültemeyer DF, Miller AG, Espie GS, Fock HP, Canvin DT (1989) Active CO_2 transport by the green alga *Chlamydomonas reinhardtii*. Plant Physiology **89:** 1213-1219

Süss KH, Prokhorenko I, Adler K (1995) *In situ* association of Calvin cycle enzymes, ribulose-1,5-bisphosphate carboxylase/oxygenase activase, ferrodoxin-$NADP^+$ reductase, and nitrite reductase with thylakoid and pyrenoid membranes of *Chlamydomonas reinhardtii* chloroplasts as revealed by immunoelectron microscopy. Plant Physiology **107:** 1387-1397

Takusagawa F, Kamitori S, Markham GD (1996) Structure and function of S-adenosylmethionine synthetase: crystal structures of S-adenosylmethionine synthetase with ADP, BrADP, and PPi at 28 angstroms resolution. Biochemistry **35:** 2586-2596

Tavío MDP, Jiménez del Ro M, Garca Reina G, Ramazanov Z (1996) The presence of ribulose 1,5-bisphosphate carboxygenase/oxygenase is required for induction of the 37k Da periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. Scientia Marina **60:** 149-154

Taylor TC, Backlund A, Bjorhall K, Spreitzer RJ, Andersson I (2001) First crystal structure of rubisco from a green alga, *Chlamydomonas reinhardtii*. Journal of Biological Chemistry **276:** 48159-48164

Terashima M, Specht M, Hippler M (2011) The chloroplast proteome: a survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features. Current Genetics **57:** 151-168

Thoms S, Pahlow M, Wolf-Gladrow DA (2001) Model of the carbon concentrating mechanism in chloroplasts of eukaryotic algae. Journal of Theoretical Biology **208:** 295-313

Thyssen C, Hermes M, Sültemeyer D (2003) Isolation and characterisation of *Chlamydomonas reinhardtii* mutants with an impaired CO_2-concentrating mechanism. Planta **217:** 102-112

Tirumani S, Kokkanti M, Chaudhari V, Shukla M, Rao BJ (2014) Regulation of CCM genes in *Chlamydomonas reinhardtii* during conditions of light-dark cycles in synchronous cultures. Plant Molecular Biology **85:** 277-286

Tokoro T, Hosokawa S, Miyoshi E, Tada K, Watanabe K, Montani S, Kayanne H, Kuwae T (2014) Net uptake of atmospheric CO_2 by coastal submerged aquatic vegetation. Global Change Biology **20:**1873-1884

Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proceedings of the National Academy of the United States of America **76:** 4350-4354.

Turkina MV, Blanco-Rivero A, Vainonen JP, Vener AV, Villarejo A (2006) CO_2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. Proteomics **6:** 2693-2704

Valledor L, Furuhashi T, Hanak AM, Weckwerth W (2013) Systemic cold stress adaptation of *Chlamydomonas reinhardtii*. Molecular and Cellular Proteomics **12:** 2032-2047

Van K, Spalding MH (1999) Periplasmic carbonic anhydrase structural gene (*Cah1*) mutant in *Chlamydomonas reinhardtii*. Plant Physiology **120:** 757-764

Van K, Wang Y, Nakamura Y, Spalding MH (2001) Insertional mutants of *Chlamydomonas reinhardtii* that require elevated CO_2 for survival. Plant Physiology **127:** 607-614

Villand P, Eriksson M, Samuelsson G (1997) Carbon dioxide and light regulation of promoters controlling the expression of mitochondrial carbonic anhydrase in *Chlamydomonas reinhardtii*. Biochemical Journal **327:** 51-57

Villarejo A, Martinez F, Plumed MD, Ramazanov Z (1996a) The induction of the CO_2 concentrating mechanism in a starch-less mutant of *Chlamydomonas reinhardtii*. Physiologia Plantarum **98:** 798-802

Villarejo A, Reina GG, Ramazanov Z (1996b) Regulation of the low-CO_2-inducible polypeptides in *Chlamydomonas reinhardtii*. Planta **199:** 481-485

Villarejo A, Shutova T, Moskvin O, Forssen M, Klimov VV, Samuelsson G (2002) A photosystem II-associated carbonic anhydrase regulates the efficiency of photosynthetic oxygen evolution. EMBO Journal **21:** 1930-1938

Vogel C, Marcotte EM (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics **13:** 227-232

von Caemmerer S, Quick WP, Furbank RT (2012) The development of C_4 rice: current progress and future challenges. Science **336:** 1671-1672

Waltenberger H, Schneid C, Grosch JO, Bareiss A, Mittag M (2001) Identification of target mRNAs for the clock-controlled RNA-binding protein Chlamy 1 from *Chlamydomonas reinhardtii*. Molecular Genetics and Genomics **265:** 180-188

Wang D, Kong D, Wang Y, Hu Y, He Y, Sun J (2003) Isolation of two plastid division *ftsZ* genes from *Chlamydomonas reinhardtii* and its evolutionary implication for the role of FtsZ in plastid division. Journal of Experimental Botany **54:** 1115-1116

Wang H, Gau B, Slade WO, Juergens M, Li P, Hicks LM (2014a) The global phosphoproteome of *Chlamydomonas reinhardtii* reveals complex organellar phosphorylation in the flagella and thylakoid membrane. Molecular and Cellular Proteomics **13:** 2337-2353

Wang L, Yamano T, Kajikawa M, Hirono M, Fukuzawa H (2014b) Isolation and characterization of novel high-CO_2-requiring mutants of *Chlamydomonas reinhardtii*. Photosynthesis Research **121:**175-184

Wang Y, Spalding MH (2006) An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of the United States of America **103:** 10110-10115

Wang Y, Spalding MH (2014) LCIB in the *Chlamydomonas* CO_2-concentrating mechanism. Photosynthesis Research **121:**185-192.

Wang Y, Sun Z, Horken KM, Im CS, Xiang Y, Grossman AR, Weeks DP (2005) Analyses of CIA5, the master regulator of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*, and its control of gene expression. Canadian Journal of Botany **83:** 765-779

Wang YJ, Duanmu DQ, Spalding MH (2011) Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: inorganic carbon transport and CO_2 recapture. Photosynthesis Research **109:** 115-122

Wanke D, Kolukisaoglu HU (2010) An update on the ABCC transporter family in plants: many genes, many proteins, but how many functions? Plant Biology **12:** 15-25

Whitehead L, Long BM, Price GD, Badger MR (2014) Comparing the in vivo function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. Plant Physiology **165:** 398-411

Williams TG, Turpin DH (1987) The role of external carbonic anhydrase in inorganic carbon acquisition by *Chlamydomonas reinhardii* at alkaline pH. Plant Physiology **83:** 92-96

Winck FV, Arvidsson S, Riaño-Pachón DM, Hempel S, Koseska A, Nikoloski Z, Gomez Urbina DA, Rupprecht J, Mueller-Roeber B (2013) Genome-wide identification of regulatory elements and reconstruction of gene regulatory networks of the green alga *Chlamydomonas reinhardtii* under carbon deprivation. PLoS One **8:** e79909

Winck FV, Riaño-Pachón DM, Sommer F, Rupprecht J, Mueller-Roeber B (2012) The nuclear proteome of the green alga *Chlamydomonas reinhardtii*. Proteomics **12:** 95-100

Winder TL, Anderson JC, Spalding MH (1992) Translational regulation of the large and small subunits of ribulose bisphosphate carboxylase/oxygenase during induction of the CO_2-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Physiology **98:** 1409-1414

Wobbe L, Schwarz C, Nickelsen J, Kruse O (2008) Translational control of photosynthetic gene expression in phototrophic eukaryotes. Physiologia Plantarum **133:** 507-515

Woodger FJ, Badger MR, Price GD (2003) Inorganic carbon limitation induces transcripts encoding components of the CO_2-concentrating mechanism in *Synechococcus* sp. PCC7942 through a redox-independent pathway. Plant Physiology **133:** 2069-2080

Woodger FJ, Badger MR, Price GD (2005) Sensing of inorganic carbon limitation in *Synechococcus* PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. Plant Physiology **139:** 1959-1969

Xiang YB, Zhang J, Weeks DP (2001) The *Cia5* gene controls formation of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. Proceedings of the National Academy of Sciences of the United States of America **98:** 5341-5346

Yamano T, Asada A, Sato E, Fukuzawa H (2014) Isolation and characterization of mutants defective in the localization of LCIB, an essential factor for the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Photosynthesis Research **121:**193-200

Yamano T, Fukuzawa H (2009) Carbon-concentrating mechanism in a green alga, *Chlamydomonas reinhardtii*, revealed by transcriptome analyses. Journal of Basic Microbiology **49:** 42-51

Yamano T, Miura K, Fukuzawa H (2008) Expression analysis of genes associated with the induction of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Physiology **147:** 340-354

Yamano T, Tsujikawa T, Hatano K, Ozawa S, Takahashi Y, Fukuzawa H (2010) Light and low-CO_2-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant and Cell Physiology **51:** 1453-1468

Yang EJ, Oh YA, Lee ES, Park AR, Cho SK, Yoo YJ, Park OK (2003) Oxygen-evolving enhancer protein 2 is phosphorylated by glycine-rich protein 3/wall-associated kinase 1 in *Arabidopsis*. Biochemical and Biophysical Research Communications **305:** 862-868

Yang MK, Qiao ZX, Zhang WY, Xiong Q, Zhang J, Li T, Ge F, Zhao JD (2013) Global phosphoproteomic analysis reveals diverse functions of serine/threonine/tyrosine phosphorylation in the model cyanobacterium *Synechococcus* sp. strain PCC 7002. Journal of Proteome Research **12:** 1909-1923

Ynalvez RA, Moroney JV (2008) Identification and characterisation of a novel inorganic carbon acquisition gene, *CIA7*, from an insertional mutant of *Chlamydomonas reinhardtii*. Functional Plant Biology **35:** 373-381

Yoshioka S, Taniguchi F, Miura K, Inoue T, Yamano T, Fukuzawa H (2004) The novel Myb transcription factor LCR1 regulates the CO_2-responsive gene *Cah1*, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Cell **16:** 1466-1477

Zaffagnini M, Bedhomme M, Groni H, Marchand CH, Puppo C, Gontero B, Cassier-Chauvat C, Decottignies P, Lemaire SD (2012a) Glutathionylation in the photosynthetic model organism *Chlamydomonas reinhardtii*: a proteomic survey. Molecular and Cellular Proteomics **11:** M111 014142

Zaffagnini M, Bedhomme M, Marchand CH, Morisse S, Trost P, Lemaire SD (2012b) Redox regulation in photosynthetic organisms: focus on glutathionylation. Antioxidants and Redox Signaling **16:** 567-586

Zhang P, Sicora CI, Vorontsova N, Allahverdiyeva Y, Battchikova N, Nixon PJ, Aro EM (2007) FtsH protease is required for induction of inorganic carbon acquisition complexes in *Synechocystis* sp. PCC 6803. Molecular Microbiology **65:** 728-740

Zhang R, Patena W, Armbruster U, Gang SR, Blum SS, Jonikas MC (2014) High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. Plant Cell **26:**1398-1409

8 APPENDICES

Appendix 1. List of proteins identified in low CO_2-adapted pyrenoid-positive wild-type (WT) and pyrenoid-negative spinach RBCS hybrid (Sp) strains using LC-MS/MS analysis (Chapter 5). Differential expression of proteins (Sig. diff.) was supported by t-test (T), G-test (G) or both t-test and G-test (B). Non-differentially expressed proteins are listed below (-). Mean and standard error of label-free intensity (Int) and spectral counts (S.C.) are given for each protein. Proteins are listed alphabetically by UniProt identifier within each category of expression.

Appendix 2. Comparison of low CO_2-adapted wild-type and spinach RBCS hybrid LC-MS/MS proteomic data (Chapter 5) with the transcriptomic datasets of Dr Moritz Meyer (WT vs hybrids, unpublished data), Brueggeman et al., (2012) and Fang et al., (2012). Protein (W/S): W, protein was significantly more abundant in the wild-type; S, protein was significantly more abundant in spinach RBCS hybrid cells; o, protein was not found to be differentially expressed. WT vs hybrids: +, mRNA was more abundant in low CO_2-adapted in wild-type; -, mRNA was more abundant in higher plant RBCS hybrid strains; =, equal expression; o, no significant expression pattern. C- and S-effect: +, mRNA was more abundant at low CO_2/in wild-type cells; -, mRNA was more abundant at high CO_2/in *cia5* mutant; o, no significant differential expression. Overall DE: +, significantly differentially expressed gene; o, not differentially expressed. Proteins are listed alphabetically by UniProt identifier within each category of protein expression (differentially expressed, non-differentially compared to all transcriptome data and non-differentially expressed compared to WT vs hybrid transcriptome data only).

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8JGP9	N/A	3.5E+09	1.2E+09	0.0E+00	0.0E+00	30.7	8.9	0.0	0.0
Q43832	N/A	0.0E+00	0.0E+00	2.6E+09	3.5E+08	0.0	0.0	105.0	10.1
A1YSB4	B	0.0E+00	0.0E+00	1.7E+07	3.7E+06	0.0	0.0	1.7	0.3
A8HS48	B	2.3E+09	9.6E+07	1.4E+09	8.0E+07	16.8	1.2	26.7	4.0
A8I2V3	B	1.2E+09	1.1E+08	3.9E+08	4.1E+07	33.7	0.7	23.8	1.6
A8I3J7	B	0.0E+00	0.0E+00	1.8E+07	6.7E+06	0.0	0.0	1.3	0.5
A8I4P5	B	1.6E+09	1.9E+08	5.1E+08	8.8E+07	20.5	1.5	12.7	2.0
A8IGE2	B	2.3E+07	5.9E+06	0.0E+00	0.0E+00	1.2	0.2	0.0	0.0
A8IGM2	B	1.4E+08	1.0E+07	0.0E+00	0.0E+00	1.5	0.3	0.0	0.0
A8III5	B	6.7E+07	9.7E+06	0.0E+00	0.0E+00	1.2	0.2	0.0	0.0
A8IVF6	B	4.6E+06	6.5E+05	0.0E+00	0.0E+00	1.2	0.2	0.0	0.0
A8IWP6	B	0.0E+00	0.0E+00	3.0E+06	1.4E+06	0.0	0.0	1.3	0.6
A8IY43	B	3.4E+07	1.9E+07	0.0E+00	0.0E+00	1.8	0.8	0.0	0.0
A8J3W1	B	6.1E+07	1.8E+07	0.0E+00	0.0E+00	1.5	0.5	0.0	0.0
A8J7X3	B	0.0E+00	0.0E+00	5.1E+07	8.0E+06	0.0	0.0	1.3	0.3
A8J9T0	B	2.9E+07	4.9E+06	0.0E+00	0.0E+00	1.2	0.2	0.0	0.0
A8JHU0	B	5.8E+08	6.1E+07	2.8E+08	2.3E+07	25.8	2.2	39.8	3.8
Q0ZAI6	B	7.6E+08	8.1E+07	3.6E+08	1.8E+07	37.5	3.6	27.2	3.1
Q2VA40	B	0.0E+00	0.0E+00	2.7E+08	6.8E+07	0.0	0.0	6.5	1.2
Q39590	B	2.7E+08	7.2E+07	7.4E+06	3.4E+06	4.5	0.3	0.7	0.3
Q42690	B	1.5E+10	9.8E+08	9.4E+09	7.3E+08	103.8	3.6	123.5	19.7
Q6RBZ1	B	1.6E+07	6.1E+06	0.0E+00	0.0E+00	1.2	0.4	0.0	0.0
A8HMM7	T	0.0E+00	0.0E+00	2.1E+06	7.1E+05	0.0	0.0	0.8	0.2
A8HP50	T	3.4E+06	2.2E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8HPY4	T	3.1E+07	2.5E+06	0.0E+00	0.0E+00	1.0	0.0	0.0	0.0
A8HQ74	T	7.4E+07	2.7E+06	4.2E+07	2.7E+06	2.3	0.3	1.7	0.5
A8HQA8	T	1.5E+07	7.2E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8HRZ9	T	3.0E+07	5.3E+06	6.9E+06	2.5E+06	5.5	1.4	5.2	2.0
A8HSJ6	T	7.2E+06	2.4E+06	0.0E+00	0.0E+00	0.8	0.2	0.0	0.0
A8HX38	T	2.2E+10	7.7E+08	1.3E+10	6.1E+08	426.8	9.4	431.2	43.1
A8HYU5	T	1.2E+09	8.8E+07	5.8E+08	5.8E+07	19.3	2.3	16.7	2.7
A8HZX4	T	0.0E+00	0.0E+00	5.2E+06	2.3E+06	0.0	0.0	0.7	0.3
A8I0A3	T	2.1E+07	6.5E+06	0.0E+00	0.0E+00	0.8	0.2	0.0	0.0
A8I3W8	T	5.0E+06	2.5E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8I531	T	1.2E+07	2.6E+06	4.3E+06	4.1E+05	1.7	0.3	1.0	0.0
A8I6P9	T	2.2E+08	1.2E+07	1.0E+08	8.6E+06	2.0	0.0	2.0	0.0
A8I9T2	T	0.0E+00	0.0E+00	1.5E+07	1.6E+07	0.0	0.0	0.5	0.3
A8IBN3	T	1.1E+08	2.8E+07	3.0E+07	5.2E+06	5.2	1.2	3.3	0.9
A8ICF4	T	0.0E+00	0.0E+00	3.0E+07	1.2E+07	0.0	0.0	0.8	0.2
A8ICG9	T	7.7E+07	4.0E+06	6.2E+07	4.1E+06	1.3	0.3	1.0	0.0
A8ICT1	T	1.8E+08	1.3E+07	6.9E+07	8.0E+06	5.0	1.2	5.2	0.7
A8ICU7	T	0.0E+00	0.0E+00	1.6E+07	4.8E+06	0.0	0.0	0.8	0.2
A8IGV4	T	1.1E+07	5.8E+06	0.0E+00	0.0E+00	1.0	0.6	0.0	0.0
A8IHW6	T	1.7E+07	6.0E+06	0.0E+00	0.0E+00	1.0	0.4	0.0	0.0
A8IJ34	T	2.3E+08	1.7E+07	1.3E+08	1.0E+07	7.7	0.9	8.3	1.2
A8IJJ8	T	1.3E+07	6.3E+06	0.0E+00	0.0E+00	1.0	0.5	0.0	0.0
A8IMS6	T	0.0E+00	0.0E+00	5.9E+06	1.2E+06	0.0	0.0	1.0	0.0
A8INX1	T	1.0E+07	3.9E+06	0.0E+00	0.0E+00	1.0	0.4	0.0	0.0
A8IT01	T	1.3E+08	1.6E+07	2.4E+07	3.8E+06	3.3	0.5	1.0	0.0
A8ITZ2	T	8.4E+06	4.1E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8IUN8	T	1.6E+06	1.0E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8IV51	T	5.9E+06	3.8E+06	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0
A8IW47	T	2.5E+08	2.1E+07	1.3E+08	9.1E+06	4.3	0.5	4.7	0.7

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IX13	T	0.0E+00	0.0E+00	3.0E+06	1.4E+06	0.0	0.0	0.7	0.3
A8IX80	T	1.2E+09	9.9E+07	5.1E+08	5.5E+07	13.7	0.5	8.3	1.2
A8IXZ0	T	1.6E+09	1.1E+08	1.0E+09	6.4E+07	71.0	4.0	63.7	7.1
A8IY40	T	7.2E+06	3.3E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8J0I0	T	4.9E+09	3.6E+08	2.7E+09	1.7E+08	91.7	6.6	108.8	11.8
A8J0R7	T	1.4E+08	1.8E+07	6.9E+07	2.1E+07	2.8	0.8	1.5	0.5
A8J2J7	T	1.2E+05	5.8E+04	0.0E+00	0.0E+00	1.0	0.5	0.0	0.0
A8J3J0	T	8.1E+06	5.2E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J448	T	1.7E+07	7.9E+06	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
A8J462	T	0.0E+00	0.0E+00	8.9E+06	4.3E+05	0.0	0.0	1.0	0.0
A8J576	T	5.0E+08	2.7E+07	2.8E+08	1.6E+07	11.3	0.7	12.8	1.4
A8J646	T	3.0E+08	5.8E+07	1.6E+08	1.2E+07	5.5	0.6	4.5	0.8
A8J6M8	T	1.3E+06	8.0E+05	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J6Q7	T	8.6E+07	9.2E+06	2.7E+07	9.6E+06	4.2	0.8	2.2	0.7
A8J7P4	T	3.9E+07	1.2E+07	8.5E+06	3.6E+06	2.0	0.4	1.2	0.4
A8J995	T	2.3E+07	1.1E+07	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8JAW4	T	1.2E+09	1.4E+08	6.2E+08	5.7E+07	14.0	1.0	10.0	1.0
A8JBX6	T	1.4E+07	6.4E+06	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
A8JC15	T	6.3E+06	4.5E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8JCS8	T	2.6E+07	1.9E+07	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8JCY4	T	2.2E+08	5.4E+07	6.1E+07	7.6E+06	5.0	0.5	2.8	0.7
A8JCZ5	T	0.0E+00	0.0E+00	3.8E+06	2.5E+06	0.0	0.0	0.5	0.3
A8JD42	T	1.2E+07	3.7E+06	0.0E+00	0.0E+00	0.8	0.2	0.0	0.0
A8JDH1	T	3.1E+07	8.9E+06	1.0E+08	1.3E+07	2.7	0.3	4.2	0.4
A8JEP0	T	0.0E+00	0.0E+00	1.8E+07	1.8E+06	0.0	0.0	1.0	0.0
A8JEQ7	T	7.5E+06	3.6E+06	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
A8JEV9	T	1.0E+08	1.1E+07	6.6E+07	5.5E+06	2.8	0.2	2.8	0.4
A8JFF0	T	1.1E+07	5.3E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8JGI9	T	3.9E+08	4.2E+07	3.0E+08	1.9E+07	9.3	1.8	10.3	1.8
A8JHA4	T	3.4E+06	1.8E+06	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
B2XY88	T	3.8E+07	1.7E+07	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
I2FKQ9	T	2.9E+08	2.3E+07	1.7E+08	2.8E+07	7.2	0.7	7.8	1.2
O22547	T	7.4E+07	6.8E+06	1.2E+07	1.8E+06	3.7	0.3	3.0	0.5
P17746	T	1.5E+09	1.6E+08	6.4E+08	4.9E+07	61.0	9.1	56.7	4.4
Q00914	T	2.5E+09	1.8E+08	1.4E+09	8.8E+07	26.7	2.6	21.5	2.0
Q540H1	T	1.5E+09	8.8E+07	8.5E+08	7.6E+07	34.2	1.9	27.2	2.7
Q5NKW4	T	2.8E+09	5.8E+08	1.7E+09	9.6E+07	22.3	2.1	16.0	1.6
Q6X898	T	3.6E+09	2.2E+08	2.1E+09	1.6E+08	147.2	7.4	157.3	8.3
Q75T33	T	5.6E+08	4.3E+07	2.8E+08	2.6E+07	9.2	1.2	6.2	1.0
Q7X8Y6	T	0.0E+00	0.0E+00	1.2E+07	8.9E+06	0.0	0.0	0.8	0.4
Q944M9	T	0.0E+00	0.0E+00	2.7E+06	1.7E+06	0.0	0.0	0.5	0.3
Q9FQ96	T	5.3E+07	1.3E+07	1.8E+07	8.2E+06	4.0	1.0	3.3	2.3
Q9GGE2	T	3.1E+07	1.1E+07	6.1E+06	1.8E+06	1.7	0.3	0.8	0.2
A6Q0K5	G	2.0E+06	1.8E+06	1.1E+07	3.2E+06	0.3	0.3	3.0	0.5
A8HNC0	G	8.7E+07	6.6E+06	3.0E+07	1.4E+07	4.7	0.6	1.7	0.6
A8HNG8	G	2.8E+08	9.5E+07	3.5E+07	2.1E+07	5.2	1.1	1.7	0.9
A8HQP0	G	6.1E+08	4.4E+07	6.2E+08	5.7E+07	21.8	2.4	32.7	3.4
A8HWZ6	G	8.5E+07	2.0E+07	3.9E+07	1.4E+07	5.2	0.4	1.7	0.5
A8IB25	G	5.8E+08	5.4E+07	1.2E+08	4.7E+07	9.5	0.7	2.7	1.2
A8IEH0	G	7.5E+08	5.2E+07	7.4E+08	9.5E+07	23.5	3.9	15.0	2.3
A8IGD9	G	1.3E+08	1.4E+08	3.8E+05	1.2E+05	7.7	0.6	1.8	0.8
A8IIK4	G	1.6E+08	3.1E+07	1.5E+07	7.4E+06	3.3	0.7	0.7	0.3
A8IJQ4	G	2.4E+07	1.5E+07	1.1E+07	8.9E+06	18.5	4.3	11.5	3.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IKX4	G	1.7E+09	1.7E+08	1.6E+09	2.8E+08	28.5	3.7	40.5	6.7
A8IMP6	G	1.4E+09	6.6E+07	1.2E+09	1.1E+08	42.3	6.6	25.8	5.8
A8IQG4	G	1.3E+09	1.6E+08	7.9E+08	1.5E+08	22.3	5.3	35.5	5.4
A8IRT6	G	2.9E+08	9.0E+06	6.9E+07	2.0E+07	5.3	0.6	1.5	0.3
A8IRW1	G	7.9E+06	4.5E+06	0.0E+00	0.0E+00	1.2	0.4	0.0	0.0
A8IRX5	G	2.7E+08	5.8E+07	2.2E+08	2.7E+07	18.3	3.7	10.0	3.1
A8IS22	G	5.4E+08	3.7E+07	3.5E+08	5.0E+07	4.8	0.6	12.3	5.7
A8IW09	G	3.1E+08	3.0E+07	8.8E+07	5.6E+07	6.7	0.5	2.7	1.3
A8IWJ8	G	1.7E+08	2.1E+07	3.2E+07	2.1E+07	3.5	0.3	0.5	0.3
A8IY85	G	5.5E+07	8.0E+06	7.8E+06	4.9E+06	2.7	0.3	0.5	0.3
A8IZS5	G	1.1E+09	7.3E+07	4.5E+08	6.2E+07	69.8	8.0	47.3	3.5
A8IZW3	G	4.6E+06	3.1E+06	9.3E+05	1.3E+06	1.8	1.2	0.2	0.2
A8J0R4	G	1.7E+08	1.7E+07	4.3E+07	2.1E+07	3.2	0.4	7.7	2.3
A8J129	G	1.6E+09	1.3E+08	1.4E+09	8.8E+07	25.2	1.5	35.3	3.8
A8J225	G	3.9E+06	4.2E+06	0.0E+00	0.0E+00	1.7	0.3	0.0	0.0
A8J244	G	1.0E+10	6.6E+08	9.7E+09	5.5E+08	384.7	16.0	423.5	36.0
A8J2S0	G	1.6E+09	6.5E+07	1.5E+09	6.5E+07	43.5	2.4	57.5	1.9
A8J3Y6	G	2.2E+08	2.5E+07	5.8E+07	4.3E+07	6.7	0.7	1.8	1.5
A8J493	G	2.5E+08	3.4E+07	7.2E+07	3.6E+07	18.5	2.2	11.7	3.0
A8J4M0	G	1.5E+07	6.4E+06	2.0E+06	2.8E+06	1.8	0.4	0.2	0.2
A8J4Q3	G	1.9E+07	2.2E+07	2.0E+08	3.5E+07	1.5	1.6	4.8	1.4
A8J5B8	G	1.4E+09	4.9E+07	9.5E+08	8.1E+07	20.3	1.6	32.2	4.1
A8J6D1	G	5.5E+09	4.4E+08	4.5E+09	3.9E+08	89.3	8.7	109.3	4.7
A8J841	G	1.3E+08	1.5E+07	2.1E+07	1.4E+07	4.0	0.5	0.8	0.7
A8JB67	G	6.9E+07	9.0E+06	5.0E+06	4.6E+06	2.5	0.3	0.3	0.3
A8JCP5	G	1.4E+08	1.5E+07	3.9E+07	1.7E+07	6.7	0.8	2.8	1.1
A8JDE1	G	7.7E+06	2.8E+06	6.0E+07	1.9E+07	1.7	0.6	6.0	0.7
A8JEA7	G	5.7E+07	2.4E+06	3.8E+07	6.5E+06	4.5	1.4	1.0	0.0
A8JEV1	G	3.1E+09	2.8E+08	2.1E+09	3.4E+08	156.0	10.9	116.0	8.9
A8JFR9	G	5.0E+09	2.8E+08	4.3E+09	7.3E+08	124.5	14.5	197.2	11.7
A8JGF4	G	1.1E+09	1.2E+08	6.7E+08	6.9E+07	23.2	1.4	15.2	1.8
A8JGW2	G	1.3E+09	2.7E+08	5.5E+08	2.0E+08	17.0	1.6	10.0	1.5
A8JHX9	G	4.1E+09	1.3E+08	2.3E+09	2.6E+08	119.3	11.9	98.3	16.5
C5HJ53	G	4.0E+09	1.6E+08	2.3E+09	3.2E+08	33.8	1.6	47.3	4.6
O22472	G	1.2E+09	9.9E+07	1.2E+09	5.1E+07	14.7	1.7	22.8	2.5
O24450	G	3.0E+08	6.6E+07	3.2E+07	1.5E+07	8.2	2.5	1.5	0.7
P00877	G	3.0E+09	6.9E+08	2.7E+09	3.1E+08	80.7	10.4	122.5	10.0
P06007	G	6.3E+09	3.6E+08	5.4E+09	6.9E+08	59.0	2.1	81.0	9.8
P06541	G	6.9E+09	4.4E+08	4.7E+09	2.1E+08	239.5	12.4	301.2	18.2
P10898	G	7.3E+09	5.8E+08	5.8E+09	5.9E+08	127.0	8.6	149.2	17.4
P11471	G	7.5E+09	5.3E+08	4.6E+09	4.5E+08	338.8	14.0	286.8	28.1
P12154	G	6.0E+09	3.7E+08	6.3E+09	6.9E+08	48.8	5.8	79.3	10.6
P26526	G	6.8E+09	4.1E+08	4.2E+09	3.6E+08	84.3	3.4	106.3	15.2
Q08365	G	7.8E+08	1.0E+08	7.9E+08	1.3E+08	18.3	1.9	26.8	3.6
Q1RS84	G	1.3E+08	2.4E+07	3.7E+07	1.3E+07	5.5	0.8	2.2	0.8
Q39568	G	6.6E+08	8.3E+07	3.8E+08	2.9E+07	13.2	1.6	7.3	0.7
Q75VY6	G	3.3E+09	5.4E+08	2.6E+09	2.9E+08	83.3	21.3	128.0	20.1
Q75VY8	G	6.5E+08	1.3E+08	4.0E+08	1.0E+08	89.0	13.2	114.0	14.6
Q84V17	G	4.9E+07	2.1E+07	1.4E+08	2.0E+07	1.2	0.2	3.8	0.7
Q84X71	G	1.4E+07	1.1E+07	6.7E+07	3.8E+06	0.8	0.7	4.0	0.4
Q8GUQ9	G	1.3E+08	2.3E+07	5.6E+07	5.9E+06	8.8	1.8	16.5	6.0
Q9AXJ2	G	1.2E+09	4.7E+07	7.7E+08	1.1E+08	25.8	4.8	14.2	8.0
Q9FNS2	G	1.5E+09	2.0E+08	5.1E+08	9.7E+07	91.2	8.9	49.3	4.5

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
Q9S9E0	G	4.9E+09	2.0E+08	4.1E+09	2.6E+08	87.7	1.9	110.5	6.2
A2BCY1	-	1.9E+08	1.5E+07	1.5E+08	3.1E+07	3.5	0.3	4.8	0.9
A2PZB5	-	1.5E+08	8.4E+06	8.6E+07	9.3E+06	3.3	0.5	5.3	0.7
A2PZC1	-	2.3E+08	2.8E+07	2.5E+08	2.6E+07	10.7	1.2	10.5	1.2
A2PZC2	-	1.3E+09	6.7E+07	5.6E+08	7.9E+07	22.8	1.5	19.8	1.8
A2PZD0	-	3.4E+06	3.8E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A2PZD2	-	2.2E+07	1.4E+07	3.9E+07	3.2E+06	0.5	0.3	1.0	0.0
A5YU13	-	4.0E+07	1.0E+07	3.1E+07	2.3E+06	7.5	1.2	5.8	1.3
A8HM32	-	2.6E+07	2.7E+06	1.7E+07	2.4E+06	1.0	0.0	1.0	0.0
A8HM76	-	2.2E+05	3.1E+05	3.2E+06	2.8E+06	0.2	0.2	0.3	0.3
A8HMA8	-	5.3E+07	1.1E+07	1.7E+07	7.6E+06	1.5	0.5	0.7	0.3
A8HMC0	-	7.9E+07	9.5E+06	8.6E+07	2.1E+07	3.2	0.4	3.5	0.6
A8HMD3	-	7.4E+06	4.6E+06	1.7E+07	1.0E+07	2.0	0.6	2.3	0.5
A8HME4	-	7.3E+08	6.8E+07	6.8E+08	7.8E+07	26.8	2.0	23.2	3.6
A8HME6	-	4.7E+08	4.5E+07	2.0E+08	4.3E+07	21.3	2.1	14.3	2.6
A8HMG7	-	7.5E+08	7.3E+07	5.7E+08	3.3E+07	11.8	1.1	13.3	1.1
A8HMQ1	-	7.7E+09	3.2E+08	3.6E+09	5.5E+08	125.3	9.4	110.5	18.1
A8HMQ3	-	2.2E+07	1.2E+07	2.0E+07	4.1E+06	1.2	0.6	2.5	0.6
A8HMW1	-	6.5E+05	5.9E+05	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8HMX2	-	3.6E+08	2.8E+07	2.4E+08	1.8E+07	8.3	0.9	9.3	1.1
A8HN02	-	3.6E+08	3.4E+07	4.3E+08	5.4E+07	3.5	0.5	5.5	0.5
A8HN48	-	4.6E+06	4.1E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8HN50	-	1.2E+09	8.5E+07	8.0E+08	9.5E+07	12.7	1.1	11.0	1.4
A8HN65	-	2.5E+07	1.1E+07	2.8E+07	1.3E+07	1.3	0.3	1.3	0.3
A8HN88	-	1.9E+06	9.1E+05	1.2E+06	8.4E+05	0.7	0.3	0.5	0.3
A8HN92	-	3.1E+07	1.5E+06	3.4E+07	1.4E+07	2.2	0.4	2.5	1.2
A8HNB4	-	7.2E+06	6.4E+06	2.1E+06	2.7E+06	0.3	0.3	0.3	0.3
A8HND3	-	6.5E+07	8.1E+06	4.1E+07	6.9E+06	2.2	0.2	2.2	0.2
A8HNE3	-	3.2E+07	8.0E+06	1.3E+07	6.0E+06	1.3	0.3	0.7	0.3
A8HNE8	-	8.2E+08	4.9E+07	4.7E+08	4.8E+07	10.7	1.2	10.0	0.7
A8HNF4	-	2.3E+06	2.1E+06	4.0E+06	1.9E+06	0.3	0.3	0.7	0.3
A8HNJ8	-	1.3E+08	1.5E+07	1.5E+07	2.1E+07	1.2	0.2	0.3	0.5
A8HNN4	-	5.6E+06	2.4E+06	6.1E+06	6.2E+06	1.5	0.3	1.2	0.2
A8HNR8	-	9.3E+06	5.2E+06	2.4E+06	3.3E+06	0.7	0.3	0.2	0.2
A8HNX3	-	2.1E+08	6.1E+07	1.2E+08	7.7E+07	4.0	1.0	1.3	1.1
A8HNZ1	-	3.2E+07	2.0E+07	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0
A8HP06	-	4.7E+08	2.9E+07	2.8E+08	2.3E+07	12.5	1.3	8.5	1.1
A8HP10	-	1.5E+06	2.1E+06	0.0E+00	0.0E+00	0.3	0.5	0.0	0.0
A8HP17	-	2.6E+07	2.2E+07	4.8E+06	4.3E+06	0.7	0.5	0.3	0.3
A8HP55	-	2.1E+08	3.8E+07	2.2E+08	2.0E+07	3.3	0.9	3.8	0.6
A8HP58	-	5.5E+08	1.2E+08	3.5E+08	4.0E+07	4.5	0.7	3.8	0.6
A8HP61	-	1.9E+08	1.9E+07	2.0E+07	1.4E+07	3.5	0.3	1.2	0.4
A8HP64	-	0.0E+00	0.0E+00	1.6E+06	2.2E+06	0.0	0.0	0.2	0.2
A8HP72	-	9.3E+07	1.9E+07	7.1E+07	1.4E+07	3.3	0.9	4.5	1.3
A8HP84	-	1.0E+10	4.7E+08	8.4E+09	6.1E+08	308.3	19.8	333.2	24.0
A8HP90	-	5.5E+08	7.6E+07	3.1E+08	9.9E+07	19.2	1.1	18.2	2.1
A8HPA0	-	5.9E+06	5.3E+06	2.5E+06	3.5E+06	0.3	0.3	0.2	0.2
A8HPF4	-	1.1E+09	1.4E+08	5.7E+08	9.6E+07	47.8	6.8	38.5	4.6
A8HPG7	-	1.0E+08	6.8E+07	6.5E+06	9.2E+06	0.5	0.3	0.2	0.2
A8HPG8	-	6.0E+07	1.6E+07	3.6E+07	8.9E+06	2.5	0.7	1.8	0.4
A8HPI1	-	3.5E+07	3.6E+06	3.5E+06	3.2E+06	1.2	0.2	0.3	0.3
A8HPJ2	-	4.1E+08	5.6E+07	3.7E+08	4.3E+07	8.3	1.3	6.8	0.9
A8HPL4	-	7.8E+06	8.7E+06	2.2E+07	9.9E+06	0.8	0.6	1.2	0.4

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8HPL8	-	1.4E+08	3.3E+07	5.4E+07	2.3E+07	3.3	0.3	3.5	1.1
A8HPP0	-	7.5E+06	3.6E+06	6.9E+06	3.3E+06	1.0	0.4	1.0	0.4
A8HPS2	-	8.7E+07	1.7E+07	7.5E+07	1.2E+07	2.3	0.3	1.8	0.4
A8HPT4	-	1.7E+06	2.4E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8HPU3	-	2.3E+07	2.3E+06	3.1E+06	4.4E+06	1.0	0.0	0.3	0.5
A8HPV3	-	4.0E+07	1.6E+07	1.4E+07	7.0E+06	1.3	0.6	0.7	0.3
A8HPW8	-	3.3E+07	6.4E+06	6.8E+06	3.2E+06	2.2	0.4	0.8	0.4
A8HPZ8	-	2.2E+07	1.2E+07	2.9E+07	1.1E+07	1.0	0.4	1.5	0.5
A8HQ48	-	3.3E+07	7.4E+06	1.5E+07	9.7E+06	2.0	0.4	1.0	0.6
A8HQ72	-	1.4E+08	2.1E+07	5.3E+07	2.3E+07	4.2	0.8	3.7	0.8
A8HQ77	-	7.6E+07	5.0E+06	3.9E+07	6.1E+06	2.0	0.0	1.7	0.3
A8HQ81	-	2.8E+08	2.3E+07	1.3E+08	2.8E+07	13.7	3.9	11.3	2.6
A8HQ93	-	2.0E+07	1.3E+07	1.5E+07	6.9E+06	0.5	0.3	0.8	0.4
A8HQ97	-	3.5E+07	2.9E+06	1.2E+07	6.0E+06	1.0	0.0	0.8	0.4
A8HQE3	-	1.6E+06	7.2E+05	8.9E+04	1.3E+05	1.2	0.6	0.2	0.2
A8HQF0	-	8.7E+06	8.8E+05	6.3E+06	3.3E+05	1.0	0.0	1.0	0.0
A8HQG3	-	2.1E+06	6.7E+05	6.4E+06	2.2E+06	1.5	0.6	2.8	0.2
A8HQG8	-	0.0E+00	0.0E+00	1.9E+06	1.7E+06	0.0	0.0	0.3	0.3
A8HQJ3	-	1.3E+06	1.6E+05	4.8E+05	2.3E+05	1.0	0.0	0.7	0.3
A8HQT1	-	1.4E+08	2.1E+07	5.2E+07	4.1E+06	4.0	0.5	1.8	0.2
A8HQU3	-	6.6E+06	4.2E+06	1.1E+07	3.8E+06	0.5	0.3	1.2	0.2
A8HQW1	-	1.8E+08	3.4E+07	5.4E+07	1.5E+07	4.0	0.5	3.3	0.3
A8HR32	-	2.3E+06	1.6E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8HR39	-	4.1E+07	3.3E+06	3.7E+07	1.4E+06	1.5	0.3	1.0	0.0
A8HR57	-	2.5E+06	3.6E+06	4.6E+06	2.2E+06	0.2	0.2	0.7	0.3
A8HR79	-	3.5E+07	4.7E+06	5.7E+07	5.3E+06	1.3	0.3	2.3	0.5
A8HRF8	-	1.4E+08	1.4E+07	7.6E+07	1.6E+07	6.2	0.6	5.3	0.8
A8HRN8	-	6.7E+07	3.2E+07	9.2E+07	1.6E+07	1.3	0.7	1.8	0.2
A8HRP1	-	1.9E+07	6.7E+06	8.8E+06	5.2E+06	2.2	0.7	1.3	0.7
A8HRR9	-	3.6E+07	5.8E+06	4.5E+06	4.6E+06	1.7	0.3	0.3	0.3
A8HRS5	-	1.5E+08	2.0E+07	6.7E+07	1.4E+07	4.2	0.7	3.5	1.3
A8HRV2	-	4.5E+07	2.3E+07	6.6E+07	8.6E+06	1.5	0.6	2.7	0.3
A8HRZ0	-	8.3E+07	2.0E+07	1.2E+07	4.8E+06	1.8	0.4	1.2	0.4
A8HS14	-	4.5E+08	2.1E+07	2.6E+08	7.4E+07	13.8	1.7	9.5	2.2
A8HS59	-	8.7E+07	1.5E+07	4.8E+07	1.6E+07	47.8	3.6	39.3	3.9
A8HSB2	-	4.8E+05	3.2E+05	5.0E+05	1.8E+05	1.2	0.8	1.5	0.5
A8HSS0	-	2.0E+07	7.0E+06	2.0E+07	3.6E+06	1.0	0.4	1.2	0.2
A8HSU1	-	1.6E+08	1.1E+08	9.6E+07	2.8E+07	1.8	0.7	1.7	0.5
A8HSU7	-	1.2E+09	4.2E+07	7.6E+08	1.0E+08	19.0	1.6	14.2	0.8
A8HSV0	-	8.8E+06	5.6E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8HT91	-	6.0E+06	8.5E+06	9.4E+06	9.3E+06	0.2	0.2	0.3	0.3
A8HTG4	-	2.4E+07	1.2E+07	5.9E+06	4.5E+06	2.8	1.4	1.5	0.9
A8HTG7	-	1.2E+07	8.1E+06	3.6E+06	3.4E+06	0.8	0.6	0.3	0.3
A8HTK2	-	1.4E+06	2.0E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8HTK7	-	5.9E+06	2.8E+06	3.0E+06	1.9E+06	0.7	0.3	0.5	0.3
A8HTQ3	-	1.4E+07	2.0E+06	7.0E+06	8.1E+05	1.0	0.0	1.0	0.0
A8HTX0	-	6.6E+07	4.6E+06	2.7E+07	1.0E+07	1.8	0.6	2.2	0.8
A8HTX3	-	4.1E+06	3.7E+06	1.0E+07	5.7E+06	0.3	0.3	1.0	0.4
A8HTX7	-	5.2E+08	8.4E+07	2.8E+08	3.5E+07	9.0	0.9	6.5	0.7
A8HTX7	-	2.7E+07	4.2E+06	4.5E+06	4.0E+06	1.0	0.0	0.3	0.3
A8HTY0	-	3.8E+08	6.0E+07	8.5E+07	3.8E+07	3.2	0.4	2.8	1.8
A8HU20	-	4.4E+06	6.2E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8HU27	-	2.8E+06	2.5E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8HU30	-	2.3E+06	3.3E+06	6.7E+06	3.2E+06	0.2	0.2	0.7	0.3
A8HUE0	-	1.2E+07	1.1E+07	0.0E+00	0.0E+00	0.5	0.5	0.0	0.0
A8HUK0	-	2.1E+07	1.9E+07	9.1E+06	9.8E+06	0.3	0.3	0.3	0.3
A8HUU9	-	1.9E+08	1.7E+07	1.6E+08	1.3E+07	2.8	0.2	3.0	0.4
A8HVD7	-	9.0E+07	2.6E+07	4.5E+07	2.1E+07	0.8	0.2	0.7	0.3
A8HVF8	-	2.7E+07	1.7E+07	5.6E+07	2.0E+06	1.0	0.5	2.0	0.0
A8HVJ9	-	1.4E+08	2.2E+07	8.9E+07	2.7E+07	4.0	0.8	5.0	1.4
A8HVK4	-	1.1E+09	1.2E+08	1.5E+09	1.1E+08	6.0	0.4	8.5	1.0
A8HVM3	-	1.3E+08	7.7E+06	1.0E+08	1.4E+07	5.5	0.3	6.8	0.7
A8HVP2	-	2.3E+09	2.3E+08	1.5E+09	7.3E+07	35.3	1.2	30.5	0.8
A8HVP7	-	1.1E+08	4.3E+07	1.0E+08	1.6E+07	4.2	1.0	3.5	0.7
A8HVQ1	-	9.3E+08	1.7E+08	6.2E+08	1.1E+08	15.8	3.2	18.3	2.0
A8HW56	-	1.8E+08	2.1E+07	9.8E+07	3.0E+07	11.5	1.4	9.5	1.6
A8HW92	-	4.8E+07	1.7E+07	1.6E+07	1.1E+07	1.5	0.6	1.0	0.6
A8HWA8	-	3.7E+08	1.9E+07	3.1E+08	3.1E+07	7.7	0.7	6.2	1.0
A8HWI6	-	9.1E+04	8.1E+04	1.5E+04	2.1E+04	0.7	0.6	0.2	0.2
A8HWS8	-	4.9E+08	3.6E+07	3.8E+08	1.6E+07	6.5	1.1	8.3	0.6
A8HWZ8	-	1.4E+08	1.9E+07	8.0E+07	4.7E+06	1.2	0.2	1.0	0.0
A8HX04	-	1.1E+08	2.3E+07	1.8E+08	1.5E+07	4.5	0.3	4.8	0.4
A8HX52	-	1.0E+06	4.6E+05	3.5E+05	3.2E+05	0.7	0.3	0.3	0.3
A8HX54	-	1.6E+09	1.5E+08	1.6E+09	9.8E+07	5.7	0.5	5.3	0.5
A8HX70	-	2.1E+07	6.2E+06	1.5E+07	4.4E+06	0.8	0.2	0.8	0.2
A8HX89	-	8.1E+07	8.6E+06	6.0E+07	3.0E+06	1.0	0.0	1.0	0.0
A8HX98	-	0.0E+00	0.0E+00	1.1E+06	9.9E+05	0.0	0.0	0.3	0.3
A8HXA2	-	0.0E+00	0.0E+00	2.0E+06	1.3E+06	0.0	0.0	0.5	0.3
A8HXA6	-	4.5E+05	2.2E+05	6.5E+05	3.9E+05	0.7	0.3	0.7	0.3
A8HXC5	-	2.9E+07	9.2E+06	1.9E+07	6.1E+06	1.8	0.4	1.5	0.6
A8HXE1	-	9.0E+07	3.0E+06	5.3E+07	1.1E+07	1.2	0.2	2.5	0.3
A8HXG5	-	1.7E+09	8.5E+07	1.0E+09	1.7E+08	5.5	0.5	8.2	0.9
A8HXL8	-	1.6E+09	6.3E+07	1.1E+09	1.3E+08	23.7	1.7	23.8	1.5
A8HXM1	-	1.8E+08	1.1E+07	9.9E+07	2.2E+07	5.2	0.2	5.7	0.3
A8HXM5	-	1.7E+06	2.4E+06	3.6E+06	2.3E+06	0.2	0.2	0.5	0.3
A8HXR2	-	1.5E+07	7.8E+06	6.0E+06	3.7E+06	1.2	0.6	0.8	0.4
A8HXS7	-	2.0E+06	2.8E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8HXS9	-	1.3E+07	6.9E+06	2.4E+06	2.2E+06	1.0	0.5	0.3	0.3
A8HXT4	-	4.3E+07	2.0E+07	7.3E+06	1.0E+07	1.0	0.5	0.2	0.2
A8HXW8	-	2.2E+08	2.7E+07	1.3E+08	3.2E+07	4.7	0.7	4.3	1.1
A8HXX4	-	1.4E+07	4.1E+06	8.9E+06	4.0E+06	0.8	0.2	0.7	0.3
A8HXX8	-	9.0E+05	1.0E+05	7.1E+05	1.2E+05	5.2	0.4	4.3	1.0
A8HXZ5	-	6.6E+06	4.2E+06	1.6E+06	2.3E+06	0.5	0.3	0.2	0.2
A8HY43	-	1.7E+08	3.1E+07	9.4E+07	1.3E+07	1.8	0.4	1.3	0.3
A8HY75	-	5.0E+06	4.5E+06	3.1E+06	2.8E+06	0.3	0.3	0.3	0.3
A8HYA9	-	8.9E+07	7.3E+06	5.7E+07	5.8E+06	2.8	0.4	3.2	0.7
A8HYJ1	-	1.3E+07	6.0E+06	1.4E+06	1.9E+06	0.8	0.4	0.2	0.2
A8HYM2	-	1.2E+06	1.7E+06	2.3E+06	3.3E+06	0.2	0.2	0.2	0.2
A8HYQ1	-	1.2E+07	3.7E+06	1.0E+06	1.2E+06	0.8	0.2	0.7	0.5
A8HYQ6	-	3.3E+07	1.1E+07	1.0E+07	5.5E+06	1.0	0.4	0.7	0.3
A8HYR2	-	1.7E+07	5.3E+06	4.0E+06	5.6E+06	0.8	0.2	0.2	0.2
A8HYU2	-	4.2E+07	2.4E+06	2.8E+07	1.6E+06	3.5	0.3	2.3	0.5
A8HZF9	-	6.7E+07	7.4E+06	4.4E+07	1.3E+07	1.8	0.4	1.2	0.4
A8HZX1	-	2.4E+06	1.1E+06	1.9E+06	5.4E+05	0.7	0.3	0.8	0.2
A8HZZ1	-	2.5E+08	3.0E+07	1.8E+08	4.8E+07	9.8	2.1	15.7	1.5
A8I023	-	5.6E+07	4.2E+06	1.3E+07	9.3E+06	2.7	0.3	0.8	0.6

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8I093	-	3.3E+06	2.4E+06	7.8E+05	1.1E+06	0.5	0.3	0.2	0.2
A8I0C9	-	1.9E+06	2.7E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I0H0	-	2.3E+07	1.0E+07	1.1E+07	1.0E+07	0.7	0.3	0.5	0.3
A8I0I1	-	2.6E+08	2.9E+07	1.4E+08	2.0E+07	3.5	0.3	3.2	0.6
A8I0M0	-	2.4E+05	2.3E+05	8.8E+04	1.2E+05	0.3	0.3	0.2	0.2
A8I0V1	-	5.0E+06	5.0E+06	2.7E+06	3.9E+06	0.5	0.5	0.2	0.2
A8I0Y2	-	3.0E+08	7.1E+07	8.6E+07	2.8E+07	9.7	1.4	7.0	2.1
A8I0Z4	-	0.0E+00	0.0E+00	2.1E+06	2.3E+06	0.0	0.0	0.3	0.3
A8I164	-	5.6E+08	5.9E+07	3.0E+08	6.2E+07	16.8	1.3	14.0	1.9
A8I175	-	4.4E+06	6.2E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I1B8	-	2.6E+05	1.3E+05	0.0E+00	0.0E+00	0.8	0.4	0.0	0.0
A8I1M5	-	1.6E+08	1.9E+07	7.8E+07	2.0E+07	2.3	0.3	2.5	0.3
A8I1Y0	-	8.5E+05	1.3E+05	5.9E+05	2.5E+05	1.0	0.0	1.0	0.4
A8I232	-	4.6E+07	1.6E+07	4.9E+07	8.3E+06	1.8	0.6	3.5	0.3
A8I247	-	3.4E+07	1.6E+07	1.2E+07	3.5E+06	1.0	0.5	0.8	0.2
A8I252	-	0.0E+00	0.0E+00	1.1E+06	1.6E+06	0.0	0.0	0.2	0.2
A8I263	-	1.7E+08	1.9E+07	1.1E+08	2.2E+07	6.7	0.9	5.7	0.8
A8I274	-	2.9E+07	6.2E+06	2.1E+07	7.8E+06	1.5	0.3	1.3	0.5
A8I282	-	3.1E+06	4.3E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I297	-	1.4E+08	6.9E+06	1.6E+08	2.3E+07	2.3	0.3	3.8	0.8
A8I2E0	-	5.9E+07	1.6E+07	4.1E+07	8.3E+06	1.8	0.6	2.2	0.6
A8I2H7	-	0.0E+00	0.0E+00	1.0E+06	1.4E+06	0.0	0.0	0.2	0.2
A8I2S1	-	1.4E+06	2.0E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I2T0	-	5.1E+08	5.7E+07	4.6E+08	3.8E+07	4.2	0.2	4.8	0.4
A8I2Y8	-	0.0E+00	0.0E+00	3.8E+06	3.0E+06	0.0	0.0	0.5	0.3
A8I305	-	4.7E+08	3.3E+07	2.4E+08	5.1E+07	25.3	0.7	17.8	3.5
A8I311	-	3.1E+07	1.0E+07	2.3E+07	2.8E+06	0.8	0.2	1.0	0.0
A8I363	-	9.8E+06	4.3E+06	1.3E+07	2.5E+06	1.3	0.5	1.7	0.5
A8I382	-	5.7E+06	3.8E+06	4.0E+06	1.8E+06	0.5	0.3	0.7	0.3
A8I3E9	-	9.6E+05	5.3E+05	2.4E+06	2.3E+06	2.0	1.0	2.3	1.1
A8I3J5	-	1.0E+07	4.5E+06	7.6E+06	3.7E+06	0.7	0.3	0.7	0.3
A8I3Q0	-	3.3E+06	4.6E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8I3W3	-	9.7E+05	9.1E+05	5.6E+06	1.9E+06	0.3	0.3	0.8	0.2
A8I403	-	3.9E+08	1.1E+08	2.2E+08	5.7E+07	24.8	4.1	17.7	4.0
A8I495	-	2.6E+07	7.1E+06	3.2E+07	1.0E+07	2.0	0.6	3.5	1.1
A8I4A8	-	8.2E+04	1.2E+05	2.4E+05	2.9E+05	0.2	0.2	0.3	0.3
A8I4H4	-	2.4E+07	9.6E+06	8.3E+06	6.5E+06	1.8	0.7	0.7	0.3
A8I4R2	-	1.2E+07	3.7E+06	1.6E+06	2.3E+06	0.8	0.2	0.2	0.2
A8I4S9	-	6.5E+07	5.3E+06	4.3E+07	5.1E+06	2.7	0.7	2.5	0.3
A8I523	-	4.1E+06	2.3E+06	5.6E+06	2.3E+06	1.3	0.7	1.8	0.7
A8I528	-	3.7E+08	6.4E+07	2.6E+08	3.5E+07	7.3	1.6	7.0	1.7
A8I542	-	3.0E+06	2.7E+06	2.7E+06	3.7E+06	0.3	0.3	0.5	0.5
A8I554	-	5.1E+05	7.2E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I5A0	-	2.6E+08	8.1E+06	1.0E+08	4.9E+07	3.8	0.6	3.2	1.7
A8I5J8	-	1.3E+08	1.4E+07	6.2E+07	8.9E+06	4.8	0.8	6.3	0.9
A8I5K1	-	1.6E+07	6.2E+06	9.5E+06	2.9E+06	1.5	0.5	2.0	0.7
A8I5Q6	-	8.7E+06	6.0E+06	5.6E+06	3.5E+06	0.7	0.5	0.7	0.5
A8I5R9	-	4.1E+06	5.7E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I5S5	-	2.3E+06	3.3E+06	4.0E+06	2.6E+06	0.2	0.2	0.5	0.3
A8I5V6	-	8.8E+05	1.2E+06	1.1E+06	1.1E+06	0.2	0.2	0.3	0.3
A8I629	-	1.7E+07	2.1E+06	8.2E+06	1.9E+06	1.7	0.3	1.7	0.5
A8I647	-	7.2E+07	1.5E+07	3.3E+07	7.3E+06	3.8	0.8	2.3	0.5
A8I6C1	-	4.1E+08	3.3E+07	2.7E+08	2.1E+07	2.7	0.3	3.3	0.5

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8I6D1	-	1.2E+08	2.2E+07	6.2E+07	3.9E+06	3.2	0.4	2.3	0.3
A8I6R4	-	2.5E+07	1.1E+07	6.7E+06	7.0E+06	1.3	0.5	0.5	0.3
A8I6T5	-	1.0E+07	6.8E+06	3.6E+06	3.2E+06	0.5	0.3	0.3	0.3
A8I6X2	-	0.0E+00	0.0E+00	1.9E+06	2.7E+06	0.0	0.0	0.2	0.2
A8I720	-	3.5E+06	4.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I741	-	2.5E+06	2.2E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8I748	-	5.3E+07	1.0E+07	6.1E+06	4.3E+06	1.7	0.3	0.5	0.3
A8I775	-	1.1E+07	1.5E+07	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I7F2	-	5.5E+07	7.9E+06	2.8E+07	2.8E+06	1.2	0.2	1.0	0.0
A8I7N8	-	1.3E+08	3.5E+07	8.2E+07	1.8E+07	2.5	0.6	3.2	0.7
A8I7T1	-	9.9E+06	3.8E+06	3.5E+06	2.3E+06	1.0	0.4	0.5	0.3
A8I7T8	-	1.9E+08	4.3E+07	8.1E+07	1.8E+07	3.5	0.6	2.5	0.6
A8I7W0	-	4.2E+08	5.7E+07	1.9E+08	4.6E+07	7.8	1.0	7.8	1.6
A8I7X0	-	1.8E+07	5.8E+06	6.0E+06	5.3E+06	1.3	0.5	0.5	0.5
A8I826	-	4.9E+07	1.5E+07	5.6E+07	1.6E+07	2.0	0.6	2.5	0.6
A8I832	-	0.0E+00	0.0E+00	6.1E+04	8.6E+04	0.0	0.0	0.2	0.2
A8I848	-	6.2E+06	4.0E+06	7.2E+06	3.3E+06	0.5	0.3	0.7	0.3
A8I893	-	1.1E+08	4.1E+06	3.1E+07	3.9E+06	3.2	0.2	1.3	0.3
A8I8A3	-	1.5E+08	1.2E+07	1.4E+08	3.3E+07	5.8	1.0	4.7	0.7
A8I8A6	-	7.2E+07	3.4E+07	6.5E+07	2.1E+07	1.7	0.5	2.0	0.6
A8I8E9	-	1.9E+08	1.7E+08	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8I8N5	-	4.9E+06	6.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8I8Q5	-	1.6E+06	5.2E+05	2.2E+06	5.9E+05	0.8	0.2	1.0	0.0
A8I8V5	-	8.4E+06	8.3E+06	1.6E+06	2.2E+06	0.5	0.5	0.2	0.2
A8I8V6	-	1.9E+08	3.6E+07	1.2E+08	1.1E+07	4.5	1.0	4.8	0.6
A8I8X8	-	5.1E+07	1.3E+07	3.0E+07	5.8E+06	2.5	0.5	1.3	0.3
A8I8Z1	-	1.5E+07	3.1E+06	3.3E+07	1.2E+07	2.2	0.2	2.2	0.6
A8I8Z4	-	1.0E+09	1.5E+08	6.8E+08	8.9E+07	12.8	1.9	11.2	2.9
A8I901	-	5.3E+07	3.0E+07	1.8E+06	1.7E+06	1.2	0.6	0.3	0.3
A8I972	-	1.6E+08	2.3E+07	1.1E+08	1.1E+07	5.2	0.9	4.2	0.4
A8I980	-	1.5E+08	1.9E+07	1.5E+08	1.7E+07	4.5	0.6	4.5	0.6
A8I982	-	8.6E+08	8.8E+07	5.8E+08	9.3E+07	10.5	0.9	14.2	1.6
A8I9A1	-	1.6E+07	4.7E+06	3.9E+06	3.6E+06	0.8	0.2	0.3	0.3
A8I9E8	-	2.0E+07	2.5E+07	3.8E+07	3.5E+07	0.5	0.3	0.5	0.5
A8I9H5	-	2.0E+07	6.3E+06	1.9E+07	3.0E+06	1.5	0.6	1.7	0.5
A8I9I9	-	2.5E+08	4.7E+07	1.0E+08	2.0E+07	13.3	2.0	8.0	0.6
A8I9L1	-	1.0E+07	3.2E+06	2.6E+08	2.3E+08	0.8	0.2	1.2	0.7
A8I9M5	-	1.9E+08	1.7E+07	1.3E+08	1.2E+07	5.3	0.6	3.0	0.5
A8I9R1	-	1.3E+07	9.2E+06	4.5E+06	6.4E+06	0.5	0.3	0.2	0.2
A8I9S6	-	2.1E+07	9.5E+05	1.1E+07	4.6E+06	1.2	0.2	1.0	0.4
A8I9S9	-	1.3E+07	7.6E+06	3.6E+06	1.0E+06	1.3	0.7	2.2	0.6
A8I9X5	-	0.0E+00	0.0E+00	2.6E+06	1.7E+06	0.0	0.0	0.5	0.3
A8IA18	-	1.0E+09	5.9E+07	5.6E+08	4.3E+07	11.7	0.6	13.3	2.1
A8IA39	-	1.4E+09	9.6E+07	7.3E+08	1.9E+08	31.5	1.4	31.3	3.4
A8IA42	-	2.0E+07	2.8E+07	1.1E+08	4.1E+07	0.2	0.2	1.2	0.4
A8IA45	-	2.9E+08	1.5E+07	2.2E+08	1.8E+07	13.3	2.4	13.7	1.2
A8IA86	-	3.7E+08	3.1E+07	2.0E+08	4.1E+07	10.3	0.8	7.2	1.3
A8IAA8	-	4.4E+07	5.3E+06	3.5E+07	4.8E+06	1.0	0.0	1.2	0.2
A8IAA9	-	1.6E+08	1.1E+07	9.2E+07	1.1E+07	1.8	0.2	2.2	0.6
A8IAB5	-	4.7E+06	3.0E+06	3.0E+06	2.7E+06	0.5	0.3	0.3	0.3
A8IAC7	-	6.1E+06	5.5E+06	2.3E+06	3.3E+06	0.7	0.6	0.2	0.2
A8IAE5	-	0.0E+00	0.0E+00	7.5E+05	1.1E+06	0.0	0.0	0.2	0.2
A8IAK1	-	1.9E+07	1.3E+07	1.1E+07	3.9E+06	1.3	0.7	1.5	0.5

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IAK9	-	2.9E+07	7.0E+06	2.6E+07	5.9E+06	1.3	0.3	1.5	0.5
A8IAL8	-	5.7E+07	1.2E+07	4.4E+07	1.0E+07	1.7	0.5	2.2	0.2
A8IAN1	-	3.5E+09	1.3E+08	2.2E+09	3.2E+08	46.8	4.9	47.2	2.4
A8IAN4	-	0.0E+00	0.0E+00	8.0E+05	1.1E+06	0.0	0.0	0.2	0.2
A8IAT4	-	1.2E+09	5.0E+07	9.5E+08	1.0E+08	18.8	1.3	21.3	2.0
A8IAV8	-	7.8E+06	7.7E+06	1.7E+06	2.3E+06	0.5	0.5	0.2	0.2
A8IAW5	-	2.3E+08	1.1E+07	1.8E+08	1.1E+07	4.7	0.3	3.2	0.2
A8IAY6	-	1.6E+08	2.1E+07	6.8E+07	2.2E+07	5.5	1.3	3.7	1.2
A8IB70	-	2.2E+05	3.1E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IBF4	-	5.6E+07	1.9E+07	4.4E+07	5.3E+06	3.0	0.7	2.0	0.5
A8IBF6	-	5.4E+06	4.8E+06	2.6E+06	2.8E+06	0.3	0.3	0.3	0.3
A8IBG1	-	2.3E+08	1.8E+07	1.3E+08	1.6E+07	3.0	0.0	2.7	0.5
A8IBU6	-	1.7E+07	7.5E+06	1.5E+07	1.2E+06	0.7	0.3	1.0	0.0
A8IBV4	-	1.4E+07	5.0E+06	4.1E+06	1.6E+06	1.7	0.5	1.5	0.5
A8IBY2	-	2.4E+08	4.2E+07	7.0E+07	2.3E+07	6.7	1.0	3.8	0.4
A8IC90	-	6.5E+06	4.1E+06	6.0E+06	3.9E+06	0.5	0.3	0.5	0.3
A8ICB0	-	7.2E+06	6.5E+06	8.5E+06	4.5E+06	0.3	0.3	1.0	0.4
A8ICB3	-	1.8E+07	1.2E+07	1.6E+07	1.5E+07	0.7	0.5	0.3	0.3
A8ICC8	-	4.0E+08	7.5E+07	3.1E+08	6.1E+07	5.0	0.6	4.0	0.5
A8ICE4	-	1.4E+07	4.0E+06	2.0E+07	5.0E+06	2.8	0.4	3.5	1.6
A8ICG7	-	5.6E+06	2.6E+06	2.5E+06	2.4E+06	0.7	0.3	0.3	0.3
A8ICM2	-	7.2E+06	4.7E+06	1.3E+07	1.2E+06	0.8	0.6	1.3	0.3
A8ICN7	-	4.4E+07	1.2E+07	4.3E+07	1.5E+07	1.5	0.3	2.5	0.3
A8ICT4	-	9.9E+07	1.6E+07	3.5E+07	1.7E+07	1.8	0.4	1.0	0.5
A8ICV4	-	2.9E+08	7.3E+07	1.2E+08	2.8E+07	4.0	1.0	4.8	1.4
A8ID17	-	8.0E+04	1.1E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8ID63	-	5.6E+06	2.7E+06	3.1E+06	2.0E+06	0.7	0.3	0.5	0.3
A8IDN1	-	4.5E+07	5.4E+06	9.6E+06	8.6E+06	1.2	0.2	0.3	0.3
A8IDY0	-	3.2E+07	1.1E+07	1.5E+07	6.9E+06	1.8	0.4	1.2	0.7
A8IDY8	-	4.7E+07	9.9E+06	2.5E+07	9.3E+06	1.7	0.5	1.5	0.3
A8IE04	-	2.0E+05	9.0E+04	9.9E+04	6.3E+04	0.7	0.3	0.5	0.3
A8IE23	-	4.4E+08	5.9E+07	2.6E+08	3.1E+07	13.5	0.9	13.8	1.8
A8IE32	-	1.1E+08	1.6E+07	9.3E+07	6.2E+06	3.3	0.5	3.3	0.3
A8IED9	-	1.8E+07	7.3E+06	5.0E+06	4.7E+06	0.8	0.2	1.0	1.2
A8IEE5	-	2.5E+06	2.3E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IEF7	-	6.1E+06	3.9E+06	8.7E+06	1.2E+06	0.5	0.3	1.0	0.0
A8IEN5	-	5.0E+06	4.7E+06	2.4E+06	1.5E+06	0.3	0.3	0.5	0.3
A8IEW6	-	2.7E+08	2.4E+07	9.2E+07	2.9E+07	7.3	1.2	6.0	1.8
A8IF06	-	1.2E+07	3.6E+06	1.7E+06	2.5E+06	0.8	0.2	0.2	0.2
A8IF08	-	1.0E+08	1.6E+07	8.0E+07	1.2E+07	3.7	0.7	2.8	0.6
A8IFC8	-	3.0E+07	1.5E+07	3.2E+07	2.0E+07	1.2	0.4	1.5	0.7
A8IFH4	-	1.0E+07	6.5E+06	5.9E+06	5.4E+06	0.5	0.3	0.3	0.3
A8IFJ3	-	3.4E+07	4.4E+06	3.0E+07	2.1E+06	1.0	0.0	1.0	0.0
A8IFN3	-	2.5E+07	1.1E+07	3.1E+07	1.1E+07	1.3	0.5	1.7	0.5
A8IFW1	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8IGI0	-	2.9E+06	4.1E+06	1.9E+06	2.6E+06	0.2	0.2	0.2	0.2
A8IGN6	-	2.5E+06	2.3E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IGU4	-	1.5E+08	2.4E+07	1.4E+08	6.7E+06	1.8	0.2	2.0	0.0
A8IGY1	-	3.9E+08	8.2E+06	2.9E+08	8.4E+07	9.3	1.6	10.7	2.3
A8IH03	-	3.3E+08	5.4E+07	1.7E+08	3.7E+07	14.0	1.0	13.7	0.6
A8IH07	-	2.9E+07	4.3E+06	6.9E+06	3.4E+06	1.7	0.3	0.7	0.3
A8IH77	-	1.3E+09	4.8E+07	1.1E+09	1.0E+08	21.8	5.8	27.2	1.6
A8IHB3	-	1.2E+08	1.3E+08	1.5E+07	5.8E+06	1.8	0.4	1.0	0.4

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IHH9	-	4.9E+07	2.1E+07	2.2E+06	2.0E+06	1.5	0.6	0.3	0.3
A8IHL6	-	6.2E+06	6.0E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IHM8	-	1.7E+06	2.4E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IHX1	-	7.2E+07	6.4E+06	3.8E+07	1.1E+07	1.8	0.4	0.8	0.2
A8II42	-	1.6E+08	3.6E+07	3.7E+07	1.7E+07	4.8	1.1	2.0	0.7
A8II71	-	1.8E+07	8.0E+06	1.9E+07	5.3E+06	0.7	0.3	1.2	0.4
A8IIA8	-	5.3E+07	4.8E+07	3.8E+07	2.8E+07	0.5	0.5	0.5	0.3
A8IID0	-	3.0E+08	6.5E+07	2.4E+08	2.9E+07	4.5	0.6	5.2	0.6
A8IIK8	-	5.1E+06	1.6E+06	2.0E+06	9.3E+05	1.8	0.4	1.7	0.5
A8IIL1	-	6.9E+07	2.1E+07	4.1E+07	1.2E+07	1.0	0.4	0.8	0.2
A8IIL8	-	4.5E+05	4.8E+05	8.1E+05	5.2E+05	0.3	0.3	0.5	0.3
A8IIP9	-	1.6E+07	6.0E+06	5.9E+06	1.9E+06	1.3	0.5	0.8	0.2
A8IJ19	-	1.3E+08	1.2E+07	5.9E+07	1.4E+07	5.0	0.5	4.2	1.1
A8IJ60	-	5.9E+07	5.8E+06	3.6E+07	9.1E+06	2.5	0.3	2.3	0.6
A8IJE5	-	3.0E+06	4.7E+05	1.8E+06	2.3E+05	1.3	0.3	1.2	0.2
A8IJK9	-	3.2E+06	2.8E+06	8.2E+05	1.2E+06	0.3	0.3	0.2	0.2
A8IJL8	-	5.3E+06	3.4E+06	5.5E+06	4.0E+06	0.5	0.3	0.5	0.3
A8IJR6	-	2.3E+08	9.1E+07	3.2E+08	4.2E+07	5.2	0.8	6.3	0.5
A8IJY7	-	2.1E+08	3.3E+07	5.7E+07	1.1E+07	5.5	0.8	3.3	0.5
A8IK58	-	4.2E+07	1.6E+07	1.9E+07	1.3E+07	1.0	0.4	0.5	0.3
A8IK69	-	2.8E+06	3.6E+06	3.2E+06	1.4E+06	0.5	0.5	1.0	0.4
A8IK91	-	2.5E+07	4.5E+06	1.3E+07	2.9E+06	1.7	0.3	1.2	0.2
A8IK94	-	0.0E+00	0.0E+00	2.1E+05	2.9E+05	0.0	0.0	0.2	0.2
A8IK98	-	3.6E+07	4.5E+06	1.6E+07	7.3E+06	2.2	0.6	1.5	0.7
A8IKA4	-	1.4E+06	2.5E+05	2.2E+05	2.0E+05	1.7	0.3	0.3	0.3
A8IKB4	-	4.6E+07	3.3E+07	2.6E+07	2.5E+07	1.0	0.5	0.8	0.2
A8IKC8	-	7.8E+08	2.7E+07	7.8E+08	1.6E+08	40.5	1.8	35.3	3.1
A8IKD6	-	1.6E+07	2.5E+06	4.0E+06	1.8E+06	1.0	0.0	0.7	0.3
A8IKE6	-	2.6E+08	1.5E+07	1.9E+08	2.0E+07	11.8	1.3	11.2	1.8
A8IKH9	-	3.6E+06	2.3E+06	5.1E+06	1.7E+06	0.8	0.6	1.0	0.4
A8IKI6	-	7.4E+04	1.0E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IKI9	-	9.2E+08	7.4E+07	5.5E+08	4.6E+07	15.2	0.8	12.3	1.7
A8IKK7	-	3.8E+06	5.3E+06	8.6E+06	5.5E+06	0.2	0.2	0.5	0.3
A8IKN8	-	1.2E+07	8.1E+06	5.7E+06	3.8E+06	0.7	0.5	0.5	0.3
A8IKP1	-	3.1E+08	7.2E+07	2.3E+08	6.4E+07	14.7	0.8	12.3	0.9
A8IKQ0	-	6.6E+08	2.8E+07	4.4E+08	7.2E+07	10.3	2.1	12.0	2.1
A8IKQ6	-	2.7E+07	2.4E+07	3.8E+07	1.6E+07	1.0	0.9	1.0	0.4
A8IKQ9	-	4.2E+06	3.9E+06	1.8E+06	2.5E+06	0.3	0.3	0.2	0.2
A8IKW6	-	4.4E+08	7.0E+07	2.7E+08	2.1E+07	15.0	4.4	9.2	2.3
A8IKZ2	-	4.5E+08	9.2E+07	1.6E+08	2.4E+07	4.3	0.8	5.0	0.5
A8IL08	-	4.5E+08	1.1E+08	1.5E+08	4.4E+07	9.2	1.0	7.8	0.9
A8IL29	-	2.5E+08	1.3E+07	1.9E+08	2.7E+07	4.5	0.3	4.8	0.4
A8IL32	-	9.7E+08	1.2E+08	8.4E+08	4.9E+07	6.0	1.6	4.5	0.8
A8IL88	-	3.0E+06	2.8E+06	1.0E+06	1.4E+06	0.3	0.3	0.2	0.2
A8ILC2	-	1.8E+08	8.2E+07	4.1E+08	2.7E+07	0.7	0.3	1.2	0.2
A8ILF4	-	0.0E+00	0.0E+00	9.5E+06	7.4E+06	0.0	0.0	0.7	0.5
A8ILG8	-	5.5E+07	3.8E+07	1.6E+07	1.4E+07	0.8	0.6	0.5	0.3
A8ILJ9	-	2.2E+06	3.2E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8ILK1	-	3.8E+06	5.4E+06	4.1E+06	3.8E+06	0.2	0.2	0.3	0.3
A8ILL5	-	1.9E+07	2.3E+06	1.9E+07	3.6E+06	3.2	0.7	3.8	0.4
A8ILN4	-	4.0E+08	4.9E+07	1.6E+08	5.5E+07	20.5	1.3	15.3	4.4
A8ILP2	-	2.8E+06	3.4E+06	2.1E+06	1.9E+06	0.3	0.3	0.3	0.3
A8ILT1	-	1.3E+06	1.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8ILX2	-	3.5E+06	3.2E+06	2.1E+06	1.9E+06	0.3	0.3	0.3	0.3
A8IM71	-	9.9E+07	9.1E+06	6.8E+07	1.1E+07	4.0	0.5	2.2	0.6
A8IML7	-	3.1E+07	9.3E+06	1.1E+07	4.1E+06	1.7	0.3	0.8	0.2
A8IMN3	-	9.4E+06	1.3E+07	2.9E+07	5.8E+06	0.3	0.5	1.5	0.3
A8IMN5	-	3.3E+08	1.4E+07	1.6E+08	2.7E+07	7.3	1.1	7.7	1.1
A8IMP0	-	0.0E+00	0.0E+00	6.9E+05	9.8E+05	0.0	0.0	0.2	0.2
A8IMU3	-	1.6E+05	7.4E+04	1.9E+04	2.7E+04	0.8	0.4	0.2	0.2
A8IMU7	-	2.4E+05	2.2E+05	5.0E+04	7.1E+04	0.3	0.3	0.2	0.2
A8IMY5	-	7.0E+07	3.2E+06	6.4E+07	1.4E+07	2.2	0.2	2.2	0.6
A8IMZ5	-	6.6E+08	8.8E+07	5.0E+08	5.6E+07	16.3	2.5	15.3	2.3
A8IN92	-	1.3E+07	5.3E+06	4.0E+06	3.6E+06	1.7	0.7	0.7	0.7
A8IN95	-	3.4E+06	3.0E+06	7.6E+06	3.7E+06	0.3	0.3	0.7	0.3
A8INI3	-	1.2E+06	1.0E+06	3.3E+06	4.7E+06	0.3	0.3	0.2	0.2
A8INR7	-	2.1E+07	4.2E+06	6.0E+06	3.8E+06	1.3	0.3	0.7	0.3
A8INS3	-	3.0E+07	2.4E+07	3.2E+05	4.5E+05	1.7	0.9	0.2	0.2
A8IP17	-	1.1E+09	1.2E+08	6.8E+08	7.8E+07	30.5	3.9	33.2	1.8
A8IP37	-	1.2E+08	1.5E+07	6.5E+07	1.2E+07	2.8	0.7	2.3	0.5
A8IP53	-	2.0E+07	1.9E+07	4.4E+07	1.3E+07	0.3	0.3	0.8	0.2
A8IPC9	-	1.2E+07	6.9E+06	9.4E+06	7.4E+06	0.8	0.4	0.7	0.3
A8IPE0	-	0.0E+00	0.0E+00	3.5E+06	3.3E+06	0.0	0.0	0.3	0.3
A8IPH5	-	2.7E+06	1.7E+06	2.0E+06	1.3E+06	0.5	0.3	0.5	0.3
A8IPI7	-	6.2E+07	2.5E+07	8.4E+06	4.3E+06	1.7	0.7	1.2	0.7
A8IPQ9	-	1.6E+08	2.8E+07	1.0E+08	2.5E+07	4.5	0.9	2.2	0.6
A8IQ09	-	2.2E+08	2.5E+07	1.1E+08	2.4E+07	5.0	0.5	4.5	0.9
A8IQ62	-	1.2E+05	1.7E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IQA9	-	2.7E+06	3.8E+06	2.9E+06	2.6E+06	0.2	0.2	0.3	0.3
A8IQB8	-	1.8E+08	2.4E+07	9.6E+07	1.2E+07	6.8	1.3	7.2	3.2
A8IQC1	-	1.1E+09	3.9E+07	9.3E+08	8.4E+07	5.0	0.7	4.8	1.5
A8IQC5	-	5.8E+07	2.2E+07	1.7E+07	9.9E+06	2.5	0.8	1.2	0.7
A8IQE3	-	1.3E+09	1.7E+08	8.2E+08	1.4E+08	16.2	2.2	18.7	1.9
A8IQL7	-	0.0E+00	0.0E+00	3.0E+06	2.7E+06	0.0	0.0	0.3	0.3
A8IQM0	-	9.2E+06	1.5E+06	1.5E+07	5.2E+06	1.0	0.0	2.2	0.2
A8IQQ1	-	9.2E+06	4.3E+06	3.9E+06	2.5E+06	0.7	0.3	0.5	0.3
A8IQS8	-	1.4E+07	7.2E+06	8.5E+06	2.8E+06	0.8	0.2	0.8	0.2
A8IQU3	-	1.9E+09	1.3E+08	1.1E+09	1.5E+08	54.2	7.1	48.8	7.5
A8IQV5	-	4.4E+07	4.4E+06	4.3E+07	1.4E+07	2.2	0.2	2.7	0.9
A8IQZ0	-	1.1E+07	5.9E+06	3.0E+06	2.7E+06	0.7	0.3	0.3	0.3
A8IR49	-	8.1E+06	5.2E+06	1.2E+07	4.8E+06	0.5	0.3	1.0	0.4
A8IR69	-	1.9E+09	7.5E+07	1.4E+09	7.2E+07	58.7	4.9	64.8	3.1
A8IR74	-	5.0E+07	1.3E+07	2.4E+07	3.0E+06	1.8	0.2	1.3	0.3
A8IR79	-	0.0E+00	0.0E+00	5.4E+04	7.7E+04	0.0	0.0	0.2	0.2
A8IR87	-	3.8E+06	5.3E+06	5.3E+06	4.7E+06	0.2	0.2	0.3	0.3
A8IR93	-	0.0E+00	0.0E+00	9.5E+05	1.3E+06	0.0	0.0	0.2	0.2
A8IR98	-	3.9E+07	7.3E+06	1.8E+07	8.6E+06	5.0	0.6	3.8	1.1
A8IRB3	-	9.8E+06	6.1E+05	1.2E+07	2.0E+06	1.2	0.2	1.0	0.0
A8IRE2	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8IRG9	-	2.3E+06	7.2E+05	1.8E+06	8.1E+05	0.8	0.2	0.7	0.3
A8IRH4	-	1.2E+06	1.7E+06	2.1E+06	1.9E+06	0.2	0.2	0.3	0.3
A8IRI7	-	0.0E+00	0.0E+00	2.1E+06	1.9E+06	0.0	0.0	0.3	0.3
A8IRK4	-	2.9E+09	1.7E+08	2.3E+09	7.5E+07	97.2	16.5	95.3	8.1
A8IRK5	-	1.4E+07	1.2E+07	2.6E+07	3.5E+06	0.3	0.3	1.0	0.0
A8IRQ1	-	3.7E+08	4.2E+07	1.3E+08	3.6E+07	5.3	0.5	2.3	0.7
A8IRR9	-	8.4E+05	7.1E+05	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IRT2	-	3.9E+07	1.6E+07	1.0E+07	1.5E+07	1.2	0.4	0.3	0.5
A8IRT4	-	2.5E+07	1.2E+07	1.5E+07	9.5E+06	0.8	0.4	0.5	0.3
A8IRT7	-	1.9E+07	1.6E+06	1.6E+06	2.2E+06	1.0	0.0	0.2	0.2
A8IRT8	-	1.1E+07	3.3E+06	6.1E+05	8.6E+05	1.0	0.4	0.2	0.2
A8IRU6	-	6.4E+08	8.7E+07	4.3E+08	8.1E+07	14.5	1.5	16.3	1.1
A8IRV0	-	3.9E+07	2.4E+07	1.3E+07	6.4E+06	1.8	0.8	0.8	0.4
A8IRV6	-	1.2E+08	1.5E+07	6.2E+07	1.3E+07	4.5	0.6	3.3	0.6
A8IRW5	-	1.6E+07	5.0E+06	2.4E+06	2.2E+06	1.0	0.4	0.3	0.3
A8IRZ2	-	7.3E+06	3.5E+06	2.3E+06	2.2E+06	1.2	0.6	0.5	0.5
A8IS44	-	1.6E+07	2.5E+06	1.9E+07	3.6E+06	1.2	0.2	1.3	0.3
A8IS47	-	1.0E+07	3.5E+06	8.4E+06	6.3E+06	0.8	0.2	0.7	0.5
A8IS72	-	5.6E+04	5.1E+04	6.5E+05	9.2E+05	0.3	0.3	0.2	0.2
A8IS88	-	3.1E+06	2.1E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8ISA8	-	8.2E+05	7.4E+05	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8ISE3	-	1.3E+08	2.9E+07	7.5E+07	1.0E+07	1.3	0.3	2.2	0.4
A8ISE5	-	3.7E+06	5.2E+06	8.2E+06	5.4E+06	0.2	0.2	0.5	0.3
A8ISG3	-	2.4E+07	1.1E+07	2.1E+06	1.4E+06	0.8	0.2	0.5	0.3
A8ISJ9	-	0.0E+00	0.0E+00	7.4E+06	4.7E+06	0.0	0.0	0.7	0.5
A8ISK2	-	0.0E+00	0.0E+00	4.3E+05	6.1E+05	0.0	0.0	0.2	0.2
A8ISN6	-	4.4E+07	2.1E+07	7.5E+06	8.4E+06	1.3	0.7	0.5	0.5
A8ISS0	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8IST1	-	4.9E+06	3.8E+06	2.0E+07	2.6E+07	0.5	0.3	0.5	0.3
A8IST3	-	2.9E+06	4.0E+06	5.1E+06	3.8E+06	0.2	0.2	0.7	0.5
A8ISV8	-	5.9E+07	1.9E+07	1.8E+07	1.2E+07	0.8	0.2	0.5	0.3
A8ISZ1	-	5.5E+08	5.6E+07	2.1E+08	3.8E+07	21.2	1.8	15.2	0.8
A8IT23	-	2.4E+07	1.8E+07	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8IT25	-	2.2E+08	2.2E+07	3.3E+07	9.5E+06	3.3	0.5	1.7	0.6
A8ITF0	-	0.0E+00	0.0E+00	1.2E+07	1.0E+07	0.0	0.0	0.5	0.5
A8ITF3	-	1.7E+07	5.3E+06	2.0E+06	2.8E+06	1.0	0.4	0.2	0.2
A8ITH8	-	1.2E+09	1.0E+08	7.3E+08	9.1E+07	15.7	1.3	14.2	1.4
A8ITJ3	-	1.6E+06	2.2E+06	1.2E+07	7.5E+06	0.2	0.2	0.5	0.3
A8ITK6	-	2.8E+08	1.4E+07	1.3E+08	4.6E+07	3.3	0.5	3.2	1.0
A8ITL0	-	5.1E+08	4.1E+07	3.5E+08	1.3E+07	14.2	1.7	15.8	1.5
A8ITL8	-	7.5E+06	4.9E+06	3.1E+06	2.8E+06	0.5	0.3	0.5	0.5
A8ITS8	-	7.6E+08	1.0E+08	7.1E+08	4.2E+07	11.7	1.3	12.7	0.9
A8ITU2	-	4.9E+08	5.5E+07	2.9E+08	3.3E+07	12.3	0.6	12.7	2.0
A8ITU6	-	1.6E+08	1.9E+07	1.3E+08	2.2E+07	2.3	0.5	2.3	0.7
A8ITV3	-	8.5E+08	1.4E+08	2.2E+08	8.3E+07	83.2	10.5	78.0	6.9
A8ITX0	-	3.1E+07	9.8E+06	2.3E+07	3.7E+06	1.3	0.3	2.0	0.4
A8ITX5	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8ITZ0	-	0.0E+00	0.0E+00	9.6E+05	1.4E+06	0.0	0.0	0.2	0.2
A8IU48	-	5.6E+06	5.0E+06	9.3E+05	1.3E+06	0.3	0.3	0.2	0.2
A8IU68	-	3.4E+07	1.2E+07	2.4E+06	3.4E+06	1.3	0.3	0.2	0.2
A8IUB8	-	1.5E+07	8.9E+06	2.3E+06	2.1E+06	1.0	0.5	0.3	0.3
A8IUC3	-	1.1E+08	7.2E+06	5.1E+07	1.1E+07	3.0	0.5	3.5	0.5
A8IUL5	-	2.5E+07	1.1E+07	8.4E+06	5.6E+06	1.2	0.4	0.5	0.3
A8IUL6	-	0.0E+00	0.0E+00	1.6E+06	2.1E+06	0.0	0.0	0.5	0.3
A8IUL9	-	1.3E+08	1.5E+07	1.0E+08	1.8E+07	2.7	0.3	2.7	0.6
A8IUR4	-	1.4E+07	1.5E+06	2.0E+07	6.2E+06	2.8	0.2	3.3	0.5
A8IUU2	-	8.7E+06	2.7E+06	7.1E+06	4.3E+06	0.8	0.2	0.8	0.4
A8IUU3	-	4.2E+08	3.2E+07	3.6E+08	2.3E+07	34.7	3.6	42.8	5.8
A8IUV7	-	1.9E+09	1.4E+08	1.1E+09	1.2E+08	35.2	2.5	33.8	3.5
A8IV18	-	7.0E+06	6.3E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IV20	-	1.8E+07	1.1E+07	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8IV36	-	4.5E+06	3.4E+06	1.8E+06	2.5E+06	0.5	0.3	0.5	0.7
A8IV40	-	1.3E+07	3.9E+06	1.5E+07	3.8E+06	2.3	0.6	2.5	0.6
A8IV67	-	3.9E+07	1.4E+07	3.7E+07	1.2E+07	1.5	0.5	1.8	0.6
A8IVB9	-	1.4E+05	6.5E+04	3.4E+04	3.3E+04	0.8	0.4	0.3	0.3
A8IVE2	-	1.7E+09	1.3E+08	1.4E+09	9.9E+07	30.3	1.1	35.2	4.8
A8IVF0	-	1.5E+07	1.6E+07	7.3E+05	1.0E+06	0.7	0.5	0.2	0.2
A8IVG0	-	3.2E+08	4.4E+07	1.5E+08	2.6E+07	8.7	1.2	5.7	1.7
A8IVG9	-	2.7E+06	2.5E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IVH2	-	1.1E+07	3.2E+06	5.6E+06	3.6E+06	1.0	0.4	0.5	0.3
A8IVK1	-	5.0E+08	2.8E+07	1.9E+08	5.5E+07	6.5	0.5	5.8	0.8
A8IVM9	-	2.9E+07	1.3E+07	1.7E+07	6.3E+06	2.0	1.2	2.0	0.5
A8IVN2	-	0.0E+00	0.0E+00	4.2E+06	3.1E+06	0.0	0.0	0.7	0.5
A8IVN6	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8IVR6	-	1.0E+08	2.0E+07	6.2E+07	1.8E+07	3.8	0.8	5.3	0.9
A8IVS6	-	0.0E+00	0.0E+00	1.5E+07	7.0E+06	0.0	0.0	0.7	0.3
A8IVU0	-	7.6E+08	4.0E+07	4.1E+08	1.5E+08	9.8	1.6	11.5	4.0
A8IVV4	-	1.0E+07	6.7E+06	3.1E+06	4.3E+06	0.5	0.3	0.2	0.2
A8IVW9	-	1.0E+07	3.0E+06	5.5E+06	3.7E+06	0.8	0.2	0.5	0.3
A8IVY4	-	4.4E+07	4.0E+07	2.0E+07	2.3E+07	0.3	0.3	0.7	0.5
A8IVZ3	-	4.0E+06	3.6E+06	7.4E+06	2.2E+06	0.3	0.3	0.8	0.2
A8IVZ9	-	1.7E+09	1.5E+08	1.4E+09	7.8E+07	43.2	5.1	44.2	5.7
A8IW20	-	9.5E+07	2.0E+07	8.2E+07	3.3E+06	1.0	0.0	1.0	0.0
A8IW34	-	4.2E+08	4.7E+07	3.3E+08	3.3E+07	10.0	0.7	12.3	1.5
A8IW39	-	9.4E+07	2.4E+07	4.2E+07	1.5E+07	2.3	0.5	3.5	0.8
A8IW44	-	1.8E+08	3.1E+07	9.7E+07	2.1E+07	2.7	0.5	1.8	0.2
A8IW49	-	3.6E+07	1.1E+07	3.4E+07	5.5E+06	1.7	0.5	2.3	0.3
A8IW72	-	2.3E+06	1.1E+06	2.0E+06	1.0E+06	0.7	0.3	0.7	0.3
A8IW84	-	5.0E+04	7.0E+04	5.4E+05	1.7E+05	0.2	0.2	1.3	0.3
A8IW99	-	3.8E+07	4.7E+06	2.2E+07	2.4E+06	1.2	0.2	1.0	0.0
A8IWA6	-	4.1E+07	5.5E+07	9.2E+04	1.3E+05	0.3	0.3	0.2	0.2
A8IWB5	-	8.1E+06	6.2E+05	5.3E+06	4.6E+05	2.3	0.3	1.5	0.3
A8IWF9	-	3.3E+06	3.0E+06	3.2E+06	2.4E+06	0.3	0.3	0.7	0.5
A8IWJ3	-	7.2E+07	4.0E+07	9.2E+07	1.6E+07	1.2	0.7	1.5	0.3
A8IWJ5	-	2.8E+08	3.5E+07	1.2E+08	3.8E+07	7.7	0.8	4.8	1.1
A8IWJ6	-	1.6E+08	6.0E+06	8.6E+07	1.9E+07	2.0	0.0	2.7	0.3
A8IWK9	-	1.7E+08	2.6E+07	1.9E+08	2.4E+07	3.5	0.8	3.2	0.2
A8IWL3	-	3.0E+07	7.1E+06	1.2E+07	2.8E+06	2.7	0.6	1.8	0.6
A8IWL4	-	9.9E+07	1.2E+07	2.2E+08	1.4E+08	2.8	0.4	2.5	0.3
A8IWP2	-	4.1E+07	1.9E+07	9.7E+07	1.8E+07	0.7	0.3	1.2	0.2
A8IWQ1	-	1.0E+07	9.6E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IWS4	-	3.3E+04	4.7E+04	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IWT7	-	1.7E+06	2.4E+06	5.1E+06	5.2E+06	0.2	0.2	0.5	0.5
A8IWU0	-	0.0E+00	0.0E+00	2.5E+06	2.3E+06	0.0	0.0	0.3	0.3
A8IWV0	-	5.0E+06	4.8E+06	0.0E+00	0.0E+00	0.5	0.5	0.0	0.0
A8IWV9	-	7.7E+06	5.1E+06	4.8E+06	6.7E+06	0.5	0.3	0.2	0.2
A8IX19	-	2.7E+07	1.3E+07	3.5E+07	1.3E+07	0.8	0.4	1.3	0.5
A8IX41	-	1.4E+07	6.3E+06	2.0E+07	2.8E+06	0.7	0.3	1.2	0.2
A8IXD1	-	1.5E+08	1.0E+07	8.9E+07	5.5E+06	4.7	0.5	2.7	0.8
A8IXE0	-	3.5E+09	2.8E+08	2.3E+09	1.2E+08	72.0	4.9	60.8	11.3
A8IXF1	-	1.0E+08	7.9E+06	7.5E+07	2.3E+07	1.0	0.0	0.8	0.2
A8IXF2	-	2.2E+07	7.6E+06	1.5E+07	1.2E+06	0.8	0.2	1.0	0.0
A8IXG3	-	9.4E+06	7.4E+06	6.4E+06	8.8E+06	1.2	0.8	0.8	0.6

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8IXH4	-	1.6E+07	7.2E+06	9.2E+06	5.9E+06	1.0	0.0	1.5	0.6
A8IXJ7	-	5.6E+05	7.9E+05	5.3E+05	7.5E+05	0.2	0.2	0.2	0.2
A8IXQ5	-	4.9E+07	1.4E+07	8.2E+06	7.8E+06	1.2	0.2	0.3	0.3
A8IXR5	-	7.5E+07	2.2E+07	4.1E+07	1.1E+07	2.3	0.6	2.5	0.6
A8IY09	-	0.0E+00	0.0E+00	8.4E+05	1.2E+06	0.0	0.0	0.2	0.2
A8IY20	-	1.3E+08	1.9E+07	2.3E+07	1.5E+07	2.5	0.3	0.8	0.6
A8IY50	-	1.3E+06	1.9E+06	2.2E+06	2.0E+06	0.2	0.2	0.3	0.3
A8IYC6	-	2.4E+05	3.4E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IYG2	-	5.9E+07	8.1E+06	2.6E+07	4.9E+06	1.7	0.3	1.7	0.5
A8IYG3	-	0.0E+00	0.0E+00	1.1E+06	1.6E+06	0.0	0.0	0.2	0.2
A8IYH1	-	3.6E+06	3.4E+06	5.0E+06	3.6E+06	0.3	0.3	0.5	0.3
A8IYJ5	-	2.1E+06	2.9E+06	1.3E+06	1.9E+06	0.2	0.2	0.2	0.2
A8IYM0	-	2.4E+07	1.2E+07	3.0E+07	1.2E+07	0.8	0.4	1.2	0.4
A8IYP3	-	2.2E+06	3.1E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IYP4	-	1.7E+09	5.0E+07	7.6E+08	1.0E+08	103.2	11.2	117.8	7.9
A8IYP5	-	3.5E+07	6.1E+06	1.3E+07	4.5E+06	1.0	0.0	0.8	0.2
A8IYR0	-	0.0E+00	0.0E+00	7.5E+06	5.0E+06	0.0	0.0	0.5	0.3
A8IYR1	-	1.7E+06	2.5E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IYR2	-	2.5E+07	4.2E+06	1.2E+07	3.6E+06	1.5	0.3	1.0	0.4
A8IYR7	-	7.5E+05	1.1E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IYS1	-	6.0E+08	3.8E+07	1.9E+08	7.2E+07	8.5	0.5	7.5	2.5
A8IYS5	-	8.2E+07	2.9E+07	7.0E+07	7.9E+06	1.8	0.8	1.7	0.5
A8IYU7	-	7.4E+06	2.2E+06	4.1E+06	1.9E+06	0.8	0.2	0.7	0.3
A8IYW8	-	1.4E+06	2.0E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8IZ29	-	1.7E+07	7.7E+06	1.5E+07	6.1E+06	0.7	0.3	1.8	0.6
A8IZ36	-	5.9E+08	5.8E+07	3.8E+08	3.6E+07	3.7	0.5	3.5	0.6
A8IZ39	-	9.7E+06	6.1E+06	1.0E+07	6.8E+05	0.7	0.3	1.0	0.0
A8IZ70	-	1.1E+07	6.7E+06	3.8E+06	5.4E+06	0.5	0.3	0.2	0.2
A8IZ72	-	1.9E+07	9.0E+06	2.0E+07	5.1E+06	0.7	0.3	1.0	0.0
A8IZ88	-	2.4E+06	3.3E+06	4.6E+07	3.7E+06	0.2	0.2	1.0	0.0
A8IZ93	-	2.3E+06	3.3E+06	2.2E+05	3.0E+05	0.2	0.2	0.2	0.2
A8IZK3	-	4.8E+08	6.0E+07	1.9E+08	2.3E+07	9.2	1.1	7.3	1.3
A8IZL1	-	1.9E+06	1.7E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8IZL5	-	0.0E+00	0.0E+00	2.8E+06	2.5E+06	0.0	0.0	0.3	0.3
A8IZR5	-	1.7E+08	2.5E+07	8.9E+07	1.6E+07	3.3	0.3	2.0	0.5
A8IZS7	-	3.9E+07	2.7E+06	4.7E+07	2.0E+07	1.2	0.2	1.8	0.7
A8IZU0	-	2.9E+08	6.7E+07	2.3E+08	1.3E+07	6.0	0.6	7.2	0.6
A8IZW5	-	3.5E+07	4.7E+06	4.5E+06	4.1E+06	1.7	0.3	0.5	0.3
A8IZW6	-	1.0E+08	1.3E+07	4.1E+07	7.9E+06	4.3	0.9	3.2	1.0
A8IZX1	-	5.5E+07	4.7E+06	3.3E+07	1.2E+07	1.0	0.0	1.2	0.4
A8IZX9	-	9.3E+07	3.7E+06	5.5E+07	2.5E+07	2.2	0.2	1.0	0.5
A8IZZ9	-	1.8E+07	9.5E+06	4.3E+07	1.1E+07	1.5	0.6	2.7	0.6
A8J014	-	1.1E+07	8.0E+06	9.8E+06	1.9E+06	0.7	0.5	1.3	0.3
A8J063	-	1.5E+07	7.6E+06	2.7E+07	7.8E+06	1.7	0.5	2.3	0.5
A8J087	-	5.8E+08	2.2E+07	3.9E+08	2.4E+07	6.5	0.8	6.3	0.7
A8J094	-	4.4E+07	5.5E+06	1.1E+07	7.1E+06	1.3	0.3	1.3	0.3
A8J0A1	-	3.6E+06	3.2E+06	6.9E+06	5.3E+06	0.3	0.3	0.5	0.3
A8J0A2	-	1.0E+06	1.5E+06	8.9E+05	1.3E+06	0.2	0.2	0.2	0.2
A8J0D6	-	2.0E+06	1.3E+06	2.2E+06	1.4E+06	0.5	0.3	0.5	0.3
A8J0E4	-	1.4E+10	5.5E+08	9.0E+09	8.5E+08	367.3	32.2	344.8	25.9
A8J0I1	-	4.8E+05	2.3E+05	1.4E+05	1.3E+05	0.7	0.3	0.3	0.3
A8J0K1	-	1.4E+07	6.3E+06	1.4E+07	4.0E+06	0.7	0.3	0.8	0.2
A8J0N1	-	2.0E+06	1.8E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8J0N6	-	2.1E+06	2.0E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J0N7	-	1.4E+09	1.1E+08	1.3E+09	1.2E+08	27.0	2.8	34.8	3.2
A8J0Q1	-	1.2E+05	7.9E+04	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J0Q2	-	1.1E+06	1.5E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J0Q8	-	0.0E+00	0.0E+00	3.8E+06	3.7E+06	0.0	0.0	0.5	0.5
A8J0R2	-	1.5E+07	8.3E+06	2.8E+07	4.9E+06	1.2	0.6	2.0	0.4
A8J0R6	-	1.1E+08	2.3E+07	6.4E+07	1.3E+07	5.0	1.0	4.2	0.4
A8J0U6	-	1.4E+05	2.0E+05	3.3E+05	3.0E+05	0.2	0.2	0.3	0.3
A8J0W9	-	4.0E+08	1.4E+07	2.5E+08	2.5E+07	6.0	0.4	4.2	0.7
A8J0Z0	-	1.5E+07	1.3E+06	6.3E+06	1.8E+06	1.0	0.0	0.8	0.2
A8J103	-	5.5E+06	3.5E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J104	-	1.2E+08	2.1E+07	1.8E+08	1.8E+07	3.8	0.7	4.7	0.5
A8J119	-	2.2E+05	3.1E+05	2.3E+05	3.2E+05	0.2	0.2	0.2	0.2
A8J122	-	1.3E+06	1.9E+06	5.3E+06	2.4E+06	0.2	0.2	0.7	0.3
A8J133	-	2.1E+08	2.4E+07	4.9E+07	3.2E+07	3.2	0.4	1.2	0.7
A8J137	-	3.0E+05	9.7E+04	1.8E+06	2.4E+06	0.8	0.2	0.8	0.4
A8J146	-	2.2E+08	1.8E+07	1.5E+08	1.7E+07	5.8	0.2	4.3	0.6
A8J152	-	2.7E+07	1.1E+07	3.6E+06	3.6E+06	1.3	0.3	0.3	0.3
A8J167	-	5.6E+04	3.8E+04	0.0E+00	0.0E+00	1.0	0.7	0.0	0.0
A8J173	-	4.7E+08	4.8E+07	2.4E+08	2.7E+07	7.3	0.5	5.5	0.5
A8J175	-	2.7E+07	7.8E+06	2.8E+07	2.8E+06	0.8	0.2	1.0	0.0
A8J195	-	1.7E+08	1.1E+07	1.0E+08	1.4E+07	2.5	0.5	2.2	0.4
A8J1A3	-	9.9E+07	1.5E+07	2.5E+07	4.8E+06	4.0	1.0	2.2	0.4
A8J1B6	-	2.3E+09	1.3E+08	1.8E+09	1.6E+08	93.8	5.5	103.5	10.7
A8J1B8	-	2.3E+04	3.3E+04	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J1C1	-	5.5E+07	4.9E+06	3.0E+07	5.6E+06	1.0	0.0	1.2	0.2
A8J1C2	-	3.7E+07	1.4E+07	4.3E+06	3.4E+06	1.8	0.4	0.5	0.3
A8J1F7	-	2.1E+07	4.9E+06	1.5E+07	2.6E+06	1.3	0.3	1.2	0.2
A8J1G8	-	2.6E+09	1.2E+08	2.0E+09	9.3E+07	14.2	0.8	16.0	1.6
A8J1H4	-	1.5E+07	1.5E+07	1.4E+06	2.0E+06	0.5	0.5	0.3	0.5
A8J1J8	-	7.9E+07	2.3E+07	8.1E+07	8.5E+06	0.8	0.2	1.0	0.0
A8J1M5	-	5.6E+05	7.9E+05	1.5E+06	1.5E+06	0.2	0.2	0.3	0.3
A8J1M9	-	1.3E+08	1.8E+07	9.2E+07	7.7E+06	3.0	0.9	2.7	0.3
A8J1P1	-	4.2E+06	6.0E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J1S7	-	2.7E+08	1.6E+07	1.8E+08	2.5E+07	7.3	0.7	7.2	1.1
A8J1S8	-	3.3E+08	1.9E+08	5.7E+07	2.3E+07	2.2	0.4	1.3	0.5
A8J1T4	-	2.6E+08	3.7E+07	1.0E+08	2.0E+07	6.7	0.9	4.5	1.1
A8J1U1	-	3.7E+08	3.8E+07	1.9E+08	3.0E+07	9.8	1.0	7.5	1.1
A8J1V0	-	2.5E+06	3.5E+06	4.5E+06	4.0E+06	0.2	0.2	0.3	0.3
A8J1X0	-	7.8E+07	5.1E+07	2.7E+07	2.1E+06	1.2	0.7	1.0	0.0
A8J1X8	-	8.9E+07	3.4E+07	8.1E+07	1.0E+07	3.5	1.1	4.3	0.9
A8J214	-	6.2E+07	1.3E+07	5.0E+07	2.9E+07	3.7	0.6	3.0	0.9
A8J216	-	9.0E+06	4.1E+06	1.2E+06	1.8E+06	0.7	0.3	0.2	0.2
A8J239	-	9.5E+08	2.7E+08	5.1E+08	1.2E+08	9.8	0.8	7.2	1.0
A8J264	-	3.4E+09	2.4E+08	2.7E+09	3.3E+08	231.2	18.4	213.2	28.1
A8J266	-	0.0E+00	0.0E+00	9.1E+05	1.3E+06	0.0	0.0	0.2	0.2
A8J282	-	1.6E+08	8.0E+06	9.5E+07	1.4E+07	3.5	0.3	2.8	0.4
A8J287	-	1.8E+09	3.8E+08	1.2E+09	1.6E+08	43.8	2.7	33.8	3.1
A8J290	-	6.4E+07	8.9E+06	2.2E+07	8.8E+06	2.7	0.8	3.0	1.3
A8J2A5	-	5.7E+07	5.8E+06	7.4E+06	6.6E+06	1.0	0.0	0.3	0.3
A8J2E5	-	9.2E+05	1.3E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J2G4	-	3.8E+08	1.5E+07	2.6E+08	2.8E+07	6.7	0.5	7.3	0.8
A8J2I5	-	7.5E+08	1.6E+08	4.4E+08	9.2E+07	28.8	2.5	23.8	3.3

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8J2J0	-	1.4E+08	2.0E+08	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J2L0	-	3.6E+08	1.3E+07	2.2E+08	1.6E+07	5.7	0.6	5.7	0.6
A8J2L7	-	5.1E+08	4.9E+07	4.0E+08	1.8E+07	2.0	0.0	1.8	0.2
A8J2N7	-	4.8E+05	4.3E+05	3.2E+05	2.9E+05	0.3	0.3	0.3	0.3
A8J2S3	-	8.0E+06	4.2E+06	2.1E+07	1.5E+06	0.7	0.3	1.7	0.3
A8J2W0	-	7.3E+07	8.1E+06	4.8E+07	3.3E+06	2.0	0.5	1.2	0.2
A8J2W4	-	2.0E+06	2.8E+06	4.9E+06	2.7E+06	0.2	0.2	0.7	0.3
A8J2X6	-	1.2E+08	4.0E+07	1.2E+08	5.3E+07	3.2	0.8	5.0	1.3
A8J2X8	-	1.1E+07	3.5E+06	2.3E+06	2.0E+06	0.8	0.2	0.3	0.3
A8J2Z6	-	1.8E+07	2.6E+07	1.0E+08	1.1E+07	0.2	0.2	1.0	0.0
A8J328	-	4.6E+07	3.8E+07	2.3E+07	6.1E+06	1.3	0.5	1.8	0.4
A8J355	-	8.3E+07	3.5E+07	3.9E+07	1.0E+07	3.3	1.1	3.0	0.4
A8J383	-	1.1E+06	1.1E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J387	-	4.7E+07	1.1E+07	1.5E+07	2.8E+06	2.0	0.5	1.0	0.0
A8J3A4	-	4.0E+07	5.3E+06	1.6E+07	5.3E+06	1.8	0.2	1.5	0.3
A8J3B2	-	1.9E+07	1.5E+06	1.1E+07	1.2E+06	1.0	0.0	1.0	0.0
A8J3C4	-	0.0E+00	0.0E+00	7.2E+05	1.0E+06	0.0	0.0	0.2	0.2
A8J3D3	-	8.4E+06	7.9E+06	1.5E+07	1.0E+06	0.5	0.5	1.0	0.0
A8J3F7	-	2.2E+09	1.4E+08	1.5E+09	1.1E+08	50.0	3.1	47.0	2.2
A8J3F8	-	1.3E+08	2.3E+07	9.3E+07	3.4E+07	5.3	1.4	6.2	1.6
A8J3K3	-	2.1E+06	3.0E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J3L2	-	2.6E+07	1.2E+07	1.5E+07	9.9E+06	0.7	0.3	0.5	0.3
A8J3L3	-	2.9E+06	2.6E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J3L5	-	5.8E+08	8.6E+07	3.2E+08	2.7E+07	9.2	0.7	7.3	3.3
A8J3L9	-	5.7E+08	3.1E+07	6.2E+08	4.3E+07	11.2	1.2	13.3	2.0
A8J3P5	-	6.5E+07	2.9E+07	4.7E+07	1.8E+07	2.7	0.9	2.0	0.8
A8J3R3	-	2.2E+06	2.0E+06	9.6E+06	7.8E+06	0.3	0.3	1.2	0.4
A8J3U9	-	1.2E+08	3.9E+07	3.3E+07	1.6E+07	6.8	0.8	5.2	1.0
A8J3Y3	-	1.4E+06	1.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J3Z3	-	7.0E+06	1.0E+07	2.7E+07	1.0E+07	0.2	0.2	1.2	0.4
A8J429	-	1.9E+05	2.6E+05	8.4E+06	1.2E+07	0.2	0.2	0.2	0.2
A8J431	-	4.4E+07	1.2E+07	4.1E+06	3.7E+06	1.8	0.4	0.3	0.3
A8J434	-	1.7E+06	2.4E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J449	-	4.5E+06	4.2E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J496	-	8.8E+06	2.0E+06	2.7E+06	1.3E+06	2.5	0.3	1.0	0.5
A8J4D3	-	1.0E+08	1.4E+07	4.2E+07	1.0E+07	2.7	0.5	1.8	0.4
A8J4E8	-	2.9E+07	1.8E+06	2.5E+07	3.0E+06	1.0	0.0	1.3	0.3
A8J4I2	-	1.4E+08	3.6E+07	8.6E+07	1.0E+07	1.5	0.3	1.5	0.7
A8J4I9	-	5.1E+06	4.6E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J4L1	-	5.8E+07	7.8E+06	5.2E+07	8.9E+06	2.5	0.5	2.8	0.4
A8J4M8	-	4.4E+07	1.1E+07	7.9E+07	1.5E+07	1.5	0.5	2.7	0.5
A8J4N7	-	1.8E+07	1.0E+07	2.2E+07	3.1E+06	1.3	0.6	1.7	0.3
A8J4Q7	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8J4R9	-	2.4E+08	5.0E+07	1.2E+08	2.3E+07	6.3	0.8	5.7	0.9
A8J4S1	-	3.4E+09	3.3E+08	1.9E+09	3.3E+08	81.8	7.2	71.7	5.8
A8J4X1	-	1.3E+06	9.4E+05	2.6E+06	6.7E+05	0.5	0.3	1.2	0.2
A8J4Z4	-	2.4E+08	2.3E+07	1.4E+08	1.5E+07	9.0	2.1	14.2	1.2
A8J500	-	1.7E+07	7.7E+06	5.8E+06	4.5E+06	1.0	0.4	0.7	0.5
A8J503	-	1.6E+09	6.6E+07	6.0E+08	2.7E+08	3.2	0.7	2.5	0.5
A8J506	-	6.4E+08	8.5E+07	3.4E+08	8.7E+06	17.0	1.3	17.8	1.4
A8J513	-	9.4E+07	1.6E+07	6.9E+07	1.3E+07	2.5	0.5	4.0	0.5
A8J524	-	1.7E+06	2.4E+06	1.4E+07	8.6E+06	0.2	0.2	0.8	0.4
A8J525	-	8.8E+07	2.7E+06	5.8E+07	3.9E+06	2.0	0.0	2.2	0.4

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8J538	-	7.1E+07	1.3E+07	5.5E+07	3.4E+06	2.2	0.4	2.0	0.0
A8J567	-	1.4E+09	9.8E+07	7.9E+08	1.0E+08	52.3	13.5	58.8	12.5
A8J597	-	3.2E+08	1.6E+07	2.0E+08	1.2E+07	2.8	1.0	4.7	0.9
A8J599	-	1.2E+07	6.3E+06	1.5E+07	4.3E+06	1.0	0.5	1.7	0.5
A8J5K0	-	1.8E+07	1.2E+07	1.2E+07	7.9E+06	0.5	0.3	0.5	0.3
A8J5K4	-	4.7E+08	3.6E+07	3.9E+08	3.8E+07	10.8	1.4	10.2	1.4
A8J5K6	-	0.0E+00	0.0E+00	3.0E+06	2.8E+06	0.0	0.0	0.3	0.3
A8J5N1	-	9.6E+06	6.1E+06	8.3E+06	4.4E+06	1.0	0.6	1.5	0.5
A8J5N6	-	3.8E+06	3.5E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J5P0	-	2.8E+07	3.9E+07	4.0E+07	3.4E+07	0.7	0.5	1.0	0.0
A8J5P7	-	6.8E+08	1.9E+07	5.5E+08	3.8E+07	8.2	1.0	10.3	0.6
A8J5Q8	-	1.3E+07	5.0E+06	6.3E+06	4.5E+06	1.0	0.4	0.7	0.5
A8J5R2	-	2.5E+06	3.5E+06	1.4E+06	2.0E+06	0.2	0.2	0.2	0.2
A8J5T0	-	2.6E+08	7.2E+06	1.7E+08	3.7E+07	5.8	1.0	6.7	1.9
A8J5T5	-	5.6E+06	6.4E+06	0.0E+00	0.0E+00	0.5	0.5	0.0	0.0
A8J5Y7	-	6.4E+07	5.3E+06	3.1E+07	5.8E+06	1.8	0.2	3.0	0.5
A8J5Z0	-	2.6E+09	9.4E+07	1.8E+09	1.6E+08	29.8	2.0	26.5	1.5
A8J605	-	7.4E+06	7.4E+06	1.2E+07	8.6E+05	0.3	0.3	1.3	0.3
A8J633	-	7.1E+07	2.4E+07	2.3E+07	1.4E+07	1.7	0.5	1.8	0.4
A8J635	-	1.5E+08	3.6E+07	1.8E+08	6.7E+06	1.5	0.3	1.7	0.3
A8J664	-	3.1E+06	3.5E+06	2.0E+06	2.8E+06	0.7	0.5	0.2	0.2
A8J6A8	-	1.5E+07	4.3E+06	1.4E+07	6.4E+06	0.8	0.2	0.7	0.3
A8J6C7	-	4.1E+08	4.7E+07	3.4E+08	6.5E+07	12.3	2.3	16.5	2.5
A8J6D5	-	5.1E+05	7.1E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J6D8	-	1.8E+06	1.1E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J6E2	-	2.5E+04	3.6E+04	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J6F2	-	1.0E+08	3.3E+07	4.3E+07	2.3E+07	1.8	0.4	1.0	0.4
A8J6G8	-	4.1E+06	2.6E+06	1.4E+06	1.9E+06	0.7	0.5	0.2	0.2
A8J6H7	-	5.2E+07	6.5E+06	6.9E+07	4.4E+06	1.3	0.3	2.0	0.0
A8J6J6	-	4.3E+08	4.2E+07	3.6E+08	4.3E+07	39.0	1.7	40.2	5.1
A8J6K9	-	6.1E+07	1.7E+07	1.6E+07	1.2E+07	2.0	0.4	0.5	0.3
A8J6M9	-	4.0E+06	2.6E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J6P2	-	1.7E+06	2.4E+06	1.6E+06	1.5E+06	0.2	0.2	0.3	0.3
A8J6R7	-	5.7E+06	5.1E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J6R9	-	2.1E+06	1.9E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J6S7	-	2.0E+06	3.1E+05	1.1E+06	1.4E+05	1.3	0.3	1.7	0.3
A8J6V1	-	3.3E+07	6.4E+06	1.4E+07	5.4E+06	1.8	0.4	1.3	0.5
A8J6Y3	-	7.8E+07	1.8E+07	4.2E+07	8.1E+06	2.3	0.6	2.2	0.2
A8J709	-	4.9E+07	1.6E+07	1.3E+07	2.6E+06	2.0	0.5	1.8	0.2
A8J724	-	1.1E+07	6.4E+06	4.4E+06	4.0E+06	1.0	0.5	0.3	0.3
A8J742	-	2.9E+07	6.5E+06	3.0E+07	1.2E+07	1.5	0.3	1.5	0.5
A8J746	-	2.2E+06	2.2E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8J768	-	1.7E+09	2.0E+08	1.5E+09	2.1E+08	10.2	1.4	8.0	1.8
A8J785	-	1.2E+08	8.3E+07	1.5E+07	7.7E+06	2.2	0.8	1.3	0.9
A8J7C8	-	9.5E+07	1.8E+07	5.5E+07	2.2E+07	3.5	0.9	2.8	1.0
A8J7F6	-	8.4E+08	4.3E+07	3.8E+08	4.2E+07	16.0	0.8	12.8	1.6
A8J7F8	-	3.0E+07	1.2E+07	2.1E+07	1.0E+07	1.3	0.5	0.8	0.4
A8J7F9	-	1.3E+08	6.1E+07	3.0E+08	9.7E+07	1.8	0.2	3.8	0.9
A8J7H3	-	8.8E+08	3.6E+07	3.6E+08	6.7E+07	14.5	1.5	9.2	1.8
A8J7H8	-	9.5E+07	1.8E+07	3.2E+07	1.3E+07	2.0	0.4	2.2	0.8
A8J7I3	-	8.3E+07	1.0E+07	2.8E+07	1.8E+07	1.0	0.0	0.5	0.3
A8J7I9	-	2.6E+07	3.3E+06	1.1E+07	3.8E+06	2.2	0.4	3.0	1.3
A8J7J2	-	6.6E+06	4.3E+06	7.7E+05	1.1E+06	0.5	0.3	0.2	0.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8J7J5	-	1.3E+07	5.9E+06	2.4E+07	2.0E+06	0.7	0.3	1.0	0.0
A8J7L5	-	1.1E+07	7.0E+06	1.3E+07	6.2E+06	0.5	0.3	0.7	0.3
A8J7L6	-	2.2E+06	3.1E+06	6.3E+06	3.0E+06	0.2	0.2	0.7	0.3
A8J7M3	-	5.2E+07	2.4E+07	3.0E+07	5.8E+06	2.0	0.4	1.3	0.3
A8J7M9	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8J7T7	-	3.6E+07	1.4E+07	3.5E+07	1.5E+07	1.7	0.8	2.0	0.8
A8J7X9	-	2.5E+08	3.2E+07	1.6E+08	1.7E+07	6.7	1.8	9.0	1.0
A8J7Z6	-	1.2E+07	8.0E+05	3.1E+06	2.8E+06	1.0	0.0	0.3	0.3
A8J807	-	9.0E+05	8.2E+05	4.6E+05	6.5E+05	0.3	0.3	0.2	0.2
A8J825	-	2.4E+07	7.0E+06	2.3E+07	7.2E+06	0.8	0.2	0.8	0.2
A8J827	-	1.8E+07	6.8E+06	2.7E+07	1.8E+07	1.2	0.4	1.2	0.2
A8J853	-	1.6E+07	1.8E+06	1.3E+07	7.7E+06	1.2	0.2	1.2	0.6
A8J8A6	-	3.8E+06	2.4E+06	9.1E+05	1.3E+06	0.5	0.3	0.2	0.2
A8J8B2	-	4.0E+08	6.9E+07	1.4E+08	3.9E+07	6.2	1.1	2.8	0.7
A8J8B3	-	5.2E+07	1.1E+07	4.5E+07	1.1E+07	1.7	0.6	1.2	0.2
A8J8D7	-	3.4E+06	4.8E+06	6.5E+06	6.8E+06	0.2	0.2	0.3	0.3
A8J8F3	-	3.7E+06	2.4E+06	4.9E+06	3.9E+06	0.5	0.3	0.7	0.5
A8J8I4	-	2.2E+07	1.0E+07	2.9E+07	2.6E+06	1.0	0.5	1.3	0.3
A8J8I8	-	3.8E+06	1.2E+06	9.7E+06	4.0E+06	3.7	0.5	2.8	0.7
A8J8J8	-	5.3E+06	3.4E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8J8L3	-	1.0E+07	4.8E+06	1.7E+06	2.5E+06	0.7	0.3	0.2	0.2
A8J8M3	-	1.0E+06	1.5E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J8M9	-	1.1E+09	3.6E+07	8.5E+08	1.2E+08	7.5	0.5	10.5	0.5
A8J8P4	-	2.6E+08	1.5E+07	2.0E+08	3.8E+07	2.5	0.3	3.0	1.2
A8J8S2	-	2.5E+07	1.4E+07	2.2E+07	1.3E+07	1.2	0.4	1.2	0.4
A8J8U1	-	2.7E+07	4.0E+06	1.4E+07	8.8E+06	1.2	0.2	0.7	0.5
A8J8V6	-	2.4E+08	1.1E+08	1.3E+08	3.8E+07	6.5	0.9	5.7	1.1
A8J8W6	-	1.8E+07	1.2E+07	3.6E+07	2.6E+06	0.5	0.3	1.2	0.2
A8J8W9	-	7.4E+06	3.4E+06	5.0E+06	3.3E+06	0.7	0.3	0.5	0.3
A8J8Y1	-	1.6E+09	1.2E+08	8.9E+08	1.0E+08	52.0	5.7	49.5	6.4
A8J8Z2	-	7.6E+08	8.3E+07	6.4E+08	6.6E+07	14.2	2.2	19.5	2.4
A8J906	-	5.3E+07	1.1E+07	8.7E+06	1.0E+06	2.5	0.3	1.5	0.3
A8J914	-	3.6E+08	9.4E+06	2.1E+08	3.9E+07	9.0	0.4	6.5	0.9
A8J927	-	3.3E+06	1.1E+06	5.1E+06	1.5E+06	0.8	0.2	0.8	0.2
A8J933	-	7.2E+06	5.6E+06	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0
A8J940	-	5.8E+05	8.2E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8J942	-	9.7E+07	5.5E+06	1.2E+07	1.1E+07	2.3	0.5	0.5	0.5
A8J951	-	8.8E+08	3.2E+07	6.2E+08	6.6E+07	26.2	0.8	24.2	3.2
A8J955	-	2.7E+06	3.8E+06	1.7E+06	2.4E+06	0.2	0.2	0.2	0.2
A8J989	-	9.3E+06	9.7E+06	1.4E+07	7.9E+06	1.2	0.6	1.2	0.6
A8J9A9	-	3.7E+07	1.9E+07	3.5E+07	2.5E+06	1.3	0.7	1.2	0.2
A8J9C6	-	3.4E+07	3.2E+06	1.3E+07	6.7E+06	2.0	0.0	1.3	0.3
A8J9D9	-	5.1E+07	1.7E+07	1.3E+08	4.2E+06	0.8	0.2	1.0	0.0
A8J9E6	-	4.0E+07	3.2E+06	3.0E+07	9.1E+06	1.0	0.0	1.0	0.4
A8J9E9	-	2.0E+07	9.3E+06	2.7E+07	2.2E+06	0.7	0.3	1.0	0.0
A8J9F3	-	2.6E+07	7.1E+06	2.8E+06	3.0E+06	2.3	0.5	0.5	0.5
A8J9F6	-	8.4E+07	1.6E+07	5.2E+07	4.6E+06	1.7	0.3	1.7	0.3
A8J9H2	-	8.0E+06	3.7E+06	2.1E+06	2.2E+06	0.7	0.3	0.3	0.3
A8J9H8	-	1.0E+09	2.0E+08	3.6E+08	7.8E+07	13.0	1.9	11.7	2.5
A8J9K0	-	8.5E+07	7.6E+07	6.0E+07	8.5E+07	0.3	0.3	0.3	0.3
A8J9K1	-	2.2E+06	2.0E+06	3.7E+06	3.3E+06	0.3	0.3	0.3	0.3
A8J9S7	-	1.0E+09	8.1E+07	6.2E+08	4.7E+07	28.8	2.5	29.5	1.8
A8J9T5	-	2.4E+08	3.0E+07	1.2E+08	1.2E+07	8.8	1.0	8.0	1.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8J9W0	-	3.0E+08	8.3E+07	1.1E+08	4.9E+07	7.5	1.3	5.3	0.6
A8J9X1	-	4.6E+07	1.2E+07	3.2E+07	7.6E+06	1.5	0.3	1.7	0.3
A8J9Y1	-	2.3E+08	8.3E+07	2.1E+08	1.3E+07	2.5	0.7	2.5	0.5
A8JA41	-	1.4E+07	1.3E+07	4.3E+06	6.0E+06	0.5	0.5	0.2	0.2
A8JA68	-	1.4E+07	1.0E+07	2.7E+06	3.8E+06	0.7	0.5	0.2	0.2
A8JAG1	-	5.5E+08	6.6E+07	2.9E+08	5.1E+07	10.0	0.8	7.0	0.9
A8JAH1	-	1.8E+07	1.7E+07	0.0E+00	0.0E+00	0.5	0.5	0.0	0.0
A8JAK5	-	4.4E+06	3.4E+06	5.8E+06	2.6E+06	0.7	0.5	1.2	0.4
A8JAL6	-	3.7E+08	3.6E+07	1.7E+08	3.0E+07	4.5	0.3	3.3	0.3
A8JAP7	-	6.8E+07	1.3E+07	6.2E+07	2.4E+07	2.3	0.3	4.0	0.4
A8JAR0	-	2.3E+06	3.3E+06	1.7E+06	2.3E+06	0.2	0.2	0.7	0.5
A8JAU8	-	3.9E+06	5.5E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JAV1	-	1.3E+09	6.9E+07	9.5E+08	6.9E+07	16.7	1.2	17.7	2.4
A8JAV8	-	6.6E+06	4.3E+06	2.0E+06	2.9E+06	0.5	0.3	0.2	0.2
A8JAX1	-	7.7E+07	2.0E+07	4.0E+07	5.7E+06	2.3	0.5	1.3	0.3
A8JB06	-	1.2E+07	1.1E+07	1.9E+07	5.8E+06	0.3	0.3	1.0	0.4
A8JB47	-	1.0E+07	8.7E+06	5.2E+06	4.7E+06	0.5	0.3	0.3	0.3
A8JB55	-	2.5E+06	2.3E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JBA7	-	0.0E+00	0.0E+00	1.7E+06	2.4E+06	0.0	0.0	0.2	0.2
A8JBB3	-	6.0E+07	4.7E+06	2.3E+07	1.1E+07	2.2	0.2	1.0	0.5
A8JBB4	-	3.5E+07	1.6E+07	1.9E+07	9.8E+06	2.0	0.6	1.7	0.7
A8JBC2	-	4.3E+05	6.1E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JBC6	-	7.6E+07	5.7E+06	8.6E+07	6.1E+06	1.3	0.5	1.0	0.0
A8JBD0	-	1.2E+06	1.7E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JBF2	-	4.8E+07	8.8E+06	8.1E+07	1.5E+07	2.5	0.5	2.2	0.7
A8JBG0	-	1.3E+07	1.2E+07	1.6E+07	7.8E+06	0.5	0.5	0.8	0.4
A8JBJ3	-	0.0E+00	0.0E+00	1.6E+06	2.3E+06	0.0	0.0	0.2	0.2
A8JBL3	-	4.6E+07	2.1E+07	4.5E+06	6.3E+06	1.0	0.6	0.2	0.2
A8JBL6	-	1.0E+08	2.3E+07	1.1E+08	2.1E+07	3.5	0.8	6.7	0.8
A8JBL7	-	9.5E+06	6.1E+06	2.0E+07	2.0E+06	0.5	0.3	1.2	0.2
A8JBM7	-	7.6E+06	4.2E+06	1.1E+07	9.7E+06	1.0	0.4	1.0	0.9
A8JBN7	-	1.6E+08	8.7E+06	8.7E+07	1.2E+07	3.3	0.3	2.5	0.5
A8JBN8	-	1.5E+07	2.4E+06	1.6E+07	4.8E+06	1.2	0.2	0.8	0.2
A8JBQ5	-	3.1E+07	1.7E+07	3.6E+07	2.3E+06	0.7	0.3	1.0	0.0
A8JBT7	-	1.2E+07	3.4E+06	2.5E+07	3.7E+06	0.8	0.2	1.5	0.3
A8JBU2	-	2.6E+06	1.8E+06	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8JBU3	-	7.5E+06	3.4E+06	4.1E+06	2.7E+06	0.7	0.3	0.5	0.3
A8JBX2	-	1.9E+06	5.9E+05	1.8E+05	2.6E+05	0.8	0.2	0.2	0.2
A8JBX4	-	6.7E+05	9.5E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JBY4	-	2.1E+05	1.9E+05	1.2E+05	8.0E+04	0.7	0.7	0.7	0.5
A8JBZ2	-	1.0E+07	4.8E+06	1.0E+07	4.6E+06	0.7	0.3	0.7	0.3
A8JBZ7	-	1.5E+07	4.4E+06	1.8E+07	1.5E+06	0.8	0.2	1.7	0.3
A8JC04	-	4.8E+09	2.3E+08	3.7E+09	1.7E+08	115.0	13.7	126.8	13.3
A8JC14	-	5.9E+07	1.2E+07	1.7E+07	1.1E+07	1.2	0.2	0.8	0.6
A8JC18	-	2.1E+06	2.9E+06	1.0E+07	9.6E+05	0.2	0.2	1.0	0.0
A8JC21	-	1.4E+08	1.0E+07	6.5E+07	2.4E+07	4.2	0.2	2.5	0.5
A8JC30	-	2.8E+07	1.4E+07	9.5E+06	8.5E+06	1.0	0.5	0.3	0.3
A8JC42	-	1.3E+08	2.5E+07	9.7E+07	1.5E+07	3.0	0.7	2.8	0.7
A8JC46	-	0.0E+00	0.0E+00	2.1E+06	2.3E+06	0.0	0.0	0.3	0.3
A8JC47	-	5.1E+06	4.8E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JC51	-	2.6E+07	1.5E+06	3.7E+06	2.7E+06	1.2	0.2	0.5	0.3
A8JC54	-	2.3E+04	3.3E+04	6.7E+04	6.0E+04	0.2	0.2	0.3	0.3
A8JCA4	-	0.0E+00	0.0E+00	1.1E+06	1.5E+06	0.0	0.0	0.2	0.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8JCA8	-	2.1E+07	7.3E+06	9.0E+06	4.2E+06	1.2	0.4	0.7	0.3
A8JCC0	-	7.1E+07	3.9E+06	6.4E+07	5.4E+06	1.5	0.3	1.7	0.3
A8JCC6	-	1.5E+06	1.3E+06	0.0E+00	0.0E+00	0.5	0.5	0.0	0.0
A8JCE1	-	5.9E+08	7.4E+07	2.8E+08	4.2E+07	13.0	1.6	17.2	3.1
A8JCE9	-	3.9E+08	9.0E+07	2.9E+08	9.1E+07	7.8	1.6	7.3	3.0
A8JCF6	-	3.5E+07	6.6E+06	5.7E+07	1.2E+07	2.2	0.6	2.7	0.6
A8JCG7	-	0.0E+00	0.0E+00	1.2E+05	1.8E+05	0.0	0.0	0.2	0.2
A8JCH8	-	3.3E+06	1.5E+06	6.5E+06	7.9E+05	0.7	0.3	1.0	0.0
A8JCJ9	-	9.0E+06	8.4E+06	3.8E+06	3.9E+06	0.3	0.3	0.5	0.5
A8JCK3	-	1.2E+06	1.7E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JCM2	-	3.0E+06	1.9E+06	6.9E+06	6.7E+05	0.5	0.3	1.2	0.2
A8JCP1	-	1.3E+07	1.2E+07	5.7E+06	8.0E+06	0.3	0.3	0.2	0.2
A8JCP8	-	9.0E+06	5.9E+05	4.8E+06	1.4E+06	1.0	0.0	0.8	0.2
A8JCP9	-	9.2E+07	4.2E+07	6.5E+07	2.6E+07	0.8	0.2	1.8	0.6
A8JCQ8	-	1.8E+08	2.6E+07	1.1E+08	1.9E+07	16.3	2.5	13.7	2.7
A8JCR1	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8JCR6	-	3.4E+08	3.5E+07	1.8E+08	2.7E+07	7.8	0.4	6.5	0.8
A8JCT1	-	0.0E+00	0.0E+00	8.3E+05	1.2E+06	0.0	0.0	0.2	0.2
A8JCV8	-	3.9E+08	3.8E+07	3.3E+08	9.0E+07	4.5	0.3	4.2	0.6
A8JCW1	-	1.2E+07	4.4E+06	1.1E+07	3.4E+06	1.0	0.4	1.2	0.2
A8JCW5	-	4.4E+07	9.2E+06	2.5E+07	3.9E+06	3.2	0.7	1.2	0.2
A8JCW7	-	4.8E+07	9.5E+06	1.1E+07	9.1E+06	1.5	0.7	0.7	0.5
A8JD31	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8JD33	-	6.9E+07	1.1E+07	6.0E+07	2.2E+07	1.7	0.3	1.2	0.2
A8JD45	-	8.4E+07	1.7E+07	5.6E+07	1.4E+07	2.5	0.3	2.2	0.4
A8JD56	-	6.5E+07	2.2E+07	5.5E+07	1.8E+07	2.8	0.7	2.8	0.7
A8JD99	-	1.8E+06	2.5E+06	5.1E+06	3.3E+06	0.2	0.2	0.5	0.3
A8JDA2	-	1.0E+08	1.2E+07	5.6E+07	4.3E+06	2.0	0.5	1.2	0.2
A8JDA7	-	3.5E+08	2.2E+07	1.4E+08	2.4E+07	6.5	0.3	6.8	1.0
A8JDC0	-	3.4E+06	6.8E+05	3.5E+07	1.4E+07	1.7	0.3	3.5	0.6
A8JDE6	-	4.1E+07	2.2E+07	2.6E+07	6.0E+06	2.5	1.1	3.0	1.1
A8JDG4	-	4.4E+07	5.4E+06	4.3E+07	6.9E+06	1.5	0.3	1.8	0.4
A8JDJ8	-	8.0E+07	4.5E+07	3.8E+07	2.1E+07	1.3	0.3	1.3	0.3
A8JDK2	-	6.4E+07	1.1E+07	5.8E+07	7.8E+06	2.8	0.8	3.0	0.5
A8JDL5	-	2.3E+08	4.7E+07	1.8E+08	1.2E+07	11.2	0.8	6.8	1.6
A8JDL8	-	3.2E+08	6.4E+07	1.7E+08	3.8E+07	6.3	0.7	3.7	1.1
A8JDM1	-	1.5E+07	7.2E+06	8.1E+06	6.1E+06	1.3	0.6	0.7	0.5
A8JDM6	-	2.6E+06	2.3E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JDN2	-	8.8E+07	3.3E+07	9.3E+07	1.0E+07	1.7	0.5	2.2	0.2
A8JDN4	-	2.7E+08	4.6E+07	9.7E+07	2.7E+07	5.2	0.2	2.8	0.6
A8JDN8	-	2.0E+08	1.9E+07	1.6E+08	3.6E+07	2.8	0.7	2.2	0.8
A8JDP0	-	0.0E+00	0.0E+00	1.3E+07	1.4E+07	0.0	0.0	0.3	0.3
A8JDP4	-	1.1E+09	1.7E+08	5.2E+08	6.0E+07	12.8	3.0	18.0	4.5
A8JDP6	-	1.9E+08	5.8E+07	1.7E+08	1.0E+07	3.3	0.5	3.0	0.0
A8JDQ7	-	4.9E+07	2.1E+07	5.1E+05	7.2E+05	1.0	0.4	0.2	0.2
A8JDV2	-	6.0E+07	1.2E+07	2.1E+07	7.7E+06	5.7	1.4	4.5	0.9
A8JDV9	-	2.3E+08	2.7E+07	1.1E+08	1.4E+07	9.0	0.7	7.0	2.4
A8JDW2	-	8.6E+07	8.5E+06	3.8E+07	1.1E+07	2.3	0.3	2.8	0.4
A8JDY3	-	3.9E+05	5.8E+04	2.3E+05	3.9E+04	1.8	0.2	2.0	0.5
A8JE04	-	2.4E+07	1.0E+07	3.7E+07	5.2E+06	1.0	0.4	1.8	0.2
A8JE05	-	4.9E+06	3.1E+06	1.5E+06	2.1E+06	0.5	0.3	0.2	0.2
A8JE06	-	2.8E+07	1.5E+06	2.6E+07	4.6E+06	1.0	0.0	1.5	0.5
A8JE07	-	3.9E+08	1.4E+08	2.8E+08	4.0E+07	11.0	3.7	14.8	3.4

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8JE34	-	6.4E+07	1.4E+07	5.9E+07	2.3E+07	1.5	0.3	2.0	0.7
A8JE35	-	4.8E+08	4.0E+07	3.3E+08	2.6E+07	13.3	0.9	12.2	1.2
A8JE40	-	1.5E+07	7.9E+06	0.0E+00	0.0E+00	0.7	0.3	0.0	0.0
A8JE81	-	1.2E+05	1.7E+05	1.7E+05	2.4E+05	0.2	0.2	0.2	0.2
A8JE83	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
A8JE91	-	1.6E+08	6.9E+06	9.2E+07	1.1E+07	4.5	0.6	4.5	0.6
A8JE93	-	6.8E+07	6.1E+07	9.2E+07	5.9E+07	0.3	0.3	0.8	0.6
A8JE98	-	2.0E+08	1.8E+07	1.0E+08	1.3E+07	3.8	0.6	3.0	1.0
A8JEF7	-	5.3E+08	3.5E+07	2.9E+08	7.5E+07	8.7	1.1	5.0	0.4
A8JEG6	-	7.1E+07	1.7E+07	1.7E+07	1.3E+07	2.2	0.4	0.7	0.3
A8JEI4	-	0.0E+00	0.0E+00	2.1E+06	1.9E+06	0.0	0.0	0.3	0.3
A8JEM6	-	0.0E+00	0.0E+00	9.1E+06	8.2E+06	0.0	0.0	0.3	0.3
A8JEP1	-	4.1E+07	5.6E+06	8.2E+06	5.4E+06	1.8	0.4	0.8	0.6
A8JEP9	-	7.3E+07	8.8E+06	6.5E+07	1.3E+07	2.8	0.2	2.8	0.6
A8JES5	-	2.7E+07	8.8E+06	1.8E+07	8.4E+06	0.8	0.2	0.7	0.3
A8JET7	-	9.0E+06	8.4E+06	5.8E+06	5.3E+06	0.3	0.3	0.3	0.3
A8JEU0	-	0.0E+00	0.0E+00	7.4E+05	1.0E+06	0.0	0.0	0.2	0.2
A8JEU4	-	1.7E+09	1.0E+08	1.0E+09	1.3E+08	26.7	0.6	28.2	3.3
A8JEV6	-	8.9E+06	4.3E+06	6.1E+06	2.7E+06	0.7	0.3	0.7	0.3
A8JEY5	-	6.2E+06	2.9E+06	6.5E+06	3.0E+06	0.7	0.3	0.7	0.3
A8JF05	-	3.4E+08	4.2E+07	4.0E+08	2.7E+07	3.5	0.6	4.5	0.5
A8JF15	-	1.2E+08	1.2E+07	7.3E+07	2.0E+07	4.0	0.4	4.0	0.4
A8JF18	-	4.7E+08	3.9E+07	5.5E+08	2.4E+07	5.8	0.7	10.2	1.6
A8JF35	-	7.6E+07	9.0E+06	3.5E+07	8.4E+06	4.5	0.9	2.8	0.6
A8JF47	-	2.3E+08	4.8E+07	2.2E+08	9.2E+06	6.7	0.8	6.8	0.8
A8JF49	-	2.8E+07	6.7E+06	5.1E+07	4.1E+06	1.8	0.4	2.5	0.3
A8JF66	-	5.6E+06	3.1E+06	5.0E+06	1.6E+06	4.7	2.7	6.5	2.9
A8JF87	-	1.1E+07	8.3E+06	1.6E+07	5.2E+06	0.7	0.5	1.0	0.4
A8JF96	-	6.5E+05	4.1E+05	1.8E+05	2.5E+05	0.8	0.6	0.2	0.2
A8JFA2	-	6.7E+06	6.0E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JFA9	-	4.8E+05	2.2E+05	4.3E+05	2.0E+05	0.7	0.3	0.7	0.3
A8JFB1	-	4.2E+08	3.7E+07	2.2E+08	5.2E+07	7.3	0.5	5.2	0.4
A8JFD6	-	1.0E+07	9.1E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JFE5	-	1.0E+06	1.4E+06	1.2E+06	1.6E+06	0.2	0.2	0.2	0.2
A8JFI7	-	1.8E+09	5.9E+07	1.6E+09	9.9E+07	23.3	1.4	24.3	1.5
A8JFJ1	-	6.7E+07	1.7E+07	3.7E+07	2.2E+07	1.3	0.5	1.2	0.4
A8JFJ2	-	6.5E+08	4.7E+07	3.6E+08	4.5E+07	11.0	1.1	7.8	1.0
A8JFJ4	-	3.7E+06	5.3E+06	8.2E+06	5.2E+06	0.2	0.2	0.5	0.3
A8JFK4	-	5.4E+08	4.4E+07	4.1E+08	3.3E+07	8.5	0.6	11.2	1.4
A8JFK6	-	7.7E+07	1.7E+07	3.1E+07	1.2E+07	1.7	0.3	1.2	0.4
A8JFM9	-	1.7E+06	2.4E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JFN4	-	3.3E+07	1.2E+07	3.7E+07	1.1E+06	1.5	0.3	2.0	0.0
A8JFN8	-	7.2E+06	4.6E+06	9.8E+06	3.1E+06	0.8	0.6	1.0	0.4
A8JFQ7	-	6.9E+07	7.6E+06	8.0E+07	1.9E+07	6.5	0.3	6.7	0.5
A8JFR4	-	2.4E+07	7.6E+06	2.7E+07	4.4E+06	2.0	0.6	2.8	0.6
A8JFT3	-	7.5E+06	6.8E+06	2.1E+07	6.7E+06	0.3	0.3	1.0	0.4
A8JFW4	-	9.0E+07	2.7E+07	3.0E+07	1.1E+07	3.5	0.9	1.3	0.6
A8JFW5	-	1.2E+07	6.9E+06	4.6E+06	3.2E+06	1.0	0.6	0.7	0.5
A8JFX7	-	2.3E+05	4.2E+04	9.4E+04	8.4E+04	1.0	0.0	0.3	0.3
A8JFZ0	-	9.0E+07	2.0E+07	5.2E+07	1.3E+07	2.8	0.2	3.3	0.6
A8JFZ2	-	3.8E+07	1.7E+07	6.0E+06	4.9E+06	1.7	0.8	0.8	0.4
A8JFZ4	-	2.8E+07	1.3E+07	2.1E+07	7.4E+06	0.7	0.3	0.8	0.2
A8JFZ7	-	2.7E+08	1.5E+07	2.1E+08	9.3E+06	4.5	0.5	3.8	0.6

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8JG03	-	2.9E+08	1.8E+07	1.5E+08	2.1E+07	8.5	0.3	5.5	1.1
A8JG04	-	1.6E+07	2.0E+06	3.6E+06	2.3E+06	1.2	0.2	0.5	0.3
A8JG06	-	0.0E+00	0.0E+00	3.8E+05	5.3E+05	0.0	0.0	0.2	0.2
A8JG07	-	2.2E+08	4.3E+07	1.0E+08	1.1E+07	6.8	1.6	8.2	1.1
A8JG49	-	0.0E+00	0.0E+00	1.9E+06	1.9E+06	0.0	0.0	0.3	0.3
A8JG55	-	1.9E+07	4.4E+06	5.3E+06	2.4E+06	1.2	0.2	0.7	0.3
A8JG56	-	1.2E+08	2.8E+07	4.8E+07	1.9E+07	3.8	0.6	4.3	0.5
A8JG58	-	7.0E+08	7.1E+07	3.3E+08	4.3E+07	16.0	2.7	12.5	2.1
A8JG70	-	6.9E+07	7.9E+06	1.4E+07	9.7E+06	2.0	0.0	0.7	0.5
A8JG73	-	2.7E+08	3.2E+07	1.7E+08	3.3E+07	3.2	0.2	2.5	0.5
A8JGB0	-	6.8E+06	3.8E+06	9.9E+05	1.4E+06	0.8	0.4	0.2	0.2
A8JGB1	-	4.1E+07	1.9E+07	5.9E+06	8.4E+06	0.7	0.3	0.2	0.2
A8JGB3	-	0.0E+00	0.0E+00	1.4E+06	2.0E+06	0.0	0.0	0.2	0.2
A8JGC8	-	0.0E+00	0.0E+00	6.1E+06	5.5E+06	0.0	0.0	0.3	0.3
A8JGD1	-	4.5E+07	5.7E+06	8.0E+07	7.4E+06	2.2	0.4	3.5	0.6
A8JGF8	-	1.3E+09	1.1E+08	5.1E+08	1.4E+08	25.2	3.4	24.0	2.2
A8JGI8	-	4.5E+07	7.4E+06	2.1E+07	6.9E+06	1.7	0.3	1.2	0.2
A8JGJ6	-	4.4E+07	9.5E+06	1.9E+07	9.2E+06	1.5	0.5	1.0	0.5
A8JGK1	-	8.8E+08	4.6E+07	3.8E+08	7.0E+07	7.2	0.8	5.7	0.7
A8JGK5	-	7.0E+07	1.0E+07	2.0E+07	9.8E+06	1.8	0.2	0.8	0.4
A8JGL0	-	4.2E+06	3.8E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
A8JGP7	-	1.1E+07	5.6E+06	1.6E+06	2.2E+06	0.7	0.3	0.2	0.2
A8JGR1	-	8.5E+07	2.5E+07	8.0E+07	4.3E+06	3.3	0.9	1.7	0.3
A8JGS2	-	2.8E+08	2.7E+07	2.1E+08	1.2E+07	4.3	0.9	5.0	1.0
A8JGS3	-	2.5E+07	1.6E+07	0.0E+00	0.0E+00	0.5	0.3	0.0	0.0
A8JGS4	-	2.6E+06	2.3E+06	4.4E+06	2.1E+06	0.3	0.3	0.7	0.3
A8JGS8	-	1.7E+07	5.1E+06	1.4E+07	1.0E+07	0.8	0.2	0.8	0.6
A8JGT1	-	7.5E+06	3.4E+06	2.4E+06	2.1E+06	0.7	0.3	0.3	0.3
A8JGU7	-	1.6E+07	1.1E+07	2.1E+07	1.9E+06	0.5	0.3	1.0	0.0
A8JGV6	-	6.6E+08	1.8E+08	2.6E+08	4.8E+07	6.8	1.6	5.2	0.8
A8JGW5	-	1.6E+07	1.1E+06	1.2E+07	1.1E+06	1.0	0.0	1.0	0.0
A8JGW7	-	3.1E+06	2.8E+06	6.6E+06	7.9E+06	0.3	0.3	0.5	0.5
A8JGX0	-	2.8E+07	1.5E+06	1.8E+07	1.9E+06	1.0	0.0	1.0	0.0
A8JGX5	-	5.3E+07	3.5E+06	4.2E+07	4.8E+06	1.2	0.2	1.2	0.2
A8JGY3	-	1.3E+08	2.7E+07	3.1E+07	9.2E+06	2.7	0.3	1.0	0.4
A8JGY8	-	5.6E+06	3.3E+06	4.7E+06	5.2E+05	1.2	0.6	1.0	0.0
A8JH01	-	1.3E+07	6.8E+06	1.4E+07	4.3E+06	0.7	0.3	0.8	0.2
A8JH10	-	3.8E+06	1.7E+06	8.7E+05	1.2E+06	1.2	0.6	0.5	0.7
A8JH12	-	3.7E+07	2.9E+07	8.7E+07	3.0E+07	1.8	0.4	2.5	0.5
A8JH37	-	3.5E+09	4.5E+08	2.0E+09	2.4E+08	79.3	13.3	73.7	6.1
A8JH39	-	2.5E+06	3.6E+06	7.2E+06	6.4E+06	0.2	0.2	0.3	0.3
A8JH47	-	5.5E+05	7.8E+05	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JH48	-	2.4E+07	1.4E+07	8.5E+06	8.7E+05	1.3	0.7	1.2	0.2
A8JH52	-	2.4E+07	1.1E+07	2.8E+06	3.9E+06	1.2	0.4	0.2	0.2
A8JH58	-	6.6E+07	1.9E+07	6.7E+07	1.6E+07	1.8	0.4	2.8	0.4
A8JH60	-	6.1E+08	5.3E+07	3.5E+08	8.1E+07	12.0	1.9	8.0	1.6
A8JH65	-	7.5E+07	2.5E+07	2.3E+07	2.0E+07	1.0	0.4	0.3	0.3
A8JH66	-	1.3E+06	1.8E+06	2.6E+06	2.3E+06	0.2	0.2	0.3	0.3
A8JH68	-	1.7E+09	2.1E+08	9.0E+08	1.6E+08	16.7	2.4	22.8	1.4
A8JH72	-	2.8E+05	1.3E+05	9.7E+04	9.4E+04	0.7	0.3	0.3	0.3
A8JH75	-	6.7E+06	9.5E+06	1.4E+07	8.8E+06	0.2	0.2	0.5	0.3
A8JH77	-	1.4E+07	1.8E+06	4.0E+06	3.1E+06	1.0	0.0	0.5	0.3
A8JH81	-	2.4E+06	3.4E+06	2.1E+05	3.0E+05	0.2	0.2	0.2	0.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
A8JH97	-	2.9E+07	2.2E+06	1.9E+07	2.7E+06	2.8	0.4	1.7	0.5
A8JH98	-	1.3E+09	5.5E+07	7.0E+08	8.0E+07	30.8	2.1	36.7	4.6
A8JHA6	-	3.1E+06	4.4E+06	7.3E+05	1.0E+06	0.3	0.5	0.2	0.2
A8JHA9	-	1.3E+07	6.2E+06	2.4E+07	3.2E+06	0.7	0.3	1.0	0.0
A8JHB2	-	1.5E+07	9.5E+06	4.6E+07	4.1E+07	0.5	0.3	1.2	0.2
A8JHB4	-	3.7E+08	4.0E+07	1.7E+08	3.1E+07	11.3	1.4	7.5	1.0
A8JHB7	-	2.5E+08	2.6E+07	1.9E+08	3.2E+07	6.3	0.7	6.2	0.9
A8JHC3	-	1.1E+08	4.1E+07	2.4E+07	8.7E+06	2.8	0.6	1.2	0.6
A8JHC9	-	1.8E+08	1.3E+07	1.5E+08	2.8E+07	4.2	0.4	4.0	0.7
A8JHI5	-	0.0E+00	0.0E+00	2.0E+05	2.9E+05	0.0	0.0	0.2	0.2
A8JHJ5	-	2.3E+08	2.0E+07	1.4E+08	4.1E+07	4.8	0.4	3.7	0.7
A8JHJ9	-	1.3E+07	4.1E+06	1.5E+07	4.2E+06	0.8	0.2	0.8	0.2
A8JHL1	-	5.4E+07	3.0E+06	3.0E+07	1.2E+07	1.5	0.3	1.0	0.4
A8JHL7	-	1.9E+06	2.7E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JHM2	-	9.8E+07	2.4E+07	1.0E+08	2.7E+07	4.7	0.9	6.0	1.2
A8JHN9	-	1.4E+09	3.6E+07	1.1E+09	1.5E+08	8.3	0.6	9.2	0.4
A8JHP2	-	3.6E+06	3.4E+06	1.2E+06	1.6E+06	0.3	0.3	0.2	0.2
A8JHQ3	-	2.7E+07	2.9E+06	1.6E+07	1.6E+06	1.2	0.2	1.0	0.0
A8JHQ7	-	6.9E+07	1.3E+07	3.6E+07	1.7E+07	2.8	0.4	1.2	0.6
A8JHU2	-	4.6E+08	3.3E+07	3.7E+08	3.3E+07	4.0	0.6	3.7	0.5
A8JHU6	-	1.6E+07	5.2E+06	7.7E+06	7.5E+06	0.8	0.2	0.7	0.6
A8JHW7	-	1.5E+07	1.0E+07	1.4E+07	1.1E+06	0.7	0.5	1.0	0.0
A8JHY4	-	2.5E+07	1.5E+07	4.3E+06	3.9E+06	1.0	0.5	0.3	0.3
A8JI60	-	2.6E+08	9.5E+07	1.1E+08	1.1E+07	4.2	0.4	3.3	0.3
A8JI83	-	3.3E+07	1.3E+07	7.8E+06	2.0E+06	3.3	0.9	2.8	0.8
A8JI94	-	3.2E+08	3.5E+07	1.9E+08	2.1E+07	6.7	1.0	5.5	1.4
A8JIB5	-	0.0E+00	0.0E+00	9.5E+06	1.3E+07	0.0	0.0	0.3	0.5
A8JIB7	-	1.0E+09	3.4E+07	7.2E+08	7.7E+07	23.7	1.7	22.5	2.2
A8JIB8	-	6.2E+06	4.0E+06	5.3E+04	4.8E+04	0.8	0.7	0.3	0.3
A8JIB9	-	0.0E+00	0.0E+00	2.3E+06	2.1E+06	0.0	0.0	0.3	0.3
A8JIC4	-	8.5E+06	3.8E+06	5.5E+06	3.6E+06	0.7	0.3	0.5	0.3
A8JID6	-	5.3E+08	1.1E+08	2.3E+08	5.8E+07	23.7	3.4	28.2	3.1
A8JIE0	-	2.0E+08	1.9E+07	7.1E+07	1.2E+07	4.5	0.5	3.7	0.3
A8JIE3	-	1.3E+06	1.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JIE5	-	1.7E+08	6.7E+07	1.8E+08	1.4E+07	3.2	1.1	3.5	0.6
A8JIJ0	-	2.0E+06	2.8E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
A8JIJ1	-	2.8E+08	5.6E+07	2.2E+08	2.0E+07	4.8	0.7	5.7	0.9
A8JIN6	-	1.9E+08	3.3E+07	3.3E+07	1.2E+07	3.8	0.4	2.5	1.0
A8JIR0	-	3.0E+08	5.4E+07	1.1E+08	3.4E+07	13.0	1.9	7.8	2.6
A8JIU6	-	6.5E+06	9.2E+06	2.4E+06	1.5E+06	0.2	0.2	0.5	0.3
A8JJ55	-	1.4E+07	9.9E+05	1.8E+06	1.7E+06	1.0	0.0	0.3	0.3
A8JJ80	-	4.5E+05	2.1E+05	2.2E+05	8.4E+04	4.3	2.0	5.2	1.5
A8JJE0	-	9.5E+08	7.7E+07	4.8E+08	9.7E+07	11.2	1.2	8.2	1.6
A8JJE1	-	2.4E+09	1.8E+08	1.6E+09	1.3E+08	38.0	1.0	48.5	2.9
A8JJH8	-	4.8E+06	6.7E+06	3.8E+07	1.0E+07	0.2	0.2	1.7	0.3
A8JJS2	-	9.2E+06	1.3E+07	0.0E+00	0.0E+00	0.8	1.2	0.0	0.0
A8JJY5	-	0.0E+00	0.0E+00	1.0E+04	1.4E+04	0.0	0.0	0.2	0.2
A8JK20	-	1.1E+08	3.5E+07	1.3E+08	2.5E+07	2.8	0.8	3.0	0.6
A8JK39	-	6.0E+08	7.5E+07	3.9E+08	2.0E+07	7.0	3.6	10.8	6.5
A8JKF2	-	4.8E+07	6.4E+06	3.6E+07	9.5E+06	2.0	0.4	1.5	0.3
B8XSL8	-	7.4E+08	6.5E+07	3.8E+08	8.1E+07	49.8	2.7	38.8	2.2
C0SPI7	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
C5HJ22	-	6.9E+06	4.5E+06	1.6E+06	2.3E+06	0.7	0.5	0.3	0.5

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
C5HJB7	-	2.4E+08	5.4E+07	1.1E+08	1.8E+07	4.5	1.1	4.3	1.9
C5HJD6	-	3.0E+07	2.8E+06	1.3E+07	2.4E+06	2.0	0.5	3.0	0.4
D5LAT9	-	0.0E+00	0.0E+00	5.7E+06	3.7E+06	0.0	0.0	0.5	0.3
D5LAY0	-	1.0E+08	5.7E+06	5.9E+07	1.3E+07	3.0	0.4	1.7	0.3
D5LAZ0	-	1.6E+05	1.4E+05	2.7E+05	1.9E+05	0.3	0.3	0.5	0.3
D5LB05	-	5.5E+06	2.6E+06	1.4E+06	1.3E+06	1.0	0.5	0.3	0.3
D5LB14	-	4.6E+07	1.5E+07	7.3E+06	3.9E+06	1.3	0.5	0.8	0.4
E3SC57	-	6.8E+08	8.3E+07	5.3E+08	1.1E+08	8.5	1.5	5.7	1.6
H6V961	-	3.1E+06	2.8E+06	1.5E+06	1.3E+06	0.3	0.3	0.3	0.3
O20032	-	2.0E+08	2.3E+07	5.6E+07	2.4E+07	7.5	1.9	10.0	4.3
O22448	-	1.5E+07	7.4E+06	2.9E+06	4.2E+06	0.7	0.3	0.2	0.2
O47027	-	3.2E+08	2.1E+07	2.1E+08	4.3E+07	11.3	1.4	11.0	2.0
O48949	-	9.6E+08	9.2E+07	1.1E+09	1.0E+08	21.2	1.6	17.5	0.7
O49822	-	5.0E+07	1.2E+07	6.9E+07	1.2E+07	1.3	0.5	1.3	0.3
O63075	-	4.2E+07	8.7E+06	1.7E+07	8.0E+06	3.8	1.2	3.5	1.8
O64985	-	1.7E+06	2.4E+06	3.8E+06	1.1E+06	0.2	0.2	0.8	0.2
O65104	-	6.3E+08	5.1E+07	3.2E+08	5.4E+07	19.8	2.3	21.0	1.8
O81349	-	4.0E+07	5.5E+06	1.4E+07	6.4E+06	2.8	0.2	1.3	0.6
O81523	-	1.3E+08	1.4E+07	9.1E+07	1.5E+07	4.0	0.5	3.0	0.8
O81648	-	6.3E+06	4.3E+06	2.5E+06	3.5E+06	0.5	0.3	0.2	0.2
O98224	-	4.8E+07	4.3E+07	6.8E+07	3.0E+07	0.3	0.3	0.8	0.4
P05726	-	2.3E+07	2.6E+06	1.2E+07	5.4E+06	1.0	0.0	0.8	0.4
P07753	-	2.1E+09	2.6E+08	2.2E+09	1.8E+08	64.0	5.0	64.8	7.7
P09144	-	3.6E+09	2.7E+08	3.7E+09	2.2E+08	79.0	8.5	83.8	10.5
P11094	-	2.6E+07	1.1E+07	5.1E+06	4.6E+06	1.2	0.4	0.3	0.3
P14149	-	1.8E+08	6.4E+06	1.1E+08	1.8E+07	4.0	0.5	6.7	1.7
P22675	-	1.4E+08	2.0E+07	7.2E+07	1.2E+07	5.0	1.4	4.5	0.6
P23489	-	9.5E+08	2.0E+08	3.5E+08	2.4E+07	15.5	1.0	9.8	2.3
P26565	-	2.6E+07	1.8E+07	4.5E+07	1.6E+07	0.5	0.3	1.3	0.6
P32974	-	3.2E+08	1.4E+08	3.8E+08	3.0E+07	1.0	0.5	1.0	0.0
P36495	-	2.5E+08	3.9E+07	2.3E+08	2.8E+07	8.0	1.2	8.0	0.6
P37255	-	4.2E+09	2.4E+08	4.9E+09	4.7E+08	315.8	12.2	323.8	14.5
P42380	-	3.0E+07	6.2E+06	2.7E+07	6.8E+06	2.2	0.8	3.0	0.8
P48267	-	1.1E+08	6.4E+07	6.2E+07	1.5E+07	1.7	0.9	1.8	0.6
P48268	-	2.7E+09	1.2E+09	1.7E+09	3.6E+08	8.8	0.6	6.8	1.2
P48270	-	8.3E+08	4.3E+07	6.7E+08	7.4E+07	9.2	1.3	12.2	3.0
P49644	-	4.0E+08	2.3E+08	2.2E+08	1.7E+08	3.8	0.6	2.8	0.6
P59775	-	2.0E+07	2.0E+07	1.4E+07	1.5E+07	0.7	0.5	0.5	0.5
P59776	-	2.3E+08	2.7E+07	1.8E+08	1.2E+07	1.2	0.2	1.0	0.0
P81831	-	1.6E+07	1.0E+07	7.9E+06	7.0E+06	0.8	0.6	0.3	0.3
P83564	-	3.7E+08	2.6E+07	1.9E+08	3.3E+07	7.3	1.1	6.5	1.9
P93106	-	3.6E+09	2.5E+08	2.2E+09	1.3E+08	97.8	6.8	88.0	8.5
Q00471	-	8.9E+08	2.0E+07	7.5E+08	4.5E+07	48.5	3.6	47.0	2.7
Q06824	-	0.0E+00	0.0E+00	1.6E+06	2.3E+06	0.0	0.0	0.2	0.2
Q08363	-	1.2E+07	4.5E+06	1.2E+07	4.0E+06	6.0	2.7	10.7	4.7
Q0ZAZ1	-	1.0E+07	6.2E+06	1.7E+06	2.4E+06	0.7	0.3	0.2	0.2
Q19VH4	-	2.2E+08	2.8E+07	1.3E+08	2.2E+07	4.3	0.5	4.0	0.4
Q19VH5	-	2.0E+06	2.8E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
Q1G2Y1	-	3.1E+06	2.2E+06	1.8E+06	1.6E+06	1.3	0.6	0.7	0.5
Q24K71	-	1.2E+08	9.4E+06	2.4E+08	1.8E+08	1.5	0.3	1.0	0.5
Q27YU5	-	2.1E+07	6.3E+06	1.6E+07	4.8E+06	0.8	0.2	1.2	0.4
Q2VA41	-	1.0E+08	2.8E+07	5.3E+07	1.8E+07	4.7	1.1	4.0	0.9
Q39576	-	4.4E+07	1.4E+07	3.9E+07	1.1E+07	0.8	0.2	0.8	0.2

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
Q39588	-	7.7E+07	1.3E+07	1.5E+07	5.4E+06	2.0	0.4	0.8	0.2
Q39595	-	2.0E+08	6.5E+07	9.1E+07	2.3E+07	5.7	0.7	4.7	0.6
Q39603	-	2.1E+09	1.5E+08	1.0E+09	1.2E+08	34.8	3.9	39.0	3.4
Q3Y8L7	-	1.3E+07	3.1E+06	1.3E+07	3.7E+06	1.0	0.0	0.8	0.2
Q42683	-	8.9E+08	1.2E+08	4.5E+08	2.1E+07	9.3	0.9	9.7	0.6
Q42685	-	8.1E+07	1.0E+07	9.6E+07	1.8E+07	3.2	0.4	4.0	0.7
Q42703	-	3.5E+07	1.7E+07	3.4E+07	5.7E+06	1.3	0.5	3.5	0.3
Q4U0V8	-	0.0E+00	0.0E+00	7.1E+05	1.0E+06	0.0	0.0	0.2	0.2
Q4U1D9	-	2.2E+07	8.1E+06	2.3E+06	2.0E+06	1.2	0.2	0.3	0.3
Q5QEB2	-	8.2E+07	1.1E+07	8.5E+07	1.2E+07	5.0	0.6	6.3	1.8
Q5S7Y5	-	4.0E+08	3.8E+07	2.2E+08	3.3E+07	4.8	0.8	3.8	0.7
Q5VLJ9	-	7.5E+05	1.1E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
Q5W9T2	-	1.1E+08	1.3E+07	1.3E+07	5.8E+06	3.2	0.2	1.3	0.3
Q65Z18	-	2.3E+07	1.8E+06	1.4E+07	4.0E+06	1.0	0.0	0.8	0.2
Q66T67	-	2.9E+07	7.7E+06	1.9E+07	9.6E+06	2.5	0.5	1.0	0.0
Q66YD0	-	7.1E+07	1.5E+07	8.3E+07	1.4E+07	3.2	0.7	4.0	0.8
Q66YD3	-	5.6E+07	9.3E+06	2.4E+07	8.9E+06	1.5	0.3	1.8	0.8
Q68RJ5	-	3.1E+05	1.1E+05	6.1E+04	5.5E+04	1.3	0.6	0.3	0.3
Q68UI8	-	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0	0.0	0.0	0.0
Q6DN05	-	2.6E+07	8.1E+06	6.1E+06	5.6E+06	1.5	0.5	0.7	0.6
Q6EMK7	-	1.5E+08	2.4E+07	1.4E+08	1.2E+07	3.5	0.8	5.8	1.2
Q6IV67	-	4.8E+07	6.0E+06	4.7E+07	4.3E+06	4.2	0.9	3.8	1.6
Q6IYG1	-	0.0E+00	0.0E+00	9.0E+05	1.3E+06	0.0	0.0	0.2	0.2
Q6J213	-	1.2E+08	1.2E+07	1.0E+08	9.6E+06	3.5	0.8	4.0	0.5
Q6QAY0	-	2.0E+06	2.9E+06	0.0E+00	0.0E+00	0.2	0.2	0.0	0.0
Q6QAY1	-	4.8E+07	9.3E+06	2.0E+07	7.2E+06	1.8	0.2	1.3	0.3
Q6QAY3	-	2.7E+08	3.9E+07	9.4E+07	4.5E+07	4.0	0.5	3.5	0.6
Q6QAY4	-	3.0E+07	1.5E+07	1.6E+07	5.8E+06	1.2	0.6	1.0	0.4
Q6QIV9	-	4.2E+06	3.7E+06	2.5E+06	3.6E+06	0.3	0.3	0.2	0.2
Q6QJE1	-	1.6E+07	7.9E+06	4.9E+06	3.1E+06	0.8	0.4	0.5	0.3
Q6R2V6	-	1.6E+06	4.6E+05	8.3E+05	7.4E+04	1.5	0.3	1.0	0.0
Q6RBZ0	-	1.4E+08	1.8E+07	1.2E+08	1.4E+07	2.0	0.4	2.5	0.3
Q6S7R7	-	3.1E+07	3.5E+06	4.0E+06	2.6E+06	1.2	0.2	0.8	0.7
Q6UKY5	-	4.5E+07	1.2E+07	1.9E+07	4.3E+06	1.7	0.5	3.3	0.6
Q6UKY6	-	9.6E+07	8.1E+06	6.0E+07	2.3E+07	2.7	0.3	2.2	0.7
Q6UP29	-	2.2E+07	1.4E+06	1.5E+07	1.7E+06	1.0	0.0	1.0	0.0
Q6UP31	-	3.1E+07	1.6E+07	3.2E+07	2.9E+06	1.2	0.6	1.2	0.2
Q6V506	-	4.1E+07	3.4E+06	1.6E+07	8.6E+06	1.3	0.3	1.3	0.7
Q6V507	-	3.1E+07	3.8E+06	8.1E+06	4.5E+06	1.5	0.3	1.5	0.6
Q6V8K6	-	0.0E+00	0.0E+00	2.0E+07	1.1E+07	0.0	0.0	0.8	0.4
Q6V9A8	-	1.5E+07	6.6E+06	3.8E+06	1.9E+06	1.7	0.7	0.8	0.4
Q6V9B0	-	2.3E+07	8.6E+06	4.8E+06	2.8E+06	0.8	0.2	0.8	0.2
Q6V9B2	-	8.0E+06	8.4E+06	0.0E+00	0.0E+00	0.3	0.3	0.0	0.0
Q6VTH1	-	0.0E+00	0.0E+00	9.5E+05	6.1E+05	0.0	0.0	0.5	0.3
Q6WEE4	-	0.0E+00	0.0E+00	2.9E+05	4.0E+05	0.0	0.0	0.2	0.2
Q6Y682	-	2.0E+08	1.6E+07	2.1E+07	5.1E+06	3.5	0.5	1.7	0.5
Q70DX8	-	2.6E+08	4.4E+07	2.6E+08	1.3E+08	6.2	0.8	9.2	0.2
Q75NZ1	-	1.2E+09	8.4E+07	6.4E+08	9.6E+07	49.0	2.1	43.8	1.9
Q75NZ3	-	1.9E+07	2.5E+06	8.6E+05	1.2E+06	1.2	0.2	0.2	0.2
Q75VY7	-	5.2E+08	1.1E+08	2.9E+08	1.4E+08	58.7	19.2	68.5	18.5
Q75VY9	-	4.9E+08	7.2E+07	2.6E+08	4.6E+07	43.7	3.0	35.7	5.7
Q75VZ0	-	4.9E+08	7.2E+07	5.0E+08	1.0E+08	56.7	9.6	65.7	7.2
Q763T6	-	2.8E+08	1.5E+07	2.5E+08	1.8E+07	6.5	0.8	9.0	0.8

Appendix 1.

UniProt ID	Sig. diff.	WT_Int	WT_Int s.e.	Sp_Int	Sp_Int s.e.	WT_SC	WT_SC s.e.	Sp_SC	Sp_SC s.e.
Q7DM26	-	4.4E+08	7.6E+07	3.4E+08	1.1E+08	23.8	2.8	26.5	5.7
Q7X7A7	-	4.4E+08	5.2E+07	3.3E+08	1.2E+07	4.8	1.0	6.3	0.9
Q7Y258	-	2.9E+08	9.2E+06	1.1E+08	2.0E+07	6.3	0.5	3.3	0.7
Q7YKX3	-	1.8E+07	2.5E+07	1.4E+08	1.4E+07	0.2	0.2	1.0	0.0
Q84JV7	-	5.8E+05	8.9E+04	4.2E+05	1.0E+05	1.0	0.0	1.2	0.2
Q84SA0	-	8.2E+07	6.1E+06	6.9E+07	3.2E+06	1.8	0.2	1.8	0.2
Q84U22	-	1.3E+08	4.4E+07	4.3E+07	1.5E+07	4.2	1.3	2.7	0.7
Q84UB2	-	7.9E+07	7.8E+06	4.3E+07	1.2E+07	3.5	0.8	3.2	1.1
Q84X75	-	7.9E+08	6.7E+07	4.4E+08	9.0E+07	8.3	0.5	8.7	0.9
Q84X77	-	4.1E+05	5.8E+04	9.9E+04	1.1E+05	1.0	0.0	0.3	0.3
Q84X82	-	2.4E+06	3.4E+06	1.1E+07	5.4E+06	0.2	0.2	1.2	0.6
Q84Y02	-	2.2E+09	1.6E+08	1.6E+09	2.2E+08	40.2	1.7	39.3	3.7
Q8GV23	-	1.1E+08	5.6E+06	1.1E+08	5.1E+06	1.0	0.0	1.5	0.3
Q8HTL1	-	7.3E+07	2.1E+07	2.7E+07	9.2E+06	1.5	0.3	1.0	0.4
Q8HTL2	-	5.4E+08	1.1E+08	6.4E+08	1.2E+08	6.2	0.8	8.5	1.1
Q8HTL3	-	3.8E+07	2.7E+07	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0
Q8HTL5	-	1.9E+08	3.7E+07	1.2E+08	2.4E+07	3.0	0.5	2.3	0.3
Q8HUG9	-	6.3E+06	4.2E+06	1.4E+07	4.8E+06	0.5	0.3	1.0	0.4
Q8HUH1	-	7.6E+07	2.9E+07	2.0E+08	6.9E+07	1.5	0.3	2.2	0.6
Q8LK22	-	2.2E+08	2.0E+07	1.3E+08	3.8E+07	8.0	1.9	6.2	1.7
Q8LKI3	-	9.1E+07	2.2E+07	1.6E+08	1.1E+07	3.8	0.7	6.3	0.9
Q8LL91	-	3.0E+07	1.0E+07	2.9E+07	8.8E+06	8.3	0.9	5.3	1.0
Q8LPD9	-	1.2E+08	1.4E+07	4.7E+07	7.8E+06	3.2	0.4	2.0	0.4
Q8LRU1	-	4.9E+07	1.7E+07	1.8E+07	5.3E+06	1.3	0.5	0.8	0.2
Q8RVB8	-	9.7E+07	1.4E+07	6.4E+07	1.5E+07	1.5	0.3	1.2	0.2
Q8S2V8	-	4.5E+08	2.7E+07	2.8E+08	1.8E+07	16.2	1.5	19.8	1.1
Q8S4W5	-	2.8E+06	3.9E+06	5.4E+06	3.5E+06	0.2	0.2	0.5	0.3
Q8SAQ6	-	3.3E+07	7.2E+06	1.3E+07	5.2E+06	2.5	0.3	1.7	0.7
Q8VZZ5	-	2.5E+07	8.6E+06	8.8E+06	5.2E+06	1.2	0.2	0.8	0.4
Q8W4V3	-	2.3E+08	2.9E+07	1.8E+08	2.3E+07	8.5	1.4	13.5	1.3
Q93VE0	-	5.4E+09	3.6E+08	4.4E+09	5.4E+08	63.8	5.8	71.5	4.1
Q93WL4	-	2.8E+09	9.7E+07	1.9E+09	3.6E+08	112.2	5.5	106.8	5.4
Q93Y49	-	2.0E+08	2.1E+07	1.8E+08	4.4E+07	5.8	0.8	3.3	0.6
Q93Z22	-	2.3E+07	4.3E+06	2.1E+07	4.9E+06	1.3	0.3	1.3	0.3
Q945T2	-	2.7E+06	3.8E+06	5.1E+06	7.2E+06	0.2	0.2	0.2	0.2
Q948T5	-	8.2E+06	6.5E+06	0.0E+00	0.0E+00	0.7	0.5	0.0	0.0
Q94CJ2	-	2.0E+07	2.0E+06	6.9E+06	6.2E+06	1.0	0.0	0.3	0.3
Q94KS3	-	9.4E+06	5.0E+06	1.8E+06	1.6E+06	0.8	0.4	0.3	0.3
Q94KV1	-	1.8E+08	2.6E+07	7.3E+07	2.1E+07	4.5	0.8	2.2	0.8
Q96550	-	4.6E+09	2.0E+08	4.0E+09	3.2E+08	31.8	1.8	34.5	4.8
Q9AU03	-	1.7E+08	1.8E+07	5.9E+07	1.8E+07	3.5	0.3	3.7	0.7
Q9AU05	-	1.2E+07	1.3E+07	3.9E+06	3.6E+06	0.3	0.3	0.3	0.3
Q9AXF6	-	4.7E+09	4.2E+08	3.1E+09	7.6E+08	141.3	9.5	128.8	14.1
Q9FEH5	-	1.3E+09	1.1E+08	8.8E+08	1.6E+08	19.7	1.7	17.2	3.3
Q9FEK6	-	3.3E+09	1.8E+08	2.1E+09	2.7E+08	124.5	15.7	118.7	5.6
Q9FXQ1	-	4.5E+07	3.7E+07	2.1E+07	1.7E+07	1.7	0.5	1.2	0.7
Q9LLL6	-	7.2E+08	4.1E+07	3.4E+08	3.7E+07	19.5	1.4	16.5	1.9
Q9S7V1	-	1.4E+08	1.7E+07	7.0E+07	1.0E+07	5.7	1.5	3.3	0.5
Q9SW75	-	6.4E+08	5.7E+07	5.6E+08	4.8E+07	4.3	0.9	5.0	0.8
Q9XGU3	-	9.8E+06	8.8E+06	1.3E+07	8.4E+06	0.3	0.3	0.5	0.3
Q9ZSJ4	-	2.4E+09	1.3E+08	1.7E+09	2.1E+08	173.3	17.9	180.8	6.5
Q9ZTA7	-	7.1E+07	7.9E+06	1.7E+07	7.9E+06	3.3	0.3	1.0	0.5

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A1YSB4	509826	Cre10.g430150.t1.2	S	o	o	o	o	o
A6Q0K5	525653	Cre08.g380250.t1.2	S	o	o	o	o	o
A8HMM7	511486	Cre01.g042200.t1.1	S	o	-	o	o	o
A8HNC0	511658	Cre01.g050100.t1.2	W	o	o	o	o	o
A8HNG8	511686	Cre01.g051500.t1.2	W	o	-	+	-	-
A8HP50	515409	Cre14.g621300.t2.1	W	o	-	+	-	+
A8HP50	515409	Cre14.g621300.t1.1	W	o	-	+	-	+
A8HPY4	511076	Cre01.g022650.t1.2	W	-	+	+	+	o
A8HQ74	511155	Cre01.g026550.t1.1	W	+	-	o	o	o
A8HQA8	511190	Cre01.g028200.t1.1	W	o	-	o	o	o
A8HQA8	511190	Cre01.g028250.t1.2	W	o	-	o	o	o
A8HQA8	511190	Cre01.g028200.t2.1	W	o	-	o	o	o
A8HQP0	511284	Cre01.g032650.t1.2	W	o	o	o	o	o
A8HQP0	511284	Cre01.g032650.t2.1	W	o	o	o	o	o
A8HRZ9	514278	Cre13.g567700.t1.2	W	o	-	o	o	o
A8HS48	514298	Cre13.g568650.t1.2	o	o	o	o	o	o
A8HSJ6	514364	Cre13.g571750.t1.2	W	o	-	+	o	+
A8HWZ6	523133	Cre06.g264350.t1.2	W	o	-	o	o	o
A8HX38	512813	Cre12.g498600.t1.2	W	o	o	o	o	o
A8HYU5	522823	Cre06.g250200.t1.2	W	+	-	o	o	o
A8HZX4	510268	Cre10.g451400.t1.1	S	o	o	o	o	o
A8I0A3	510326	Cre10.g454150.t1.2	W	o	o	o	o	o
A8I2V3	519296	Cre02.g114600.t1.2	W	o	o	+	o	+
A8I3J7	519187	Cre02.g109550.t1.2	S	-	-	+	-	-
A8I3W8	519136	Cre02.g107200.t1.1	W	-	o	+	+	o
A8I4P5	519030	Cre02.g102250.t1.2	W	-	o	o	o	o
A8I531	522598	Cre05.g242000.t1.2	W	+	-	o	o	o
A8I6P9	524314	Cre07.g318450.t1.2	W	+	-	o	o	o
A8I9T2	518407	Cre02.g073750.t1.2	S	+	o	o	o	o
A8IB25	509880	Cre10.g432800.t1.2	W	o	o	o	o	o
A8IBN3	509806	Cre10.g429100.t1.2	W	o	o	+	-	-
A8ICF4	509690	Cre10.g423500.t1.2	S	o	o	o	o	o
A8ICG9	509683	Cre10.g423250.t1.2	W	-	+	+	+	+
A8ICT1	509631	Cre10.g420750.t1.2	W	=	o	o	o	o
A8ICU7	509626	Cre10.g420550.t1.1	S	+	-	o	o	o
A8IEH0	520905	Cre03.g172300.t1.2	S	o	-	o	o	o
A8IGD9	509959	Cre10.g436550.t1.2	W	-	+	+	+	+
A8IGE2	509958	Cre10.g436500.t1.2	W	+	-	+	-	o
A8IGM2	509931	Cre10.g435300.t1.2	W	o	o	o	o	o
A8IGM2	509931	Cre10.g435300.t2.1	W	o	o	o	o	o
A8IGV4	524609	Cre07.g332300.t1.2	W	-	+	+	+	o
A8IHW6	510130	Cre10.g444800.t1.2	W	o	-	o	o	o
A8III5	510035	Cre10.g440450.t1.2	W	+	-	+	-	-
A8IIK4	510026	Cre10.g440050.t1.2	W	o	o	o	o	o
A8IJ34	512854	Cre12.g500500.t2.1	W	+	o	+	o	-
A8IJ34	512854	Cre12.g500500.t1.1	W	o	o	+	o	-
A8IJJ8	520788	Cre12.g503300.t1.2	W	o	o	o	o	o
A8IJQ4	512935	Cre12.g504200.t1.2	W	o	o	o	o	o
A8IKX4	513105	Cre12.g512300.t1.2	S	o	-	+	o	+

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IMP6	524062	Cre06.g308250.t1.2	W	o	o	o	o	o
A8IMS6	524049	Cre06.g307600.t1.2	S	o	o	o	o	o
A8INX1	523881	Cre06.g299650.t1.2	W	o	o	o	o	o
A8IQG4	517100	Cre17.g700750.t1.2	S	o	o	o	o	o
A8IRT6	521268	Cre03.g189400.t1.2	W	o	-	+	-	o
A8IRW1	521296	Cre03.g190400.t1.1	W	o	o	o	o	o
A8IRX5	521309	Cre03.g191050.t1.2	S	+	-	o	o	o
A8IS22	516709	Cre16.g682300.t1.2	S	o	o	o	o	o
A8IT01	522126	Cre04.g223100.t1.2	W	o	+	+	+	+
A8ITL0	524777	Cre07.g340350.t1.1	o	+	o	o	o	o
A8ITZ2	524920	Cre07.g347150.t1.2	W	o	o	o	o	o
A8IUN8	525170	Cre14.g632775.t1.1	W	o	o	o	o	o
A8IV51	515514	Cre14.g626200.t1.1	W	+	o	+	-	+
A8IVF6	513636	Cre12.g537450.t1.2	W	o	o	o	o	o
A8IW09	513483	Cre12.g530300.t1.2	W	o	o	o	o	o
A8IW47	517819	Cre17.g734500.t1.2	W	o	-	o	o	o
A8IWJ8	521366	Cre03.g193850.t1.2	W	o	o	o	o	o
A8IWP6	521401	Cre03.g195500.t1.2	S	-	o	+	+	o
A8IWP6	521401	Cre03.g195500.t2.1	S	-	o	+	+	o
A8IX13	521668	Cre03.g208100.t1.1	S	o	o	o	o	o
A8IX80	521637	Cre03.g206600.t1.2	W	o	-	o	o	o
A8IXZ0	513735	Cre12.g549550.t1.2	W	+	o	o	o	o
A8IXZ0	513735	Cre12.g542250.t1.1	W	+	o	o	o	o
A8IY40	513772	Cre12.g544050.t1.1	W	+	-	+	-	-
A8IY43	513775	Cre12.g544150.t1.2	W	o	o	o	o	o
A8IY85	513815	Cre12.g546050.t1.2	W	o	-	o	o	o
A8IZS5	525923	Cre09.g392350.t1.2	W	o	o	o	o	o
A8IZS5	525923	Cre09.g392350.t2.1	W	o	o	o	o	o
A8IZW3	525969	Cre09.g394350.t1.2	W	+	o	o	o	o
A8J0I0	519717	Cre09.g397697.t1.1	W	=	o	+	o	-
A8J0R4	519902	Cre02.g143050.t1.2	S	=	o	o	o	o
A8J0R7	519907	Cre02.g143250.t1.2	W	+	-	o	o	o
A8J129	519460	Cre09.g387726.t1.1	S	-	o	o	o	o
A8J225	523507	Cre06.g281600.t2.1	W	o	+	+	+	+
A8J225	523507	Cre06.g281600.t1.2	W	o	+	+	+	+
A8J244	523531	Cre06.g282800.t1.2	S	-	o	o	o	o
A8J2J7	523717	Cre06.g291700.t1.2	W	o	o	o	o	o
A8J2S0	520404	Cre03.g149100.t1.2	S	o	-	o	o	o
A8J3J0	516534	Cre16.g674300.t1.1	W	o	-	o	o	o
A8J3W1	516676	Cre16.g680700.t1.2	W	o	o	o	o	o
A8J3Y6	525320	Cre08.g364800.t1.2	W	o	o	o	o	o
A8J448	525395	Cre08.g368400.t1.2	W	o	-	o	o	o
A8J462	525405	Cre08.g368900.t1.2	S	o	-	o	o	o
A8J462	525405	Cre08.g368900.t2.1	S	o	-	o	o	o
A8J493	525240	Cre08.g360900.t1.2	W	o	o	o	o	o
A8J4M0	526291	Cre09.g409850.t1.2	W	+	-	o	o	o
A8J4Q3	526318	Cre09.g411100.t1.2	S	-	+	o	o	o
A8J4Q3	526318	Cre09.g411100.t2.1	S	o	+	o	o	o
A8J576	513463	Cre12.g529400.t1.2	W	=	o	o	o	o
A8J5B8	513432	Cre12.g528000.t1.2	S	o	-	+	-	-
A8J646	513248	Cre12.g519100.t1.2	W	o	o	o	o	o
A8J6D1	517514	Cre17.g720250.t1.2	S	=	-	o	o	o

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J6M8	517621	Cre17.g725350.t1.1	W	o	-	o	o	o
A8J6Q7	517651	Cre17.g726750.t1.2	W	o	o	o	o	o
A8J7P4	520495	Cre03.g153450.t1.2	W	o	-	+	-	+
A8J7X3	519813	NA	S	NA	o	+	+	+
A8J841	522983	Cre05.g240850.t1.2	S	o	-	o	o	o
A8J995	516128	Cre16.g655150.t1.2	W	o	o	o	o	o
A8J9T0	NA	Cre04.g214503.t2.1	W	=	o	o	o	o
A8J9T0	NA	Cre04.g214503.t1.1	W	o	o	o	o	o
A8JAW4	515051	Cre13.g604650.t2.1	W	o	-	o	o	o
A8JAW4	515051	Cre13.g604650.t1.2	W	o	-	o	o	o
A8JB67	516886	Cre16.g690500.t1.2	W	o	-	+	-	-
A8JBX6	518319	Cre03.g213537.t1.1	W	+	o	o	o	o
A8JC15	512139	Cre11.g467400.t1.2	W	+	o	+	+	+
A8JCP5	511747	Cre01.g054500.t1.2	W	o	+	o	o	o
A8JCS8	520129	Cre01.g061077.t1.1	W	o	o	o	o	o
A8JCY4	510743	Cre01.g006950.t2.1	W	o	o	+	o	-
A8JCY4	510743	Cre01.g006950.t1.1	W	o	o	+	o	-
A8JCZ5	510810	Cre01.g009950.t1.2	S	o	-	o	o	o
A8JD42	510621	Cre01.g000900.t1.2	W	o	o	o	o	o
A8JDE1	517366	Cre17.g713450.t1.2	S	o	o	o	o	o
A8JDH1	517327	Cre17.g711750.t1.2	S	o	+	+	o	o
A8JEA7	517991	Cre17.g742200.t1.1	W	+	o	o	o	o
A8JEP0	521074	Cre04.g217933.t1.1	S	o	o	o	o	o
A8JEQ7	521726	Cre04.g217910.t1.1	W	o	-	o	o	o
A8JEV1	525487	Cre08.g372450.t1.2	W	o	-	o	o	o
A8JEV9	525497	Cre08.g372950.t1.2	W	o	o	o	o	o
A8JFF0	517758	Cre17.g731571.t1.1	W	-	o	o	o	o
A8JFR9	525057	Cre07.g353450.t1.2	S	=	o	o	o	o
A8JGF4	525207	Cre08.g359350.t1.2	W	o	o	o	o	o
A8JGI9	512819	Cre12.g498900.t1.2	W	=	o	+	o	-
A8JGW2	514086	Cre12.g558900.t1.2	W	o	-	o	o	o
A8JHA4	513135	Cre12.g513500.t1.2	W	+	-	+	-	+
A8JHU0	512496	Cre12.g483950.t1.2	o	o	o	o	o	o
A8JHX9	513191	Cre12.g516200.t1.2	W	o	-	+	-	o
B2XY88	NA	NA	W	NA	o	o	o	o
C5HJ53	NA	NA	S	NA	o	o	o	o
I2FKQ9	524078	Cre06.g309100.t1.2	W	o	-	o	o	o
O22472	526443	Cre09.g417150.t1.2	S	-	o	o	o	o
O22547	520369	Cre09.g386758.t1.1	W	-	o	o	o	o
O24450	522130	Cre04.g223300.t1.2	W	o	+	+	+	+
P00877	NA	NA	S	o	o	o	o	o
P06007	NA	NA	S	NA	o	o	o	o
P06541	NA	NA	S	NA	o	o	o	o
P10898	NA	NA	S	NA	o	o	o	o
P11471	513916	Cre12.g550850.t1.2	W	o	-	o	o	o
P12154	NA	NA	S	o	o	o	o	o
P17746	NA	NA	W	NA	o	o	o	o
P26526	NA	NA	S	NA	o	o	o	o
Q00914	NA	NA	W	NA	o	o	o	o
Q08365	NA	NA	o	o	o	o	o	o
Q0ZAI6	510298	Cre10.g452800.t1.2	W	o	+	+	+	+
Q1RS84	518335	Cre17.g746997.t1.1	W	o	o	o	o	o

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
Q2VA40	524706	Cre07.g336950.t1.1	S	+	o	o	o	o
Q39568	511273	Cre01.g032300.t1.2	W	o	-	+	-	-
Q39590	522733	Cre05.g248450.t1.2	W	=	+	+	+	+
Q42690	522439	Cre05.g234550.t1.2	o	-	o	+	o	-
Q42690	522439	Cre05.g234550.t2.1	o	o	o	+	o	-
Q540H1	521991	Cre03.g190950.t1.2	W	+	o	o	o	o
Q540H1	521991	Cre04.g216850.t1.2	W	+	o	o	o	o
Q5NKW4	522766	Cre05.g238332.t1.1	W	=	-	o	o	o
Q6RBZ1	522835	Cre06.g250800.t1.1	W	+	-	+	-	+
Q6X898	511816	Cre03.g144807.t1.1	W	o	o	o	o	o
Q75T33	512786	Cre12.g497300.t1.2	W	o	-	o	o	o
Q75VY6	514933	Cre06.g278213.t1.1	W	o	-	o	o	o
Q75VY8	509740	Cre10.g425900.t1.2	S	+	-	o	o	o
Q7X8Y6	517131	Cre17.g702500.t1.2	S	o	o	o	o	o
Q84V17	NA	Cre12.g483650.t1.2	S	o	o	o	o	o
Q84X71	520929/30	Cre03.g173350.t1.2	S	o	o	o	o	o
Q84X71	520929/30	Cre03.g173350.t2.1	S	o	o	o	o	o
Q8GUQ9	520259	Cre07.g325746.t1.1	S	o	o	o	o	o
Q944M9	514302	Cre13.g568850.t1.2	S	o	-	o	o	o
Q9AXJ2	518601	Cre02.g082500.t1.1	W	o	o	+	-	-
Q9FNS2	522982	Cre06.g257601.t1.2	W	o	o	o	o	o
Q9FQ96	515778	Cre15.g638500.t1.2	W	o	o	o	o	o
Q9GGE2	NA	NA	W	NA	o	o	o	o
Q9S9E0	512341	Cre11.g476750.t1.2	S	+	o	o	o	o
A8HMQ3	511498	Cre01.g042800.t1.2	o	o	-	+	-	-
A8HNX3	515335	Cre14.g617900.t1.2	o	o	o	o	o	o
A8HP64	515420	Cre14.g621751.t1.1	o	+	-	+	-	o
A8HPU3	511025	Cre01.g020350.t1.2	o	o	o	o	o	o
A8HQU3	511311	Cre01.g034000.t1.2	o	o	-	o	o	o
A8HTK7	514510	Cre13.g578750.t1.2	o	o	+	+	+	+
A8HU20	514591	Cre13.g582671.t1.1	o	o	o	o	o	o
A8HU30	514594	Cre13.g582800.t1.2	o	o	o	o	o	o
A8HUK0	514667	Cre13.g586300.t1.2	o	+	o	o	o	o
A8HUU9	514707	Cre13.g588100.t1.2	o	+	o	+	+	+
A8HVD7	523369	Cre06.g275100.t1.2	o	o	-	o	o	o
A8HWA8	523232	Cre06.g269050.t1.2	o	o	-	o	o	o
A8HWZ8	523132	Cre06.g264300.t1.2	o	o	-	o	o	o
A8HX98	523087	NA	o	NA	o	o	o	o
A8HXR2	523017	Cre06.g259150.t1.2	o	o	-	+	-	o
A8HXS7	515194	Cre06.g258850.t1.2	o	+	o	o	o	o
A8HY43	522952	Cre06.g256250.t1.2	o	o	-	o	o	o
A8HYA9	514003	Cre06.g254400.t1.1	o	o	-	+	-	-
A8HYM2	522851	Cre06.g251600.t1.2	o	-	-	o	o	o
A8I0H0	510362	Cre10.g456000.t1.2	o	o	o	o	o	o
A8I252	518941	Cre02.g098150.t1.2	o	o	-	+	+	o
A8I274	518932	Cre02.g097650.t1.2	o	=	o	o	o	o
A8I403	519124	Cre02.g106600.t1.2	o	o	o	o	o	o
A8I629	524211	Cre07.g313700.t1.1	o	o	-	o	o	o
A8I9A1	518797	Cre02.g091550.t1.2	o	o	o	o	o	o
A8I9L1	518854	Cre02.g094100.t1.2	o	o	o	o	o	o
A8IAV8	509909	Cre10.g434200.t1.1	o	+	-	o	o	o
A8ICM2	509655	Cre10.g421850.t1.2	o	o	o	o	o	o

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IE32	520970	Cre03.g175200.t1.2	o	+	-	o	o	o
A8IEW6	520843	Cre03.g169400.t1.2	o	+	-	+	o	+
A8IGU4	524612	Cre07.g332450.t1.2	o	+	o	o	o	o
A8IJ19	512847	Cre12.g500150.t1.1	o	-	+	o	o	o
A8IJR6	512942	Cre12.g504550.t1.2	o	o	o	o	o	o
A8IJY7	512968	Cre12.g505850.t1.2	o	o	-	o	o	o
A8IKC8	513033	Cre12.g508750.t1.2	o	o	-	o	o	o
A8IKE6	513040	Cre12.g509050.t1.1	o	o	-	o	o	o
A8IKI9	513055	Cre12.g509750.t1.2	o	o	-	o	o	o
A8IKK7	513061	Cre12.g510050.t1.2	o	o	-	o	o	o
A8IKZ2	513111	Cre12.g512600.t1.2	o	o	o	o	o	o
A8ILP2	512644	Cre12.g490500.t1.2	o	+	-	+	-	-
A8IMU3	524041	Cre06.g307250.t1.2	o	o	o	o	o	o
A8IQU3	517045	Cre17.g698000.t1.2	o	+	o	o	o	o
A8IR87	517257	Cre17.g708300.t1.2	o	o	o	o	o	o
A8IRE2	517191	Cre17.g705400.t1.2	o	o	o	o	o	o
A8IRG9	521162	Cre03.g184550.t1.2	o	o	-	o	o	o
A8ISZ1	522118	Cre04.g222700.t1.2	o	=	o	o	o	o
A8IT25	522145	Cre04.g223876.t1.1	o	o	o	+	+	+
A8IT25	522145	Cre04.g223850.t1.2	o	o	o	+	+	+
A8ITH8	524751	Cre07.g339150.t1.1	o	o	-	o	o	o
A8ITL8	524782	Cre07.g340650.t1.2	o	o	o	+	o	+
A8IU68	524989	Cre07.g350500.t1.2	o	-	-	o	o	o
A8IUL6	515627	Cre14.g633650.t1.1	o	o	o	+	-	-
A8IUU3	515610	Cre14.g630847.t1.1	o	o	o	o	o	o
A8IUU3	515610	Cre14.g630835.t1.1	o	o	o	o	o	o
A8IV20	515544	Cre14.g627650.t1.1	o	o	o	o	o	o
A8IVB9	513657	Cre12.g538500.t1.1	o	o	-	+	-	o
A8IVG9	523886	Cre12.g537050.t1.2	o	+	-	o	o	o
A8IVV4	513520	Cre12.g532100.t1.2	o	o	-	+	-	-
A8IVW9	513509	Cre12.g531550.t1.2	o	o	-	o	o	o
A8IWF9	521346	Cre03.g192850.t1.2	o	o	-	+	-	+
A8IWS4	521420	Cre03.g196450.t1.2	o	-	o	+	+	+
A8IXG3	521571	Cre03.g203450.t1.2	o	o	o	o	o	o
A8IYC6	512845	Cre12.g547351.t1.2	o	o	o	o	o	o
A8IYP5	514004	Cre12.g554850.t1.2	o	o	+	o	o	o
A8IYU7	513699	Cre12.g540450.t1.2	o	+	o	o	o	o
A8IZL5	525847	Cre09.g388750.t1.2	o	o	-	o	o	o
A8IZX1	525978	Cre09.g394850.t1.2	o	+	o	o	o	o
A8J0N6	519866	Cre02.g141350.t1.2	o	+	o	o	o	o
A8J133	519464	Cre09.g387875.t1.1	o	o	-	+	-	o
A8J137	519469	Cre09.g388097.t1.1	o	o	o	o	o	o
A8J167	519509	Cre09.g389393.t1.1	o	o	-	o	o	o
A8J1P1	526183	Cre09.g404650.t1.2	o	+	o	+	+	+
A8J1S7	521747	Cre18.g749447.t2.1	o	o	-	o	o	o
A8J1S7	521747	Cre18.g749447.t1.1	o	o	-	o	o	o
A8J1X0	523444	Cre06.g278750.t1.2	o	o	o	+	+	+
A8J216	523496	Cre06.g281050.t1.2	o	o	o	+	+	-
A8J264	523554	Cre06.g283950.t1.2	o	+	-	o	o	o
A8J2X8	511881	Cre03.g146167.t1.1	o	o	-	o	o	o
A8J496	525241	Cre08.g360950.t1.1	o	o	o	+	+	-
A8J4E8	512576	Cre08.g362900.t1.1	o	+	o	o	o	o

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J567	513469	Cre12.g529651.t1.1	o	o	o	o	o	o
A8J5T0	513318	Cre12.g522600.t1.2	o	o	-	o	o	o
A8J6D5	517518	Cre17.g720400.t1.1	o	o	+	+	+	+
A8J6D8	517521	Cre17.g720550.t1.2	o	+	o	o	o	o
A8J6H7	517565	Cre17.g722750.t1.2	o	o	-	o	o	o
A8J6P2	514498	Cre17.g726000.t1.2	o	o	o	o	o	o
A8J6R7	517664	Cre17.g727300.t1.2	o	o	o	o	o	o
A8J6S7	517679	Cre17.g727950.t1.1	o	o	o	+	o	-
A8J746	512393	Cre11.g479050.t1.2	o	+	o	o	o	o
A8J785	516980	Cre11.g481450.t1.2	o	o	o	o	o	o
A8J7F6	520617	Cre03.g158900.t1.2	o	o	-	o	o	o
A8J7L5	520531	Cre03.g155001.t1.1	o	-	+	o	o	o
A8J8I8	516183	Cre16.g657650.t1.2	o	+	+	o	o	o
A8J8U1	516310	Cre16.g663500.t1.2	o	o	o	+	+	+
A8JAR0	515273	Cre14.g615100.t1.2	o	-	o	+	+	-
A8JBB4	511953/4	Cre16.g683081.t1.1	o	o	-	o	o	o
A8JBB4	511953/4	Cre16.g682725.t1.1	o	o	-	o	o	o
A8JBN8	518200	Cre03.g199759.t1.1	o	o	-	o	o	o
A8JBN8	518200	Cre03.g199759.t2.1	o	o	-	o	o	o
A8JBU2	518270	Cre03.g207713.t1.1	o	o	o	o	o	o
A8JBU3	518271	Cre03.g207825.t2.1	o	o	-	o	o	o
A8JBU3	518271	Cre03.g207825.t1.1	o	o	-	o	o	o
A8JBU3	518271	Cre03.g207825.t3.1	o	o	-	o	o	o
A8JBX4	518317	Cre03.g213313.t1.1	o	-	o	o	o	o
A8JBY4	520078	Cre11.g467741.t1.1	o	+	o	o	o	o
A8JBY4	520078	Cre11.g467742.t1.1	o	o	o	o	o	o
A8JBZ7	520062	Cre11.g467759.t1.1	o	o	o	o	o	o
A8JC21	512145	Cre11.g467700.t1.1	o	+	o	o	o	o
A8JC47	512167	Cre11.g468750.t1.2	o	o	o	o	o	o
A8JCC0	517890/1/2	Cre17.g738000	o	o	o	o	o	o
A8JCE1	524737	Cre07.g338451.t1.1	o	-	+	+	+	o
A8JCK3	524668	Cre07.g335200.t1.1	o	o	o	+	o	-
A8JD33	510610	Cre01.g000400.t1.2	o	o	o	o	o	o
A8JD33	510610	Cre01.g000350.t1.1	o	o	o	o	o	o
A8JDA2	517409	Cre17.g715500.t1.2	o	+	-	+	-	-
A8JDG4	517336	Cre17.g712100.t1.2	o	o	-	o	o	o
A8JDM6	512742	Cre12.g495300.t1.2	o	o	-	o	o	o
A8JE91	517973	Cre17.g741450.t1.2	o	o	-	+	-	-
A8JE98	517984	Cre17.g741850.t1.2	o	o	-	+	-	-
A8JEG6	512291	Cre02.g095146.t1.1	o	+	-	o	o	o
A8JEI4	512267/8	Cre02.g095122.t1.1	o	+	o	o	o	o
A8JEU4	525480	Cre08.g372100.t1.2	o	o	-	+	-	-
A8JFI7	523787	Cre06.g294950.t1.1	o	-	o	o	o	o
A8JG55	522425	Cre05.g233950.t1.2	o	o	o	o	o	o
A8JG58	522422	Cre05.g233800.t1.2	o	o	-	+	-	o
A8JG73	522405	Cre05.g233050.t1.2	o	o	-	o	o	o
A8JGB0	522352	Cre24.g755097.t1.1	o	-	+	+	+	o
A8JGP7	519416	Cre02.g120301.t1.1	o	+	+	+	o	-
A8JGY3	514060	Cre12.g557600.t1.1	o	o	-	o	o	o
A8JH98	513127	Cre12.g513200.t1.2	o	+	-	+	-	-
A8JHB4	513147	Cre12.g514050.t1.2	o	o	o	o	o	o
A8JHP2	514128	Cre12.g560800.t1.2	o	o	o	o	o	o

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JIE0	520331	NA	o	NA	-	o	o	o
A8JIU6	515779-82	Cre15.g638550.t1.1	o	o	o	o	o	o
D5LAY0	522849	Cre06.g251500.t1.2	o	o	-	o	o	o
D5LB14	522873	Cre06.g252650.t1.2	o	o	-	o	o	o
O81349	524464	Cre07.g325500.t1.1	o	+	-	+	-	-
O98224	510652	Cre01.g002500.t1.2	o	o	o	o	o	o
P42380	513295	NA	o	NA	o	o	o	o
P93106	521388	Cre03.g194850.t1.2	o	o	o	o	o	o
Q0ZAZ1	520461	Cre03.g151700.t1.2	o	o	o	o	o	o
Q0ZAZ1	520461	Cre03.g151750.t1.2	o	+	o	o	o	o
Q0ZAZ1	520461	Cre03.g151700.t2.1	o	o	o	o	o	o
Q19VH5	510365	Cre10.g456100.t1.2	o	+	-	+	o	-
Q39588	526413	Cre09.g415700.t1.2	o	o	o	+	+	+
Q39603	522821	Cre06.g250100	o	o	-	o	o	o
Q6DN05	524437	Cre07.g324200.t1.2	o	o	-	+	+	+
Q6IV67	513097	Cre12.g511850.t1.2	o	o	o	o	o	o
Q6IYG1	521721	Cre04.g217915.t1.1	o	o	-	+	o	+
Q6QAY0	525603	Cre08.g378050.t1.2	o	o	-	o	o	o
Q6QIV9	521596	Cre03.g204650.t1.2	o	o	o	o	o	o
Q6UKY6	513601	Cre12.g535950.t1.2	o	o	o	o	o	o
Q6UP31	520055	Cre11.g467767.t1.1	o	o	o	o	o	o
Q6V9A8	526209	Cre09.g405850.t1.1	o	o	o	o	o	o
Q7X7A7	522979	Cre06.g257500.t1.2	o	o	o	o	o	o
Q84SA0	513892	Cre12.g549750.t1.2	o	o	o	o	o	o
Q84U22	512402	Cre11.g479500.t1.2	o	o	-	o	o	o
Q84Y02	524874	Cre16.g687900.t1.2	o	o	o	o	o	o
Q8LK22	516919	Cre16.g691850.t1.2	o	o	o	o	o	o
Q8RVB8	510990	Cre01.g018800.t1.2	o	o	o	o	o	o
Q8VZZ5	514684	Cre13.g587050.t1.2	o	o	-	o	o	o
Q948T5	520827	Cre03.g168700.t1.2	o	o	o	o	o	o
Q96550	519346	Cre02.g116750.t1.1	o	o	-	o	o	o
Q96550	519346	Cre02.g116750.t2.1	o	o	-	o	o	o
Q9XGU3	523738	Cre06.g292550.t1.2	o	+	o	o	o	o
A2BCY1	NA	Cre02.g073550.t1.2	o	o	NA	NA	NA	NA
A2PZB5	NA	Cre04.g214502.t1.1	o	+	NA	NA	NA	NA
A2PZC1	NA	Cre06.g272900.t1.2	o	o	NA	NA	NA	NA
A2PZC2	NA	Cre13.g565800.t1.2	o	+	NA	NA	NA	NA
A2PZD0	NA	Cre07.g314650.t1.2	o	o	NA	NA	NA	NA
A2PZD2	NA	Cre09.g387171.t2.1	o	o	NA	NA	NA	NA
A2PZD2	NA	Cre09.g387171.t1.1	o	+	NA	NA	NA	NA
A5YU13	NA	Cre08.g373100.t1.1	o	o	NA	NA	NA	NA
A8HM32	NA	Cre01.g035350.t1.2	o	-	NA	NA	NA	NA
A8HM76	NA	Cre01.g036850.t1.2	o	o	NA	NA	NA	NA
A8HMA8	NA	Cre01.g037850.t1.1	o	o	NA	NA	NA	NA
A8HMC0	NA	Cre01.g038400.t1.2	o	+	NA	NA	NA	NA
A8HMD3	NA	Cre01.g038850.t1.1	o	o	NA	NA	NA	NA
A8HME4	NA	Cre01.g039250.t1.1	o	o	NA	NA	NA	NA
A8HME4	NA	Cre01.g039250.t2.1	o	o	NA	NA	NA	NA
A8HME6	NA	Cre01.g039300.t1.2	o	o	NA	NA	NA	NA
A8HMG7	NA	Cre01.g040000.t1.2	o	o	NA	NA	NA	NA
A8HMQ1	NA	Cre01.g042750.t1.2	o	o	NA	NA	NA	NA
A8HMW1	NA	Cre01.g044400.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8HMX2	NA	Cre01.g044800.t1.2	o	o	NA	NA	NA	NA
A8HN02	NA	Cre01.g045550.t1.2	o	+	NA	NA	NA	NA
A8HN48	NA	Cre01.g047700.t1.2	o	o	NA	NA	NA	NA
A8HN50	NA	Cre01.g047750.t1.2	o	o	NA	NA	NA	NA
A8HN65	NA	Cre01.g048150.t1.2	o	o	NA	NA	NA	NA
A8HN88	NA	Cre01.g048750.t1.2	o	o	NA	NA	NA	NA
A8HN92	NA	Cre01.g048950.t1.2	o	o	NA	NA	NA	NA
A8HNB4	NA	Cre01.g049900.t1.1	o	o	NA	NA	NA	NA
A8HND3	NA	Cre01.g050550.t1.2	o	o	NA	NA	NA	NA
A8HNE3	NA	Cre01.g050850.t2.1	o	o	NA	NA	NA	NA
A8HNE3	NA	Cre01.g050850.t1.2	o	o	NA	NA	NA	NA
A8HNE8	NA	Cre01.g050950.t1.2	o	o	NA	NA	NA	NA
A8HNF4	NA	Cre01.g051100.t1.2	o	o	NA	NA	NA	NA
A8HNJ8	NA	Cre01.g052100.t1.2	o	o	NA	NA	NA	NA
A8HNN4	NA	Cre01.g053000.t1.2	o	o	NA	NA	NA	NA
A8HNR8	NA	Cre14.g616600.t1.2	o	+	NA	NA	NA	NA
A8HNZ1	NA	Cre14.g618550.t1.2	o	o	NA	NA	NA	NA
A8HP06	NA	Cre14.g619133.t1.1	o	o	NA	NA	NA	NA
A8HP10	NA	Cre14.g619300.t1.1	o	o	NA	NA	NA	NA
A8HP17	NA	Cre14.g619550.t1.2	o	o	NA	NA	NA	NA
A8HP17	NA	Cre14.g619550.t2.1	o	o	NA	NA	NA	NA
A8HP55	NA	Cre14.g621450.t1.2	o	o	NA	NA	NA	NA
A8HP58	NA	NA	o	NA	NA	NA	NA	NA
A8HP61	NA	Cre14.g621650.t1.1	o	o	NA	NA	NA	NA
A8HP72	NA	Cre01.g010400.t1.2	o	o	NA	NA	NA	NA
A8HP84	NA	Cre01.g010900.t1.2	o	o	NA	NA	NA	NA
A8HP90	NA	Cre01.g011000.t1.2	o	=	NA	NA	NA	NA
A8HPA0	NA	Cre01.g011500.t1.2	o	o	NA	NA	NA	NA
A8HPF4	NA	Cre01.g013700.t2.1	o	o	NA	NA	NA	NA
A8HPF4	NA	Cre01.g013700.t1.2	o	o	NA	NA	NA	NA
A8HPG7	NA	Cre01.g014300.t2.1	o	+	NA	NA	NA	NA
A8HPG7	NA	Cre01.g014300.t1.1	o	o	NA	NA	NA	NA
A8HPG8	NA	Cre01.g014350.t1.2	o	+	NA	NA	NA	NA
A8HPI1	NA	Cre01.g015000.t1.2	o	o	NA	NA	NA	NA
A8HPJ2	NA	Cre01.g015350.t1.1	o	+	NA	NA	NA	NA
A8HPL4	NA	Cre01.g016300.t1.1	o	o	NA	NA	NA	NA
A8HPL8	NA	Cre01.g016500.t1.1	o	o	NA	NA	NA	NA
A8HPL8	NA	Cre01.g016514.t1.1	o	o	NA	NA	NA	NA
A8HPP0	NA	Cre01.g017750.t1.1	o	o	NA	NA	NA	NA
A8HPP0	NA	Cre01.g017750.t2.1	o	o	NA	NA	NA	NA
A8HPS2	NA	Cre01.g019250.t1.2	o	o	NA	NA	NA	NA
A8HPT4	NA	Cre01.g019850.t1.1	o	o	NA	NA	NA	NA
A8HPV3	NA	NA	o	NA	NA	NA	NA	NA
A8HPW8	NA	Cre01.g021600.t1.2	o	o	NA	NA	NA	NA
A8HPZ8	NA	Cre01.g023350.t1.1	o	o	NA	NA	NA	NA
A8HQ48	NA	Cre01.g025250.t1.1	o	o	NA	NA	NA	NA
A8HQ72	NA	Cre01.g026450.t1.2	o	o	NA	NA	NA	NA
A8HQ77	NA	Cre01.g026600.t1.2	o	o	NA	NA	NA	NA
A8HQ81	NA	Cre01.g027000.t1.2	o	=	NA	NA	NA	NA
A8HQ93	NA	Cre01.g027450.t1.2	o	o	NA	NA	NA	NA
A8HQ97	NA	Cre01.g027800.t1.2	o	o	NA	NA	NA	NA
A8HQE3	NA	Cre01.g029450.t1.1	o	+	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8HQF0	NA	Cre01.g029650.t1.1	o	o	NA	NA	NA	NA
A8HQG3	NA	Cre01.g030050.t1.2	o	o	NA	NA	NA	NA
A8HQG8	NA	Cre01.g030200.t1.2	o	o	NA	NA	NA	NA
A8HQJ3	NA	Cre01.g031004.t1.2	o	-	NA	NA	NA	NA
A8HQT1	NA	Cre01.g033550.t1.2	o	o	NA	NA	NA	NA
A8HQW1	NA	Cre01.g034400.t1.1	o	o	NA	NA	NA	NA
A8HR32	NA	Cre13.g562062.t1.2	o	o	NA	NA	NA	NA
A8HR32	NA	Cre13.g562031.t1.1	o	-	NA	NA	NA	NA
A8HR39	NA	Cre13.g562150.t1.2	o	+	NA	NA	NA	NA
A8HR57	NA	Cre13.g562450.t1.2	o	o	NA	NA	NA	NA
A8HR79	NA	Cre13.g562850.t1.2	o	o	NA	NA	NA	NA
A8HRF8	NA	Cre13.g564250.t1.2	o	+	NA	NA	NA	NA
A8HRN8	NA	Cre13.g565450.t1.2	o	o	NA	NA	NA	NA
A8HRP1	NA	Cre13.g565450.t1.2	o	o	NA	NA	NA	NA
A8HRR9	NA	Cre13.g565850.t1.1	o	o	NA	NA	NA	NA
A8HRS5	NA	Cre13.g566000.t1.2	o	o	NA	NA	NA	NA
A8HRV2	NA	Cre13.g566650.t2.1	o	o	NA	NA	NA	NA
A8HRV2	NA	Cre13.g566650.t1.2	o	o	NA	NA	NA	NA
A8HRZ0	NA	Cre13.g567450.t1.2	o	o	NA	NA	NA	NA
A8HS14	NA	Cre13.g567950.t1.2	o	+	NA	NA	NA	NA
A8HS59	NA	Cre13.g568900.t1.2	o	o	NA	NA	NA	NA
A8HSB2	NA	Cre13.g570050.t1.2	o	o	NA	NA	NA	NA
A8HSS0	NA	Cre13.g572900.t2.1	o	o	NA	NA	NA	NA
A8HSS0	NA	Cre13.g572900.t1.1	o	o	NA	NA	NA	NA
A8HSU1	NA	Cre13.g573250.t1.2	o	o	NA	NA	NA	NA
A8HSU7	NA	Cre13.g573351.t1.2	o	o	NA	NA	NA	NA
A8HSV0	NA	Cre13.g573400.t1.2	o	o	NA	NA	NA	NA
A8HT91	NA	Cre13.g576400.t1.2	o	o	NA	NA	NA	NA
A8HTG4	NA	Cre13.g577850.t1.1	o	o	NA	NA	NA	NA
A8HTG7	NA	Cre13.g577900.t1.2	o	o	NA	NA	NA	NA
A8HTK2	NA	Cre13.g578650.t1.1	o	o	NA	NA	NA	NA
A8HTQ3	NA	Cre13.g579550.t1.1	o	o	NA	NA	NA	NA
A8HTX0	NA	Cre13.g581400.t1.1	o	o	NA	NA	NA	NA
A8HTX3	NA	Cre13.g581450.t1.2	o	o	NA	NA	NA	NA
A8HTX7	NA	Cre13.g581600.t1.2	o	+	NA	NA	NA	NA
A8HTX7	NA	Cre13.g581600.t1.2	o	+	NA	NA	NA	NA
A8HTY0	NA	Cre13.g581650.t1.2	o	o	NA	NA	NA	NA
A8HU27	NA	NA	o	NA	NA	NA	NA	NA
A8HUE0	NA	Cre13.g585150.t1.2	o	o	NA	NA	NA	NA
A8HVF8	NA	Cre06.g274700.t1.1	o	-	NA	NA	NA	NA
A8HVJ9	NA	Cre06.g273700.t1.2	o	o	NA	NA	NA	NA
A8HVK4	NA	Cre06.g273600.t1.2	o	o	NA	NA	NA	NA
A8HVM3	NA	Cre06.g273300.t1.2	o	o	NA	NA	NA	NA
A8HVP2	NA	Cre06.g272950.t1.1	o	o	NA	NA	NA	NA
A8HVP7	NA	Cre06.g272850.t1.2	o	o	NA	NA	NA	NA
A8HVQ1	NA	Cre06.g272800.t1.2	o	o	NA	NA	NA	NA
A8HW56	NA	Cre06.g269950.t1.2	o	+	NA	NA	NA	NA
A8HW92	NA	Cre06.g269450.t1.2	o	+	NA	NA	NA	NA
A8HWI6	NA	Cre06.g267500.t1.2	o	o	NA	NA	NA	NA
A8HWS8	NA	Cre06.g265800.t1.2	o	o	NA	NA	NA	NA
A8HX04	NA	Cre06.g264200.t1.2	o	o	NA	NA	NA	NA
A8HX52	NA	NA	o	NA	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8HX54	NA	Cre06.g263250.t1.1	o	o	NA	NA	NA	NA
A8HX70	NA	Cre06.g262900.t1.2	o	o	NA	NA	NA	NA
A8HX89	NA	Cre06.g262700.t1.1	o	o	NA	NA	NA	NA
A8HXA2	NA	Cre06.g262250.t1.2	o	o	NA	NA	NA	NA
A8HXA6	NA	Cre06.g262100.t1.1	o	o	NA	NA	NA	NA
A8HXC5	NA	Cre06.g261800.t1.2	o	o	NA	NA	NA	NA
A8HXE1	NA	Cre06.g261450.t1.2	o	o	NA	NA	NA	NA
A8HXG5	NA	Cre06.g261000.t1.2	o	o	NA	NA	NA	NA
A8HXL8	NA	Cre06.g259900.t1.2	o	o	NA	NA	NA	NA
A8HXM1	NA	Cre06.g259850.t1.2	o	o	NA	NA	NA	NA
A8HXM5	NA	Cre06.g259800.t1.2	o	o	NA	NA	NA	NA
A8HXS9	NA	Cre06.g258733.t1.1	o	o	NA	NA	NA	NA
A8HXS9	NA	Cre06.g258733.t2.1	o	+	NA	NA	NA	NA
A8HXT4	NA	Cre06.g258700.t1.1	o	+	NA	NA	NA	NA
A8HXW8	NA	Cre06.g257950.t1.2	o	o	NA	NA	NA	NA
A8HXX4	NA	Cre06.g257850.t1.2	o	o	NA	NA	NA	NA
A8HXX8	NA	Cre06.g257700.t1.1	o	+	NA	NA	NA	NA
A8HXZ5	NA	Cre06.g257450.t1.2	o	o	NA	NA	NA	NA
A8HY75	NA	Cre06.g255400.t1.2	o	o	NA	NA	NA	NA
A8HYJ1	NA	Cre06.g252200.t1.2	o	o	NA	NA	NA	NA
A8HYQ1	NA	Cre06.g251100.t1.2	o	o	NA	NA	NA	NA
A8HYQ6	NA	Cre06.g251000.t1.2	o	o	NA	NA	NA	NA
A8HYR2	NA	Cre06.g250902.t1.1	o	o	NA	NA	NA	NA
A8HYU2	NA	Cre06.g250250.t1.2	o	o	NA	NA	NA	NA
A8HZF9	NA	Cre10.g447950.t1.2	o	o	NA	NA	NA	NA
A8HZX1	NA	Cre10.g451350.t1.1	o	o	NA	NA	NA	NA
A8HZZ1	NA	Cre10.g451900.t1.1	o	o	NA	NA	NA	NA
A8I023	NA	Cre10.g452450.t1.2	o	o	NA	NA	NA	NA
A8I093	NA	Cre10.g453900.t1.2	o	o	NA	NA	NA	NA
A8I0C9	NA	Cre10.g454734.t1.1	o	o	NA	NA	NA	NA
A8I0C9	NA	Cre10.g454734.t2.1	o	o	NA	NA	NA	NA
A8I0I1	NA	Cre10.g456200.t1.2	o	o	NA	NA	NA	NA
A8I0M0	NA	Cre10.g457000.t1.1	o	o	NA	NA	NA	NA
A8I0V1	NA	Cre10.g458500.t1.2	o	o	NA	NA	NA	NA
A8I0Y2	NA	Cre10.g459250.t1.2	o	o	NA	NA	NA	NA
A8I0Z4	NA	Cre10.g459500.t1.1	o	-	NA	NA	NA	NA
A8I164	NA	Cre10.g461050.t1.2	o	o	NA	NA	NA	NA
A8I175	NA	Cre10.g461250.t1.1	o	o	NA	NA	NA	NA
A8I1B8	NA	Cre10.g462250.t1.2	o	o	NA	NA	NA	NA
A8I1M5	NA	Cre10.g465550.t1.1	o	o	NA	NA	NA	NA
A8I1Y0	NA	Cre02.g099300.t1.2	o	+	NA	NA	NA	NA
A8I232	NA	Cre02.g098450.t1.2	o	o	NA	NA	NA	NA
A8I247	NA	Cre02.g098250.t1.2	o	o	NA	NA	NA	NA
A8I263	NA	Cre02.g097900.t1.2	o	+	NA	NA	NA	NA
A8I282	NA	Cre02.g097550.t1.2	o	o	NA	NA	NA	NA
A8I297	NA	Cre02.g097400.t1.2	o	o	NA	NA	NA	NA
A8I2E0	NA	Cre02.g096150.t1.2	o	o	NA	NA	NA	NA
A8I2H7	NA	Cre02.g095200.t1.2	o	o	NA	NA	NA	NA
A8I2S1	NA	Cre02.g115400.t1.1	o	+	NA	NA	NA	NA
A8I2T0	NA	Cre02.g115200.t1.2	o	o	NA	NA	NA	NA
A8I2Y8	NA	Cre02.g113550.t1.2	o	o	NA	NA	NA	NA
A8I305	NA	Cre02.g113200.t1.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8I311	NA	Cre02.g113100.t1.2	o	+	NA	NA	NA	NA
A8I363	NA	Cre02.g111800.t1.2	o	o	NA	NA	NA	NA
A8I382	NA	Cre02.g111450.t1.2	o	o	NA	NA	NA	NA
A8I3E9	NA	Cre02.g110350.t1.2	o	o	NA	NA	NA	NA
A8I3J5	NA	Cre02.g109600.t1.2	o	o	NA	NA	NA	NA
A8I3Q0	NA	Cre02.g108450.t1.2	o	o	NA	NA	NA	NA
A8I3W3	NA	Cre02.g107300.t1.2	o	o	NA	NA	NA	NA
A8I495	NA	Cre02.g104900.t1.2	o	o	NA	NA	NA	NA
A8I4A8	NA	Cre02.g104650.t1.1	o	o	NA	NA	NA	NA
A8I4A8	NA	Cre02.g104650.t2.1	o	o	NA	NA	NA	NA
A8I4H4	NA	Cre02.g103550.t1.2	o	o	NA	NA	NA	NA
A8I4R2	NA	Cre02.g101950.t1.2	o	o	NA	NA	NA	NA
A8I4S9	NA	Cre02.g101400.t1.2	o	o	NA	NA	NA	NA
A8I523	NA	Cre05.g241850.t1.2	o	o	NA	NA	NA	NA
A8I528	NA	Cre05.g241950.t1.2	o	o	NA	NA	NA	NA
A8I542	NA	Cre05.g242350.t1.2	o	+	NA	NA	NA	NA
A8I554	NA	Cre05.g242550.t1.2	o	+	NA	NA	NA	NA
A8I5A0	NA	Cre05.g243800.t1.2	o	o	NA	NA	NA	NA
A8I5J8	NA	Cre05.g245900.t1.2	o	o	NA	NA	NA	NA
A8I5K1	NA	Cre05.g245950.t1.1	o	o	NA	NA	NA	NA
A8I5Q6	NA	Cre05.g247400.t1.2	o	o	NA	NA	NA	NA
A8I5Q6	NA	Cre05.g247450.t1.2	o	+	NA	NA	NA	NA
A8I5R9	NA	Cre05.g247851.t1.1	o	-	NA	NA	NA	NA
A8I5S5	NA	Cre05.g247950.t1.2	o	-	NA	NA	NA	NA
A8I5V6	NA	Cre07.g312050.t1.1	o	o	NA	NA	NA	NA
A8I647	NA	Cre07.g314150.t1.2	o	o	NA	NA	NA	NA
A8I6C1	NA	NA	o	NA	NA	NA	NA	NA
A8I6D1	NA	Cre07.g316000.t1.2	o	o	NA	NA	NA	NA
A8I6R4	NA	Cre07.g318750.t1.2	o	o	NA	NA	NA	NA
A8I6T5	NA	Cre07.g319300.t1.1	o	o	NA	NA	NA	NA
A8I6X2	NA	Cre07.g319950.t1.2	o	o	NA	NA	NA	NA
A8I720	NA	Cre07.g320900.t1.2	o	o	NA	NA	NA	NA
A8I741	NA	Cre07.g321300.t1.2	o	o	NA	NA	NA	NA
A8I748	NA	Cre07.g321400.t1.1	o	o	NA	NA	NA	NA
A8I775	NA	Cre07.g321800.t1.2	o	o	NA	NA	NA	NA
A8I7F2	NA	Cre07.g323600.t1.1	o	o	NA	NA	NA	NA
A8I7N8	NA	Cre07.g325400.t1.2	o	o	NA	NA	NA	NA
A8I7T1	NA	Cre02.g080650.t1.2	o	o	NA	NA	NA	NA
A8I7T8	NA	Cre02.g080700.t1.2	o	o	NA	NA	NA	NA
A8I7W0	NA	Cre02.g081050.t1.2	o	+	NA	NA	NA	NA
A8I7X0	NA	Cre02.g081250.t1.2	o	o	NA	NA	NA	NA
A8I826	NA	Cre02.g082250.t1.1	o	o	NA	NA	NA	NA
A8I832	NA	Cre02.g082350.t1.1	o	o	NA	NA	NA	NA
A8I832	NA	Cre02.g082350.t2.1	o	o	NA	NA	NA	NA
A8I848	NA	Cre02.g082877.t1.1	o	o	NA	NA	NA	NA
A8I893	NA	Cre02.g083800.t2.1	o	+	NA	NA	NA	NA
A8I893	NA	Cre02.g083800.t1.2	o	-	NA	NA	NA	NA
A8I8A3	NA	Cre02.g083950.t1.1	o	o	NA	NA	NA	NA
A8I8A6	NA	Cre02.g084000.t1.2	o	o	NA	NA	NA	NA
A8I8E9	NA	Cre02.g084900.t1.2	o	+	NA	NA	NA	NA
A8I8N5	NA	Cre02.g086600.t1.2	o	o	NA	NA	NA	NA
A8I8Q5	NA	Cre02.g087150.t1.2	o	-	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8I8V5	NA	Cre02.g087950.t1.2	o	o	NA	NA	NA	NA
A8I8V6	NA	Cre02.g088000.t1.2	o	o	NA	NA	NA	NA
A8I8X8	NA	Cre02.g088600.t1.2	o	o	NA	NA	NA	NA
A8I8Z1	NA	Cre02.g088850.t1.2	o	o	NA	NA	NA	NA
A8I8Z4	NA	Cre02.g088900.t1.2	o	o	NA	NA	NA	NA
A8I901	NA	Cre02.g089100.t1.2	o	+	NA	NA	NA	NA
A8I972	NA	Cre02.g090850.t1.1	o	o	NA	NA	NA	NA
A8I980	NA	Cre02.g091050.t1.2	o	o	NA	NA	NA	NA
A8I982	NA	Cre02.g091100.t1.2	o	o	NA	NA	NA	NA
A8I9E8	NA	Cre02.g092700.t1.2	o	o	NA	NA	NA	NA
A8I9H5	NA	Cre02.g093450.t1.2	o	o	NA	NA	NA	NA
A8I9I9	NA	Cre02.g093650.t1.2	o	o	NA	NA	NA	NA
A8I9M5	NA	Cre02.g094250.t1.1	o	o	NA	NA	NA	NA
A8I9R1	NA	Cre02.g073200.t1.2	o	+	NA	NA	NA	NA
A8I9S6	NA	Cre02.g073650.t2.1	o	o	NA	NA	NA	NA
A8I9S6	NA	Cre02.g073650.t1.2	o	o	NA	NA	NA	NA
A8I9S9	NA	Cre02.g073700.t1.2	o	o	NA	NA	NA	NA
A8I9S9	NA	Cre02.g073700.t2.1	o	o	NA	NA	NA	NA
A8I9X5	NA	Cre02.g074650.t1.2	o	o	NA	NA	NA	NA
A8I9X5	NA	Cre02.g074626.t1.1	o	o	NA	NA	NA	NA
A8IA18	NA	Cre02.g075700.t1.2	o	o	NA	NA	NA	NA
A8IA39	NA	Cre02.g076250.t1.1	o	o	NA	NA	NA	NA
A8IA42	NA	Cre02.g076300.t1.1	o	o	NA	NA	NA	NA
A8IA45	NA	Cre02.g076350.t1.2	o	o	NA	NA	NA	NA
A8IA86	NA	Cre02.g077300.t1.2	o	o	NA	NA	NA	NA
A8IAA8	NA	Cre02.g077750.t1.2	o	o	NA	NA	NA	NA
A8IAA9	NA	Cre02.g077800.t1.2	o	+	NA	NA	NA	NA
A8IAB5	NA	Cre02.g077850.t1.2	o	+	NA	NA	NA	NA
A8IAC7	NA	Cre02.g078100.t1.1	o	o	NA	NA	NA	NA
A8IAE5	NA	Cre02.g078507.t1.2	o	o	NA	NA	NA	NA
A8IAK1	NA	Cre02.g079550.t1.2	o	o	NA	NA	NA	NA
A8IAK9	NA	Cre02.g079700.t1.2	o	+	NA	NA	NA	NA
A8IAL8	NA	Cre02.g079800.t1.2	o	+	NA	NA	NA	NA
A8IAN1	NA	Cre02.g080200.t1.2	o	o	NA	NA	NA	NA
A8IAN4	NA	Cre02.g080250.t1.2	o	o	NA	NA	NA	NA
A8IAT4	NA	Cre10.g434750.t1.2	o	o	NA	NA	NA	NA
A8IAW5	NA	Cre10.g433950.t1.1	o	o	NA	NA	NA	NA
A8IAY6	NA	Cre10.g433600.t1.2	o	o	NA	NA	NA	NA
A8IB70	NA	Cre10.g432000.t1.2	o	o	NA	NA	NA	NA
A8IBF4	NA	Cre10.g430501.t1.1	o	+	NA	NA	NA	NA
A8IBF6	NA	Cre10.g430501.t1.1	o	+	NA	NA	NA	NA
A8IBG1	NA	Cre10.g430400.t1.2	o	o	NA	NA	NA	NA
A8IBU6	NA	Cre10.g428200.t1.2	o	o	NA	NA	NA	NA
A8IBV4	NA	Cre10.g428050.t1.1	o	o	NA	NA	NA	NA
A8IBY2	NA	Cre10.g427700.t1.2	o	o	NA	NA	NA	NA
A8IC90	NA	Cre10.g424850.t1.2	o	o	NA	NA	NA	NA
A8ICB0	NA	Cre10.g424450.t1.2	o	+	NA	NA	NA	NA
A8ICB3	NA	Cre10.g424400.t1.2	o	o	NA	NA	NA	NA
A8ICC8	NA	Cre10.g424100.t1.2	o	-	NA	NA	NA	NA
A8ICC8	NA	Cre10.g424050.t1.1	o	o	NA	NA	NA	NA
A8ICC8	NA	Cre10.g424000.t1.1	o	o	NA	NA	NA	NA
A8ICE4	NA	Cre10.g423650.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8ICG7	NA	Cre10.g423300.t1.1	o	-	NA	NA	NA	NA
A8ICN7	NA	Cre10.g421600.t1.2	o	-	NA	NA	NA	NA
A8ICT4	NA	Cre10.g420700.t1.2	o	o	NA	NA	NA	NA
A8ICV4	NA	Cre10.g420350.t1.2	o	o	NA	NA	NA	NA
A8ID17	NA	Cre10.g418950.t1.1	o	o	NA	NA	NA	NA
A8ID17	NA	Cre10.g418950.t2.1	o	o	NA	NA	NA	NA
A8ID63	NA	Cre10.g418100.t1.1	o	+	NA	NA	NA	NA
A8IDN1	NA	Cre03.g178500.t1.2	o	o	NA	NA	NA	NA
A8IDN1	NA	Cre03.g178450.t1.2	o	NA	NA	NA	NA	NA
A8IDY0	NA	Cre03.g176250.t1.2	o	o	NA	NA	NA	NA
A8IDY8	NA	Cre03.g176100.t1.2	o	o	NA	NA	NA	NA
A8IE04	NA	Cre03.g175750.t1.2	o	o	NA	NA	NA	NA
A8IE23	NA	Cre03.g175400.t2.1	o	o	NA	NA	NA	NA
A8IE23	NA	Cre03.g175400.t1.2	o	o	NA	NA	NA	NA
A8IED9	NA	Cre03.g172950.t1.2	o	o	NA	NA	NA	NA
A8IEE5	NA	Cre03.g172850.t1.2	o	o	NA	NA	NA	NA
A8IEF7	NA	Cre03.g172500.t1.1	o	o	NA	NA	NA	NA
A8IEN5	NA	Cre03.g170900.t1.1	o	-	NA	NA	NA	NA
A8IF06	NA	Cre03.g168550.t1.1	o	+	NA	NA	NA	NA
A8IF08	NA	Cre03.g168450.t1.2	o	+	NA	NA	NA	NA
A8IFC8	NA	Cre03.g166050.t1.2	o	-	NA	NA	NA	NA
A8IFH4	NA	Cre03.g165000.t1.2	o	o	NA	NA	NA	NA
A8IFJ3	NA	Cre03.g164700.t1.2	o	+	NA	NA	NA	NA
A8IFJ3	NA	Cre03.g164750.t1.2	o	o	NA	NA	NA	NA
A8IFN3	NA	Cre03.g164000.t1.2	o	o	NA	NA	NA	NA
A8IFW1	NA	Cre03.g162550.t1.1	o	+	NA	NA	NA	NA
A8IGI0	NA	Cre10.g435850.t1.2	o	+	NA	NA	NA	NA
A8IGN6	NA	Cre07.g333150.t1.2	o	o	NA	NA	NA	NA
A8IGY1	NA	Cre07.g331900.t1.2	o	o	NA	NA	NA	NA
A8IH03	NA	Cre07.g331550.t1.2	o	o	NA	NA	NA	NA
A8IH07	NA	Cre07.g331500.t1.1	o	o	NA	NA	NA	NA
A8IH77	NA	Cre07.g330250.t1.2	o	+	NA	NA	NA	NA
A8IHB3	NA	Cre07.g329700.t1.2	o	+	NA	NA	NA	NA
A8IHH9	NA	Cre07.g328226.t1.1	o	o	NA	NA	NA	NA
A8IHH9	NA	Cre07.g328200.t1.2	o	+	NA	NA	NA	NA
A8IHL6	NA	Cre07.g327350.t1.2	o	o	NA	NA	NA	NA
A8IHM8	NA	Cre07.g326800.t1.1	o	+	NA	NA	NA	NA
A8IHX1	NA	Cre10.g444700.t1.1	o	-	NA	NA	NA	NA
A8II42	NA	Cre10.g443250.t1.2	o	+	NA	NA	NA	NA
A8II71	NA	Cre10.g442700.t1.2	o	+	NA	NA	NA	NA
A8IIA8	NA	Cre10.g441850.t2.1	o	+	NA	NA	NA	NA
A8IIA8	NA	Cre10.g441850.t1.1	o	+	NA	NA	NA	NA
A8IID0	NA	Cre10.g441400.t1.2	o	+	NA	NA	NA	NA
A8IIK8	NA	Cre10.g439900.t1.2	o	o	NA	NA	NA	NA
A8IIL1	NA	Cre10.g439850.t1.2	o	+	NA	NA	NA	NA
A8IIL8	NA	Cre10.g439700.t1.2	o	o	NA	NA	NA	NA
A8IIP9	NA	Cre10.g439100.t1.2	o	+	NA	NA	NA	NA
A8IJ60	NA	Cre12.g500950.t1.2	o	o	NA	NA	NA	NA
A8IJE5	NA	Cre12.g502300.t1.2	o	+	NA	NA	NA	NA
A8IJK9	NA	Cre12.g503500.t1.1	o	+	NA	NA	NA	NA
A8IJL8	NA	Cre12.g503650.t1.2	o	+	NA	NA	NA	NA
A8IK58	NA	Cre12.g507400.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IK69	NA	Cre11.g467717.t1.1	o	o	NA	NA	NA	NA
A8IK91	NA	Cre12.g508000.t1.1	o	o	NA	NA	NA	NA
A8IK94	NA	Cre12.g508050.t1.2	o	o	NA	NA	NA	NA
A8IK98	NA	Cre12.g508150.t1.2	o	-	NA	NA	NA	NA
A8IKA4	NA	Cre12.g508250.t1.1	o	o	NA	NA	NA	NA
A8IKB4	NA	Cre12.g508500.t1.2	o	o	NA	NA	NA	NA
A8IKD6	NA	Cre12.g508900.t1.2	o	+	NA	NA	NA	NA
A8IKH9	NA	Cre12.g509550.t1.1	o	o	NA	NA	NA	NA
A8IKI6	NA	Cre12.g509700.t1.1	o	o	NA	NA	NA	NA
A8IKN8	NA	Cre12.g510400.t1.1	o	o	NA	NA	NA	NA
A8IKP1	NA	Cre12.g510450.t1.2	o	o	NA	NA	NA	NA
A8IKQ0	NA	Cre12.g510650.t1.2	o	o	NA	NA	NA	NA
A8IKQ6	NA	Cre12.g510800.t1.2	o	+	NA	NA	NA	NA
A8IKQ9	NA	Cre12.g510850.t1.2	o	o	NA	NA	NA	NA
A8IKW6	NA	Cre12.g511900.t1.2	o	-	NA	NA	NA	NA
A8IL08	NA	Cre12.g485800.t1.2	o	o	NA	NA	NA	NA
A8IL29	NA	Cre12.g486250.t1.2	o	o	NA	NA	NA	NA
A8IL32	NA	Cre12.g486300.t1.2	o	+	NA	NA	NA	NA
A8IL88	NA	Cre12.g487450.t1.1	o	o	NA	NA	NA	NA
A8ILC2	NA	Cre12.g488173.t1.2	o	o	NA	NA	NA	NA
A8ILF4	NA	Cre12.g488850.t1.1	o	o	NA	NA	NA	NA
A8ILF4	NA	Cre12.g488850.t2.1	o	o	NA	NA	NA	NA
A8ILG8	NA	Cre12.g489153.t1.1	o	=	NA	NA	NA	NA
A8ILJ9	NA	Cre12.g489700.t1.2	o	+	NA	NA	NA	NA
A8ILK1	NA	Cre12.g489750.t1.2	o	+	NA	NA	NA	NA
A8ILL5	NA	Cre12.g490000.t1.2	o	o	NA	NA	NA	NA
A8ILN4	NA	Cre12.g490350.t1.1	o	o	NA	NA	NA	NA
A8ILT1	NA	Cre12.g491500.t1.2	o	o	NA	NA	NA	NA
A8ILX2	NA	Cre06.g311850.t1.1	o	o	NA	NA	NA	NA
A8ILX2	NA	Cre06.g311900.t1.2	o	o	NA	NA	NA	NA
A8IM71	NA	Cre06.g310750.t1.2	o	+	NA	NA	NA	NA
A8IML7	NA	Cre06.g308850.t1.2	o	o	NA	NA	NA	NA
A8IMN3	NA	Cre06.g308533.t2.1	o	o	NA	NA	NA	NA
A8IMN3	NA	Cre06.g308533.t1.1	o	o	NA	NA	NA	NA
A8IMN5	NA	Cre06.g308500.t1.2	o	+	NA	NA	NA	NA
A8IMP0	NA	Cre06.g308350.t1.2	o	o	NA	NA	NA	NA
A8IMU7	NA	Cre06.g307250.t1.2	o	o	NA	NA	NA	NA
A8IMY5	NA	Cre06.g306601.t1.1	o	o	NA	NA	NA	NA
A8IMZ5	NA	Cre06.g306300.t1.2	o	+	NA	NA	NA	NA
A8IN92	NA	Cre06.g304350.t1.2	o	o	NA	NA	NA	NA
A8IN95	NA	Cre06.g304300.t1.2	o	o	NA	NA	NA	NA
A8INI3	NA	Cre06.g302700.t1.1	o	+	NA	NA	NA	NA
A8INR7	NA	Cre06.g300800.t1.2	o	o	NA	NA	NA	NA
A8INS3	NA	Cre06.g300700.t1.1	o	o	NA	NA	NA	NA
A8IP17	NA	Cre06.g298650.t1.2	o	=	NA	NA	NA	NA
A8IP37	NA	Cre06.g298400.t1.2	o	o	NA	NA	NA	NA
A8IP37	NA	Cre06.g298350.t1.2	o	+	NA	NA	NA	NA
A8IP53	NA	Cre06.g298100.t1.2	o	o	NA	NA	NA	NA
A8IPC9	NA	Cre06.g296600.t1.2	o	o	NA	NA	NA	NA
A8IPE0	NA	Cre06.g296400.t1.2	o	-	NA	NA	NA	NA
A8IPH5	NA	Cre06.g295700.t1.1	o	+	NA	NA	NA	NA
A8IPI7	NA	Cre06.g295450.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IPQ9	NA	Cre17.g705000.t1.2	o	o	NA	NA	NA	NA
A8IQ09	NA	Cre17.g703700.t1.1	o	o	NA	NA	NA	NA
A8IQ62	NA	Cre17.g702800.t1.2	o	o	NA	NA	NA	NA
A8IQA9	NA	Cre17.g702150.t1.2	o	o	NA	NA	NA	NA
A8IQB8	NA	Cre17.g701700.t2.1	o	-	NA	NA	NA	NA
A8IQB8	NA	Cre17.g701700.t1.2	o	o	NA	NA	NA	NA
A8IQC1	NA	Cre17.g701650.t1.2	o	o	NA	NA	NA	NA
A8IQC5	NA	Cre17.g701500.t1.2	o	o	NA	NA	NA	NA
A8IQE3	NA	Cre17.g701200.t2.1	o	o	NA	NA	NA	NA
A8IQE3	NA	Cre17.g701200.t1.2	o	+	NA	NA	NA	NA
A8IQL7	NA	Cre17.g699600.t1.2	o	o	NA	NA	NA	NA
A8IQM0	NA	Cre17.g699550.t1.1	o	o	NA	NA	NA	NA
A8IQQ1	NA	Cre17.g699000.t1.2	o	o	NA	NA	NA	NA
A8IQQ1	NA	Cre17.g699000.t2.1	o	o	NA	NA	NA	NA
A8IQS8	NA	Cre17.g698450.t1.2	o	+	NA	NA	NA	NA
A8IQV5	NA	Cre17.g697450.t1.2	o	+	NA	NA	NA	NA
A8IQZ0	NA	Cre17.g696350.t1.2	o	+	NA	NA	NA	NA
A8IR49	NA	Cre17.g709850.t1.2	o	o	NA	NA	NA	NA
A8IR69	NA	Cre17.g709150.t1.2	o	o	NA	NA	NA	NA
A8IR74	NA	Cre17.g708800.t1.1	o	+	NA	NA	NA	NA
A8IR79	NA	Cre17.g708600.t1.2	o	o	NA	NA	NA	NA
A8IR93	NA	Cre17.g708100.t1.2	o	o	NA	NA	NA	NA
A8IR98	NA	Cre17.g707900.t1.2	o	o	NA	NA	NA	NA
A8IRB3	NA	Cre17.g706950.t1.2	o	-	NA	NA	NA	NA
A8IRH4	NA	Cre03.g184700.t2.1	o	+	NA	NA	NA	NA
A8IRH4	NA	Cre03.g184700.t1.2	o	+	NA	NA	NA	NA
A8IRI7	NA	Cre03.g185300.t1.2	o	o	NA	NA	NA	NA
A8IRK4	NA	Cre03.g185550.t1.2	o	o	NA	NA	NA	NA
A8IRK5	NA	Cre03.g185600.t1.1	o	o	NA	NA	NA	NA
A8IRQ1	NA	Cre03.g187450.t1.2	o	o	NA	NA	NA	NA
A8IRR9	NA	Cre03.g188650.t1.2	o	o	NA	NA	NA	NA
A8IRT2	NA	Cre03.g189250.t1.2	o	o	NA	NA	NA	NA
A8IRT4	NA	Cre03.g189300.t1.1	o	o	NA	NA	NA	NA
A8IRT7	NA	Cre03.g189450.t1.2	o	+	NA	NA	NA	NA
A8IRT7	NA	Cre03.g189500.t1.2	o	o	NA	NA	NA	NA
A8IRT8	NA	Cre03.g189500.t1.2	o	o	NA	NA	NA	NA
A8IRU6	NA	Cre03.g189800.t1.2	o	o	NA	NA	NA	NA
A8IRV0	NA	Cre03.g189950.t1.2	o	+	NA	NA	NA	NA
A8IRV6	NA	Cre03.g190100.t1.1	o	o	NA	NA	NA	NA
A8IRW5	NA	Cre03.g190500.t1.2	o	o	NA	NA	NA	NA
A8IRZ2	NA	Cre16.g680900.t1.2	o	o	NA	NA	NA	NA
A8IS44	NA	Cre16.g682800.t1.2	o	o	NA	NA	NA	NA
A8IS47	NA	Cre16.g682900.t2.1	o	o	NA	NA	NA	NA
A8IS47	NA	Cre16.g682900.t1.2	o	+	NA	NA	NA	NA
A8IS72	NA	Cre16.g683950.t1.2	o	+	NA	NA	NA	NA
A8IS88	NA	Cre16.g684650.t1.2	o	o	NA	NA	NA	NA
A8ISA8	NA	Cre16.g685500.t2.1	o	o	NA	NA	NA	NA
A8ISA8	NA	Cre16.g685500.t1.2	o	o	NA	NA	NA	NA
A8ISE3	NA	Cre16.g687300.t1.2	o	-	NA	NA	NA	NA
A8ISE5	NA	Cre16.g687350.t1.2	o	-	NA	NA	NA	NA
A8ISG3	NA	Cre16.g688050.t1.1	o	+	NA	NA	NA	NA
A8ISJ9	NA	Cre16.g669525.t1.1	o	+	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8ISK2	NA	Cre16.g667750.t1.2	o	o	NA	NA	NA	NA
A8ISN6	NA	Cre04.g218250.t1.2	o	o	NA	NA	NA	NA
A8ISS0	NA	Cre04.g219576.t1.1	o	o	NA	NA	NA	NA
A8ISS0	NA	Cre04.g219600.t1.2	o	o	NA	NA	NA	NA
A8IST1	NA	Cre04.g220200.t1.2	o	o	NA	NA	NA	NA
A8IST1	NA	Cre04.g220200.t2.1	o	o	NA	NA	NA	NA
A8IST1	NA	Cre04.g220200.t3.1	o	o	NA	NA	NA	NA
A8IST3	NA	Cre04.g220350.t2.1	o	o	NA	NA	NA	NA
A8IST3	NA	Cre04.g220350.t1.2	o	+	NA	NA	NA	NA
A8ISV8	NA	Cre04.g221400.t1.2	o	o	NA	NA	NA	NA
A8IT23	NA	Cre04.g223800.t2.1	o	-	NA	NA	NA	NA
A8IT23	NA	Cre04.g223800.t1.2	o	o	NA	NA	NA	NA
A8ITF0	NA	Cre04.g229550.t1.1	o	o	NA	NA	NA	NA
A8ITF3	NA	Cre04.g229700.t1.2	o	o	NA	NA	NA	NA
A8ITJ3	NA	Cre07.g339700.t1.2	o	o	NA	NA	NA	NA
A8ITK6	NA	Cre07.g340200.t1.1	o	o	NA	NA	NA	NA
A8ITS8	NA	Cre07.g343700.t1.2	o	o	NA	NA	NA	NA
A8ITU2	NA	Cre07.g344400.t1.2	o	o	NA	NA	NA	NA
A8ITU6	NA	Cre07.g344550.t1.2	o	+	NA	NA	NA	NA
A8ITV3	NA	Cre07.g344950.t1.2	o	o	NA	NA	NA	NA
A8ITX0	NA	Cre07.g346050.t2.1	o	o	NA	NA	NA	NA
A8ITX0	NA	Cre07.g346050.t1.2	o	o	NA	NA	NA	NA
A8ITX5	NA	Cre07.g346350.t1.2	o	o	NA	NA	NA	NA
A8ITX5	NA	Cre07.g346350.t2.1	o	o	NA	NA	NA	NA
A8ITZ0	NA	Cre07.g347100.t1.2	o	o	NA	NA	NA	NA
A8IU48	NA	Cre07.g349650.t1.2	o	-	NA	NA	NA	NA
A8IUB8	NA	Cre07.g352650.t1.2	o	o	NA	NA	NA	NA
A8IUC3	NA	Cre07.g352850.t1.2	o	o	NA	NA	NA	NA
A8IUL5	NA	Cre14.g633750.t1.1	o	+	NA	NA	NA	NA
A8IUL9	NA	Cre14.g633550.t1.2	o	o	NA	NA	NA	NA
A8IUR4	NA	Cre14.g631900.t1.2	o	+	NA	NA	NA	NA
A8IUU2	NA	Cre14.g630859.t1.1	o	o	NA	NA	NA	NA
A8IUV7	NA	Cre14.g630100.t1.2	o	=	NA	NA	NA	NA
A8IV18	NA	NA	o	NA	NA	NA	NA	NA
A8IV36	NA	Cre14.g626900.t1.2	o	+	NA	NA	NA	NA
A8IV40	NA	Cre14.g626700.t1.2	o	o	NA	NA	NA	NA
A8IV67	NA	Cre14.g625400.t1.1	o	o	NA	NA	NA	NA
A8IVE2	NA	Cre12.g537800.t1.2	o	o	NA	NA	NA	NA
A8IVF0	NA	Cre12.g537581.t1.1	o	o	NA	NA	NA	NA
A8IVG0	NA	Cre12.g537200.t1.2	o	o	NA	NA	NA	NA
A8IVH2	NA	Cre12.g536800.t1.2	o	-	NA	NA	NA	NA
A8IVK1	NA	Cre12.g535851.t1.1	o	o	NA	NA	NA	NA
A8IVM9	NA	Cre12.g534800.t1.1	o	o	NA	NA	NA	NA
A8IVN2	NA	Cre12.g534700.t1.2	o	o	NA	NA	NA	NA
A8IVN6	NA	Cre12.g534600.t1.2	o	+	NA	NA	NA	NA
A8IVR6	NA	Cre12.g533550.t1.1	o	o	NA	NA	NA	NA
A8IVS6	NA	Cre12.g533050.t1.2	o	+	NA	NA	NA	NA
A8IVU0	NA	Cre12.g532550.t1.1	o	o	NA	NA	NA	NA
A8IVY4	NA	Cre12.g531100.t2.1	o	o	NA	NA	NA	NA
A8IVY4	NA	Cre12.g531100.t1.2	o	+	NA	NA	NA	NA
A8IVZ3	NA	Cre12.g530850.t1.1	o	o	NA	NA	NA	NA
A8IVZ3	NA	Cre12.g530850.t3.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IVZ3	NA	Cre12.g530850.t2.1	o	o	NA	NA	NA	NA
A8IVZ9	NA	Cre12.g530650.t2.1	o	-	NA	NA	NA	NA
A8IVZ9	NA	Cre12.g530650.t1.1	o	o	NA	NA	NA	NA
A8IW20	NA	Cre12.g529950.t1.2	o	o	NA	NA	NA	NA
A8IW34	NA	Cre17.g734100.t1.2	o	o	NA	NA	NA	NA
A8IW39	NA	Cre17.g734200.t1.2	o	o	NA	NA	NA	NA
A8IW44	NA	Cre17.g734450.t1.2	o	o	NA	NA	NA	NA
A8IW49	NA	Cre17.g734548.t1.1	o	o	NA	NA	NA	NA
A8IW72	NA	Cre13.g590650.t1.2	o	-	NA	NA	NA	NA
A8IW84	NA	Cre13.g591200.t1.2	o	o	NA	NA	NA	NA
A8IW99	NA	Cre13.g591900.t1.2	o	o	NA	NA	NA	NA
A8IWA6	NA	Cre13.g592200.t1.2	o	o	NA	NA	NA	NA
A8IWB5	NA	Cre13.g592450.t1.2	o	+	NA	NA	NA	NA
A8IWJ3	NA	Cre03.g193750.t1.1	o	o	NA	NA	NA	NA
A8IWJ5	NA	Cre03.g193800.t1.1	o	o	NA	NA	NA	NA
A8IWJ6	NA	Cre03.g193800.t1.1	o	o	NA	NA	NA	NA
A8IWK9	NA	Cre03.g194200.t1.2	o	o	NA	NA	NA	NA
A8IWL3	NA	Cre03.g194350.t1.2	o	o	NA	NA	NA	NA
A8IWL4	NA	Cre03.g194400.t1.2	o	o	NA	NA	NA	NA
A8IWP2	NA	Cre03.g195388.t1.1	o	o	NA	NA	NA	NA
A8IWQ1	NA	Cre03.g195650.t1.2	o	o	NA	NA	NA	NA
A8IWT7	NA	Cre03.g196900.t1.2	o	+	NA	NA	NA	NA
A8IWU0	NA	Cre03.g197000.t1.2	o	o	NA	NA	NA	NA
A8IWV0	NA	Cre03.g197350.t1.2	o	+	NA	NA	NA	NA
A8IWV9	NA	Cre03.g197500.t1.2	o	o	NA	NA	NA	NA
A8IWV9	NA	Cre03.g197500.t2.1	o	o	NA	NA	NA	NA
A8IX19	NA	Cre03.g208000.t1.1	o	o	NA	NA	NA	NA
A8IX41	NA	Cre03.g207700.t1.1	o	+	NA	NA	NA	NA
A8IXD1	NA	Cre03.g204601.t1.1	o	+	NA	NA	NA	NA
A8IXE0	NA	Cre03.g204250.t1.2	o	+	NA	NA	NA	NA
A8IXF1	NA	Cre03.g203850.t1.2	o	o	NA	NA	NA	NA
A8IXF2	NA	Cre03.g203800.t1.2	o	-	NA	NA	NA	NA
A8IXH4	NA	Cre03.g202950.t1.1	o	-	NA	NA	NA	NA
A8IXJ7	NA	Cre03.g202100.t1.1	o	o	NA	NA	NA	NA
A8IXJ7	NA	Cre03.g202150.t1.1	o	o	NA	NA	NA	NA
A8IXQ5	NA	Cre03.g200250.t1.2	o	o	NA	NA	NA	NA
A8IXR5	NA	Cre03.g199900.t1.2	o	+	NA	NA	NA	NA
A8IY09	NA	Cre12.g542950.t1.1	o	+	NA	NA	NA	NA
A8IY09	NA	Cre12.g543000.t1.1	o	-	NA	NA	NA	NA
A8IY20	NA	Cre12.g543350.t3.1	o	+	NA	NA	NA	NA
A8IY20	NA	Cre12.g543350.t1.1	o	+	NA	NA	NA	NA
A8IY20	NA	Cre12.g543350.t2.1	o	o	NA	NA	NA	NA
A8IY20	NA	Cre12.g543400.t1.2	o	o	NA	NA	NA	NA
A8IY50	NA	Cre12.g544450.t1.2	o	=	NA	NA	NA	NA
A8IYG2	NA	Cre12.g549350.t1.2	o	+	NA	NA	NA	NA
A8IYG3	NA	Cre12.g549450.t1.2	o	o	NA	NA	NA	NA
A8IYH1	NA	Cre12.g550400.t1.2	o	o	NA	NA	NA	NA
A8IYJ5	NA	Cre12.g551900.t1.2	o	o	NA	NA	NA	NA
A8IYM0	NA	Cre12.g553250.t1.2	o	o	NA	NA	NA	NA
A8IYP3	NA	Cre12.g554750.t1.2	o	o	NA	NA	NA	NA
A8IYP4	NA	Cre12.g554800.t1.2	o	o	NA	NA	NA	NA
A8IYR0	NA	Cre12.g555450.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8IYR1	NA	NA	o	NA	NA	NA	NA	NA
A8IYR2	NA	Cre12.g555600.t1.1	o	o	NA	NA	NA	NA
A8IYR2	NA	Cre12.g555550.t1.2	o	+	NA	NA	NA	NA
A8IYR7	NA	Cre12.g555951.t1.1	o	o	NA	NA	NA	NA
A8IYS1	NA	Cre12.g556050.t1.2	o	o	NA	NA	NA	NA
A8IYS5	NA	Cre12.g556250.t1.2	o	+	NA	NA	NA	NA
A8IYW8	NA	Cre12.g541800.t1.2	o	o	NA	NA	NA	NA
A8IZ29	NA	Cre08.g382800.t1.2	o	o	NA	NA	NA	NA
A8IZ36	NA	Cre08.g382476.t1.1	o	o	NA	NA	NA	NA
A8IZ36	NA	Cre08.g382500.t1.2	o	=	NA	NA	NA	NA
A8IZ39	NA	Cre08.g382350.t1.2	o	+	NA	NA	NA	NA
A8IZ70	NA	Cre08.g380300.t1.1	o	-	NA	NA	NA	NA
A8IZ72	NA	Cre08.g380201.t1.1	o	o	NA	NA	NA	NA
A8IZ88	NA	Cre08.g379200.t1.1	o	o	NA	NA	NA	NA
A8IZ93	NA	Cre08.g378850.t1.2	o	o	NA	NA	NA	NA
A8IZK3	NA	Cre09.g388200.t1.1	o	o	NA	NA	NA	NA
A8IZL1	NA	Cre09.g388500.t1.2	o	o	NA	NA	NA	NA
A8IZR5	NA	Cre09.g391900.t1.1	o	o	NA	NA	NA	NA
A8IZS7	NA	Cre09.g392500.t1.2	o	o	NA	NA	NA	NA
A8IZU0	NA	Cre09.g393200.t1.2	o	o	NA	NA	NA	NA
A8IZW5	NA	Cre09.g394450.t1.2	o	+	NA	NA	NA	NA
A8IZW6	NA	Cre09.g394550.t1.2	o	-	NA	NA	NA	NA
A8IZX9	NA	Cre09.g395350.t1.1	o	+	NA	NA	NA	NA
A8IZZ9	NA	Cre09.g396650.t1.2	o	o	NA	NA	NA	NA
A8J014	NA	Cre09.g397200.t1.2	o	+	NA	NA	NA	NA
A8J063	NA	Cre09.g392282.t1.1	o	o	NA	NA	NA	NA
A8J087	NA	Cre09.g393358.t1.1	o	o	NA	NA	NA	NA
A8J094	NA	Cre09.g393765.t1.1	o	o	NA	NA	NA	NA
A8J0A1	NA	Cre09.g394102.t1.1	o	+	NA	NA	NA	NA
A8J0A2	NA	Cre09.g394102.t1.1	o	+	NA	NA	NA	NA
A8J0D6	NA	Cre09.g395621.t1.1	o	+	NA	NA	NA	NA
A8J0E4	NA	Cre09.g396213.t1.1	o	o	NA	NA	NA	NA
A8J0I1	NA	Cre09.g397734.t1.1	o	o	NA	NA	NA	NA
A8J0K1	NA	Cre09.g398993.t1.1	o	o	NA	NA	NA	NA
A8J0N1	NA	Cre02.g141100.t1.2	o	o	NA	NA	NA	NA
A8J0N7	NA	Cre02.g141400.t1.2	o	o	NA	NA	NA	NA
A8J0Q1	NA	Cre02.g142150.t1.2	o	o	NA	NA	NA	NA
A8J0Q2	NA	Cre02.g142200.t1.2	o	-	NA	NA	NA	NA
A8J0Q8	NA	Cre02.g142800.t1.2	o	o	NA	NA	NA	NA
A8J0R2	NA	Cre02.g143000.t1.2	o	-	NA	NA	NA	NA
A8J0R6	NA	Cre02.g143200.t1.1	o	o	NA	NA	NA	NA
A8J0U6	NA	Cre02.g144850.t1.2	o	o	NA	NA	NA	NA
A8J0W9	NA	Cre02.g145800.t1.2	o	o	NA	NA	NA	NA
A8J0Z0	NA	Cre02.g147100.t1.1	o	o	NA	NA	NA	NA
A8J103	NA	Cre02.g147850.t1.2	o	o	NA	NA	NA	NA
A8J104	NA	Cre02.g147900.t4.1	o	o	NA	NA	NA	NA
A8J104	NA	Cre02.g147900.t3.1	o	o	NA	NA	NA	NA
A8J104	NA	Cre02.g147900.t2.1	o	-	NA	NA	NA	NA
A8J104	NA	Cre02.g147900.t1.1	o	o	NA	NA	NA	NA
A8J119	NA	Cre09.g387245.t1.1	o	o	NA	NA	NA	NA
A8J122	NA	Cre09.g387393.t1.1	o	+	NA	NA	NA	NA
A8J146	NA	Cre09.g388467.t1.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J152	NA	Cre09.g388726.t2.1	o	o	NA	NA	NA	NA
A8J152	NA	Cre09.g388726.t1.1	o	o	NA	NA	NA	NA
A8J152	NA	Cre09.g388726.t3.1	o	o	NA	NA	NA	NA
A8J173	NA	Cre09.g389689.t1.1	o	o	NA	NA	NA	NA
A8J175	NA	Cre09.g389763.t1.1	o	o	NA	NA	NA	NA
A8J195	NA	Cre09.g390763.t1.1	o	o	NA	NA	NA	NA
A8J1A3	NA	Cre09.g391097.t1.1	o	o	NA	NA	NA	NA
A8J1A3	NA	Cre09.g391097.t2.1	o	o	NA	NA	NA	NA
A8J1B6	NA	Cre09.g386650.t2.1	o	o	NA	NA	NA	NA
A8J1B6	NA	Cre09.g386650.t1.2	o	+	NA	NA	NA	NA
A8J1B8	NA	Cre09.g386550.t1.2	o	+	NA	NA	NA	NA
A8J1C1	NA	Cre09.g386400.t1.2	o	o	NA	NA	NA	NA
A8J1C2	NA	Cre09.g386350.t1.1	o	+	NA	NA	NA	NA
A8J1F7	NA	NA	o	NA	NA	NA	NA	NA
A8J1G8	NA	Cre09.g400650.t1.2	o	o	NA	NA	NA	NA
A8J1H4	NA	Cre09.g400850.t1.1	o	o	NA	NA	NA	NA
A8J1J8	NA	Cre09.g402300.t1.2	o	+	NA	NA	NA	NA
A8J1M5	NA	Cre09.g403800.t1.2	o	o	NA	NA	NA	NA
A8J1M9	NA	Cre09.g404000.t1.1	o	o	NA	NA	NA	NA
A8J1S8	NA	Cre18.g749497.t1.1	o	o	NA	NA	NA	NA
A8J1T4	NA	Cre18.g749847.t2.1	o	o	NA	NA	NA	NA
A8J1T4	NA	Cre18.g749847.t1.1	o	o	NA	NA	NA	NA
A8J1U1	NA	Cre09.g386750.t1.2	o	+	NA	NA	NA	NA
A8J1V0	NA	Cre09.g386743.t1.1	o	+	NA	NA	NA	NA
A8J1X8	NA	Cre06.g279150.t1.2	o	o	NA	NA	NA	NA
A8J214	NA	Cre06.g280950.t1.2	o	o	NA	NA	NA	NA
A8J239	NA	Cre06.g282500.t1.2	o	o	NA	NA	NA	NA
A8J266	NA	Cre06.g284050.t1.2	o	o	NA	NA	NA	NA
A8J282	NA	Cre06.g284900.t1.2	o	o	NA	NA	NA	NA
A8J282	NA	Cre06.g284900.t2.1	o	o	NA	NA	NA	NA
A8J287	NA	Cre06.g285250.t1.2	o	o	NA	NA	NA	NA
A8J290	NA	Cre06.g285401.t1.1	o	+	NA	NA	NA	NA
A8J2A5	NA	Cre06.g286250.t1.2	o	+	NA	NA	NA	NA
A8J2E5	NA	Cre06.g288500.t1.2	o	+	NA	NA	NA	NA
A8J2G4	NA	Cre06.g289550.t1.2	o	o	NA	NA	NA	NA
A8J2I5	NA	Cre06.g290950.t1.2	o	=	NA	NA	NA	NA
A8J2J0	NA	Cre06.g291250.t1.1	o	o	NA	NA	NA	NA
A8J2L0	NA	NA	o	o	NA	NA	NA	NA
A8J2L7	NA	Cre03.g145747.t1.1	o	o	NA	NA	NA	NA
A8J2N7	NA	Cre03.g150900.t1.2	o	-	NA	NA	NA	NA
A8J2S3	NA	Cre03.g148950.t1.2	o	o	NA	NA	NA	NA
A8J2W0	NA	Cre03.g146527.t1.1	o	-	NA	NA	NA	NA
A8J2W4	NA	Cre03.g146447.t1.1	o	o	NA	NA	NA	NA
A8J2X6	NA	Cre03.g146187.t1.1	o	+	NA	NA	NA	NA
A8J2Z6	NA	Cre03.g143967.t1.1	o	-	NA	NA	NA	NA
A8J328	NA	Cre03.g145247.t1.1	o	o	NA	NA	NA	NA
A8J355	NA	Cre03.g144627.t1.1	o	o	NA	NA	NA	NA
A8J383	NA	Cre08.g365400.t1.2	o	o	NA	NA	NA	NA
A8J387	NA	Cre03.g143887.t1.1	o	o	NA	NA	NA	NA
A8J3A4	NA	Cre16.g676197.t1.1	o	+	NA	NA	NA	NA
A8J3B2	NA	Cre16.g674964.t1.1	o	o	NA	NA	NA	NA
A8J3C4	NA	Cre16.g673617.t1.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J3D3	NA	Cre16.g672385.t1.1	o	o	NA	NA	NA	NA
A8J3F7	NA	Cre16.g672650.t1.2	o	o	NA	NA	NA	NA
A8J3F8	NA	Cre16.g672750.t1.2	o	+	NA	NA	NA	NA
A8J3K3	NA	Cre16.g674950.t1.2	o	o	NA	NA	NA	NA
A8J3L2	NA	Cre16.g675450.t3.1	o	o	NA	NA	NA	NA
A8J3L2	NA	Cre16.g675450.t2.1	o	o	NA	NA	NA	NA
A8J3L2	NA	Cre16.g675450.t1.2	o	o	NA	NA	NA	NA
A8J3L3	NA	Cre16.g675500.t1.2	o	o	NA	NA	NA	NA
A8J3L5	NA	Cre16.g675550.t3.1	o	+	NA	NA	NA	NA
A8J3L5	NA	Cre16.g675550.t2.1	o	o	NA	NA	NA	NA
A8J3L5	NA	Cre16.g675550.t1.2	o	o	NA	NA	NA	NA
A8J3L9	NA	Cre16.g675650.t1.2	o	-	NA	NA	NA	NA
A8J3P5	NA	Cre16.g677000.t1.2	o	o	NA	NA	NA	NA
A8J3R3	NA	Cre16.g677950.t1.1	o	-	NA	NA	NA	NA
A8J3U9	NA	Cre16.g680000.t1.2	o	o	NA	NA	NA	NA
A8J3Y3	NA	Cre08.g364700.t2.1	o	o	NA	NA	NA	NA
A8J3Y3	NA	Cre08.g364700.t1.2	o	NA	NA	NA	NA	NA
A8J3Z3	NA	Cre08.g367500.t1.1	o	o	NA	NA	NA	NA
A8J429	NA	Cre08.g367300.t1.1	o	o	NA	NA	NA	NA
A8J431	NA	Cre08.g367500.t1.1	o	o	NA	NA	NA	NA
A8J431	NA	Cre08.g367400.t1.1	o	o	NA	NA	NA	NA
A8J434	NA	Cre08.g367600.t1.2	o	-	NA	NA	NA	NA
A8J449	NA	Cre08.g368450.t2.1	o	o	NA	NA	NA	NA
A8J449	NA	Cre08.g368450.t1.2	o	o	NA	NA	NA	NA
A8J4D3	NA	Cre08.g362450.t1.2	o	o	NA	NA	NA	NA
A8J4I2	NA	Cre09.g408200.t1.2	o	+	NA	NA	NA	NA
A8J4I9	NA	Cre09.g408600.t1.2	o	o	NA	NA	NA	NA
A8J4L1	NA	Cre09.g409350.t1.2	o	o	NA	NA	NA	NA
A8J4M8	NA	Cre09.g410250.t1.1	o	o	NA	NA	NA	NA
A8J4N7	NA	Cre09.g410600.t1.1	o	o	NA	NA	NA	NA
A8J4Q7	NA	Cre09.g411200.t1.2	o	o	NA	NA	NA	NA
A8J4R9	NA	Cre09.g411900.t1.1	o	o	NA	NA	NA	NA
A8J4S1	NA	Cre09.g412100.t1.2	o	o	NA	NA	NA	NA
A8J4X1	NA	Cre09.g414200.t1.2	o	o	NA	NA	NA	NA
A8J4Z4	NA	Cre09.g415550.t1.2	o	o	NA	NA	NA	NA
A8J500	NA	Cre09.g415800.t1.2	o	o	NA	NA	NA	NA
A8J503	NA	Cre09.g415950.t1.2	o	o	NA	NA	NA	NA
A8J506	NA	Cre09.g416050.t1.2	o	+	NA	NA	NA	NA
A8J513	NA	Cre09.g416350.t1.2	o	o	NA	NA	NA	NA
A8J524	NA	Cre09.g416750.t1.2	o	+	NA	NA	NA	NA
A8J525	NA	Cre09.g416800.t1.2	o	o	NA	NA	NA	NA
A8J538	NA	Cre09.g417200.t1.2	o	+	NA	NA	NA	NA
A8J597	NA	Cre12.g528750.t1.2	o	o	NA	NA	NA	NA
A8J599	NA	Cre12.g528700.t1.2	o	o	NA	NA	NA	NA
A8J5K0	NA	Cre12.g525400.t1.1	o	o	NA	NA	NA	NA
A8J5K0	NA	Cre12.g525350.t1.1	o	o	NA	NA	NA	NA
A8J5K4	NA	Cre12.g525200.t1.2	o	o	NA	NA	NA	NA
A8J5K6	NA	Cre12.g525150.t1.2	o	o	NA	NA	NA	NA
A8J5N1	NA	Cre12.g524400.t1.2	o	o	NA	NA	NA	NA
A8J5N6	NA	Cre12.g524300.t1.2	o	+	NA	NA	NA	NA
A8J5P0	NA	Cre12.g524100.t1.2	o	o	NA	NA	NA	NA
A8J5P7	NA	Cre12.g523850.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J5Q8	NA	Cre12.g523300.t1.2	o	o	NA	NA	NA	NA
A8J5R2	NA	Cre12.g523150.t1.2	o	+	NA	NA	NA	NA
A8J5T5	NA	Cre12.g522450.t1.2	o	+	NA	NA	NA	NA
A8J5Y7	NA	Cre12.g520600.t1.2	o	o	NA	NA	NA	NA
A8J5Z0	NA	Cre12.g520500.t1.1	o	o	NA	NA	NA	NA
A8J5Z0	NA	Cre12.g520550.t1.2	o	+	NA	NA	NA	NA
A8J605	NA	Cre12.g520200.t1.1	o	o	NA	NA	NA	NA
A8J633	NA	Cre12.g519350.t1.2	o	o	NA	NA	NA	NA
A8J635	NA	Cre12.g519300.t1.2	o	o	NA	NA	NA	NA
A8J664	NA	Cre12.g518650.t1.2	o	o	NA	NA	NA	NA
A8J6A8	NA	Cre12.g517100.t1.1	o	o	NA	NA	NA	NA
A8J6C7	NA	Cre17.g720050.t1.2	o	o	NA	NA	NA	NA
A8J6E2	NA	Cre17.g720750.t1.2	o	o	NA	NA	NA	NA
A8J6F2	NA	Cre17.g721300.t1.2	o	o	NA	NA	NA	NA
A8J6G8	NA	Cre17.g722150.t1.2	o	o	NA	NA	NA	NA
A8J6J6	NA	Cre17.g723650.t1.2	o	-	NA	NA	NA	NA
A8J6K9	NA	Cre17.g724350.t1.2	o	+	NA	NA	NA	NA
A8J6M9	NA	Cre17.g725350.t2.1	o	o	NA	NA	NA	NA
A8J6M9	NA	Cre17.g725350.t1.1	o	o	NA	NA	NA	NA
A8J6R9	NA	Cre17.g727350.t1.2	o	o	NA	NA	NA	NA
A8J6V1	NA	Cre17.g728800.t1.2	o	o	NA	NA	NA	NA
A8J6Y3	NA	Cre11.g476550.t1.2	o	o	NA	NA	NA	NA
A8J709	NA	Cre11.g477300.t1.1	o	+	NA	NA	NA	NA
A8J724	NA	Cre11.g477850.t1.1	o	o	NA	NA	NA	NA
A8J742	NA	Cre11.g478800.t1.2	o	+	NA	NA	NA	NA
A8J768	NA	Cre11.g480150.t1.2	o	o	NA	NA	NA	NA
A8J7C8	NA	Cre03.g160500.t1.2	o	o	NA	NA	NA	NA
A8J7F8	NA	Cre03.g158800.t2.1	o	o	NA	NA	NA	NA
A8J7F8	NA	Cre03.g158800.t1.1	o	o	NA	NA	NA	NA
A8J7F8	NA	Cre03.g158800.t3.1	o	o	NA	NA	NA	NA
A8J7F9	NA	Cre03.g158750.t1.1	o	o	NA	NA	NA	NA
A8J7H3	NA	Cre03.g158000.t1.2	o	+	NA	NA	NA	NA
A8J7H8	NA	Cre03.g157700.t1.2	o	o	NA	NA	NA	NA
A8J7I3	NA	Cre03.g157300.t1.2	o	o	NA	NA	NA	NA
A8J7I9	NA	Cre03.g156950.t1.1	o	o	NA	NA	NA	NA
A8J7J2	NA	Cre03.g156750.t1.2	o	+	NA	NA	NA	NA
A8J7J5	NA	Cre03.g156050.t1.2	o	o	NA	NA	NA	NA
A8J7L6	NA	Cre03.g154950.t1.2	o	o	NA	NA	NA	NA
A8J7M3	NA	Cre03.g154550.t1.1	o	+	NA	NA	NA	NA
A8J7M9	NA	Cre03.g154250.t1.1	o	-	NA	NA	NA	NA
A8J7T7	NA	Cre09.g407700.t1.2	o	o	NA	NA	NA	NA
A8J7X9	NA	Cre09.g401886.t1.1	o	o	NA	NA	NA	NA
A8J7Z6	NA	Cre09.g402775.t1.1	o	o	NA	NA	NA	NA
A8J807	NA	Cre09.g403293.t1.1	o	o	NA	NA	NA	NA
A8J825	NA	Cre09.g406050.t1.2	o	o	NA	NA	NA	NA
A8J827	NA	Cre09.g406200.t1.2	o	+	NA	NA	NA	NA
A8J853	NA	Cre36.g759747.t1.1	o	o	NA	NA	NA	NA
A8J8A6	NA	Cre05.g237700.t1.2	o	o	NA	NA	NA	NA
A8J8B2	NA	Cre05.g237450.t1.2	o	o	NA	NA	NA	NA
A8J8B3	NA	Cre05.g237400.t1.2	o	o	NA	NA	NA	NA
A8J8D7	NA	Cre05.g236050.t1.1	o	o	NA	NA	NA	NA
A8J8F3	NA	Cre16.g655600.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8J8I4	NA	Cre16.g657350.t1.2	o	o	NA	NA	NA	NA
A8J8J8	NA	Cre16.g658526.t1.1	o	-	NA	NA	NA	NA
A8J8J8	NA	Cre16.g658450.t1.2	o	o	NA	NA	NA	NA
A8J8L3	NA	Cre16.g659350.t1.2	o	o	NA	NA	NA	NA
A8J8M3	NA	Cre16.g659850.t1.2	o	o	NA	NA	NA	NA
A8J8M9	NA	Cre16.g660150.t1.2	o	o	NA	NA	NA	NA
A8J8P4	NA	Cre16.g661050.t1.2	o	o	NA	NA	NA	NA
A8J8S2	NA	Cre16.g662600.t1.2	o	o	NA	NA	NA	NA
A8J8V6	NA	Cre06.g278255.t1.1	o	-	NA	NA	NA	NA
A8J8W6	NA	Cre06.g278239.t2.1	o	o	NA	NA	NA	NA
A8J8W6	NA	Cre06.g278239.t1.1	o	+	NA	NA	NA	NA
A8J8W9	NA	Cre06.g278236.t1.1	o	o	NA	NA	NA	NA
A8J8Y1	NA	Cre06.g278222.t1.1	o	o	NA	NA	NA	NA
A8J8Z2	NA	Cre06.g278210.t1.1	o	o	NA	NA	NA	NA
A8J906	NA	Cre06.g278195.t1.1	o	o	NA	NA	NA	NA
A8J914	NA	Cre06.g278185.t1.1	o	+	NA	NA	NA	NA
A8J927	NA	Cre06.g278169.t1.1	o	o	NA	NA	NA	NA
A8J933	NA	Cre06.g278163.t1.1	o	o	NA	NA	NA	NA
A8J940	NA	Cre06.g278154.t1.1	o	o	NA	NA	NA	NA
A8J942	NA	Cre06.g278148.t1.1	o	o	NA	NA	NA	NA
A8J951	NA	Cre06.g278135.t1.1	o	o	NA	NA	NA	NA
A8J955	NA	Cre06.g278125.t1.1	o	o	NA	NA	NA	NA
A8J989	NA	Cre06.g278087.t1.1	o	o	NA	NA	NA	NA
A8J9A9	NA	Cre16.g654500.t1.1	o	o	NA	NA	NA	NA
A8J9C6	NA	Cre16.g653300.t1.2	o	o	NA	NA	NA	NA
A8J9C6	NA	Cre16.g653338.t1.1	o	+	NA	NA	NA	NA
A8J9C6	NA	Cre16.g653350.t1.2	o	o	NA	NA	NA	NA
A8J9D9	NA	Cre16.g652550.t1.2	o	o	NA	NA	NA	NA
A8J9E6	NA	Cre16.g652100.t1.2	o	o	NA	NA	NA	NA
A8J9E9	NA	Cre16.g651900.t1.1	o	o	NA	NA	NA	NA
A8J9E9	NA	Cre16.g651923.t1.2	o	+	NA	NA	NA	NA
A8J9F3	NA	Cre16.g651750.t1.2	o	o	NA	NA	NA	NA
A8J9F6	NA	Cre16.g651550.t1.2	o	o	NA	NA	NA	NA
A8J9H2	NA	Cre16.g650800.t1.1	o	+	NA	NA	NA	NA
A8J9H8	NA	Cre16.g650550.t2.1	o	+	NA	NA	NA	NA
A8J9H8	NA	Cre16.g650550.t1.2	o	+	NA	NA	NA	NA
A8J9K0	NA	Cre16.g649400.t1.2	o	o	NA	NA	NA	NA
A8J9K1	NA	Cre16.g649350.t1.1	o	+	NA	NA	NA	NA
A8J9S7	NA	Cre04.g214500.t1.1	o	o	NA	NA	NA	NA
A8J9T5	NA	Cre04.g214150.t1.1	o	o	NA	NA	NA	NA
A8J9W0	NA	Cre04.g211800.t1.2	o	o	NA	NA	NA	NA
A8J9X1	NA	N/A	o	NA	NA	NA	NA	NA
A8J9Y1	NA	Cre11.g467689.t1.1	o	o	NA	NA	NA	NA
A8JA41	NA	Cre11.g467617.t1.1	o	o	NA	NA	NA	NA
A8JA68	NA	Cre16.g687518.t1.1	o	o	NA	NA	NA	NA
A8JAG1	NA	Cre16.g690431.t2.1	o	+	NA	NA	NA	NA
A8JAG1	NA	Cre16.g690431.t1.1	o	+	NA	NA	NA	NA
A8JAH1	NA	Cre16.g689087.t1.1	o	o	NA	NA	NA	NA
A8JAK5	NA	Cre14.g611950.t1.1	o	+	NA	NA	NA	NA
A8JAL6	NA	Cre14.g612450.t1.2	o	-	NA	NA	NA	NA
A8JAP7	NA	Cre14.g614300.t1.2	o	o	NA	NA	NA	NA
A8JAU8	NA	Cre13.g603600.t1.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JAV1	NA	Cre13.g603700.t1.2	o	o	NA	NA	NA	NA
A8JAV1	NA	Cre13.g603750.t1.1	o	-	NA	NA	NA	NA
A8JAV8	NA	NA	o	NA	NA	NA	NA	NA
A8JAX1	NA	Cre13.g605150.t1.2	o	o	NA	NA	NA	NA
A8JB06	NA	Cre13.g607050.t1.2	o	o	NA	NA	NA	NA
A8JB47	NA	Cre16.g669550.t1.2	o	o	NA	NA	NA	NA
A8JB55	NA	Cre16.g670100.t1.1	o	o	NA	NA	NA	NA
A8JBA7	NA	Cre16.g688550.t1.2	o	o	NA	NA	NA	NA
A8JBB3	NA	Cre16.g683081.t1.1	o	o	NA	NA	NA	NA
A8JBC2	NA	Cre16.g679164.t1.1	o	o	NA	NA	NA	NA
A8JBC6	NA	Cre16.g677026.t1.1	o	+	NA	NA	NA	NA
A8JBD0	NA	Cre16.g675602.t1.1	o	o	NA	NA	NA	NA
A8JBF2	NA	Cre16.g691000.t1.2	o	o	NA	NA	NA	NA
A8JBG0	NA	Cre16.g691450.t1.2	o	o	NA	NA	NA	NA
A8JBJ3	NA	Cre16.g693700.t1.2	o	o	NA	NA	NA	NA
A8JBL3	NA	Cre16.g694850.t1.2	o	o	NA	NA	NA	NA
A8JBL6	NA	Cre16.g695050.t1.2	o	o	NA	NA	NA	NA
A8JBL7	NA	Cre16.g695100.t1.1	o	-	NA	NA	NA	NA
A8JBM7	NA	Cre16.g696000.t1.2	o	o	NA	NA	NA	NA
A8JBN7	NA	Cre03.g199647.t1.1	o	+	NA	NA	NA	NA
A8JBQ5	NA	Cre03.g202113.t1.1	o	o	NA	NA	NA	NA
A8JBT7	NA	Cre03.g206929.t1.1	o	o	NA	NA	NA	NA
A8JBX2	NA	Cre03.g212753.t1.1	o	o	NA	NA	NA	NA
A8JBZ2	NA	Cre11.g467755.t1.1	o	o	NA	NA	NA	NA
A8JC04	NA	Cre11.g467770.t1.1	o	o	NA	NA	NA	NA
A8JC14	NA	Cre11.g467350.t1.2	o	o	NA	NA	NA	NA
A8JC18	NA	Cre11.g467550.t1.2	o	+	NA	NA	NA	NA
A8JC30	NA	Cre11.g468300.t1.2	o	o	NA	NA	NA	NA
A8JC42	NA	Cre11.g468550.t1.2	o	o	NA	NA	NA	NA
A8JC46	NA	Cre11.g468700.t1.2	o	o	NA	NA	NA	NA
A8JC51	NA	Cre11.g468950.t1.2	o	o	NA	NA	NA	NA
A8JC54	NA	Cre11.g469150.t1.2	o	o	NA	NA	NA	NA
A8JCA4	NA	Cre17.g737050.t1.2	o	o	NA	NA	NA	NA
A8JCA8	NA	Cre17.g737250.t1.2	o	+	NA	NA	NA	NA
A8JCC6	NA	Cre17.g738300.t1.2	o	o	NA	NA	NA	NA
A8JCE9	NA	Cre07.g338050.t1.2	o	+	NA	NA	NA	NA
A8JCF6	NA	Cre07.g337650.t1.2	o	o	NA	NA	NA	NA
A8JCF6	NA	Cre07.g337700.t1.1	o	o	NA	NA	NA	NA
A8JCG7	NA	Cre07.g337100.t1.2	o	-	NA	NA	NA	NA
A8JCH8	NA	Cre07.g336600.t2.1	o	o	NA	NA	NA	NA
A8JCH8	NA	Cre07.g336600.t1.1	o	o	NA	NA	NA	NA
A8JCJ9	NA	Cre07.g335400.t1.1	o	-	NA	NA	NA	NA
A8JCM2	NA	Cre07.g334250.t1.1	o	o	NA	NA	NA	NA
A8JCP1	NA	NA	o	NA	NA	NA	NA	NA
A8JCP8	NA	Cre01.g054800.t1.2	o	-	NA	NA	NA	NA
A8JCP9	NA	Cre01.g054850.t1.2	o	o	NA	NA	NA	NA
A8JCQ8	NA	Cre01.g055408.t1.1	o	o	NA	NA	NA	NA
A8JCR1	NA	Cre01.g055420.t1.1	o	o	NA	NA	NA	NA
A8JCR6	NA	Cre01.g055453.t1.1	o	o	NA	NA	NA	NA
A8JCT1	NA	Cre01.g062172.t1.1	o	o	NA	NA	NA	NA
A8JCV8	NA	Cre01.g071662.t1.1	o	=	NA	NA	NA	NA
A8JCW1	NA	Cre01.g007850.t1.2	o	+	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JCW5	NA	Cre01.g007700.t1.1	o	o	NA	NA	NA	NA
A8JCW7	NA	Cre01.g007600.t1.1	o	o	NA	NA	NA	NA
A8JCW7	NA	Cre01.g007600.t2.1	o	o	NA	NA	NA	NA
A8JD31	NA	Cre01.g000250.t1.2	o	o	NA	NA	NA	NA
A8JD45	NA	Cre01.g001100.t1.2	o	+	NA	NA	NA	NA
A8JD56	NA	Cre01.g001750.t1.2	o	o	NA	NA	NA	NA
A8JD99	NA	Cre17.g715750.t1.2	o	+	NA	NA	NA	NA
A8JDA7	NA	Cre17.g715250.t1.2	o	o	NA	NA	NA	NA
A8JDC0	NA	Cre17.g714550.t1.2	o	o	NA	NA	NA	NA
A8JDE6	NA	Cre17.g713200.t1.1	o	o	NA	NA	NA	NA
A8JDJ8	NA	Cre04.g217967.t1.1	o	o	NA	NA	NA	NA
A8JDK2	NA	Cre02.g142146.t1.1	o	o	NA	NA	NA	NA
A8JDL5	NA	Cre12.g495951.t1.1	o	+	NA	NA	NA	NA
A8JDL5	NA	Cre12.g495951.t2.1	o	o	NA	NA	NA	NA
A8JDL8	NA	Cre12.g495850.t1.2	o	-	NA	NA	NA	NA
A8JDM1	NA	Cre12.g495650.t1.2	o	+	NA	NA	NA	NA
A8JDN2	NA	Cre12.g494850.t1.2	o	o	NA	NA	NA	NA
A8JDN4	NA	Cre12.g494750.t1.2	o	o	NA	NA	NA	NA
A8JDN8	NA	Cre12.g494450.t1.2	o	o	NA	NA	NA	NA
A8JDP0	NA	Cre12.g494350.t1.2	o	o	NA	NA	NA	NA
A8JDP4	NA	Cre12.g494050.t1.2	o	=	NA	NA	NA	NA
A8JDP6	NA	Cre12.g493950.t1.2	o	o	NA	NA	NA	NA
A8JDQ7	NA	Cre12.g493250.t1.2	o	o	NA	NA	NA	NA
A8JDV2	NA	Cre15.g635600.t1.2	o	o	NA	NA	NA	NA
A8JDV9	NA	Cre15.g635850.t2.1	o	+	NA	NA	NA	NA
A8JDV9	NA	Cre15.g635850.t1.2	o	+	NA	NA	NA	NA
A8JDW2	NA	Cre15.g636050.t1.2	o	-	NA	NA	NA	NA
A8JDY3	NA	Cre05.g234665.t1.1	o	o	NA	NA	NA	NA
A8JE04	NA	Cre05.g234639.t1.1	o	+	NA	NA	NA	NA
A8JE05	NA	Cre05.g234638.t1.1	o	o	NA	NA	NA	NA
A8JE06	NA	Cre05.g234638.t1.1	o	o	NA	NA	NA	NA
A8JE07	NA	Cre05.g234637.t1.1	o	o	NA	NA	NA	NA
A8JE34	NA	Cre15.g644051.t1.1	o	+	NA	NA	NA	NA
A8JE35	NA	Cre48.g761197.t1.1	o	o	NA	NA	NA	NA
A8JE40	NA	Cre15.g643600.t1.2	o	o	NA	NA	NA	NA
A8JE81	NA	Cre17.g740900.t1.2	o	-	NA	NA	NA	NA
A8JE81	NA	Cre17.g740850.t1.2	o	o	NA	NA	NA	NA
A8JE83	NA	Cre17.g741000.t1.2	o	o	NA	NA	NA	NA
A8JE93	NA	Cre17.g741500.t1.1	o	+	NA	NA	NA	NA
A8JEF7	NA	Cre11.g467723.t2.1	o	o	NA	NA	NA	NA
A8JEF7	NA	Cre11.g467723.t1.1	o	o	NA	NA	NA	NA
A8JEM6	NA	Cre04.g217952.t2.1	o	o	NA	NA	NA	NA
A8JEM6	NA	Cre04.g217952.t1.1	o	o	NA	NA	NA	NA
A8JEM6	NA	Cre04.g217951.t1.1	o	o	NA	NA	NA	NA
A8JEP1	NA	Cre04.g217932.t1.1	o	o	NA	NA	NA	NA
A8JEP9	NA	Cre04.g217921.t1.1	o	+	NA	NA	NA	NA
A8JES5	NA	Cre08.g370650.t1.2	o	o	NA	NA	NA	NA
A8JET7	NA	Cre08.g371450.t2.1	o	+	NA	NA	NA	NA
A8JET7	NA	Cre08.g371450.t1.2	o	+	NA	NA	NA	NA
A8JEU0	NA	Cre08.g371650.t1.2	o	o	NA	NA	NA	NA
A8JEV6	NA	Cre08.g372800.t1.1	o	o	NA	NA	NA	NA
A8JEY5	NA	Cre14.g609030.t1.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JF05	NA	Cre11.g467578.t1.1	o	o	NA	NA	NA	NA
A8JF15	NA	Cre11.g467569.t1.1	o	o	NA	NA	NA	NA
A8JF18	NA	Cre01.g007051.t1.2	o	=	NA	NA	NA	NA
A8JF35	NA	Cre11.g467547.t1.1	o	o	NA	NA	NA	NA
A8JF47	NA	Cre11.g467535.t1.1	o	+	NA	NA	NA	NA
A8JF49	NA	Cre11.g467531.t1.1	o	-	NA	NA	NA	NA
A8JF66	NA	Cre16.g666301.t1.2	o	=	NA	NA	NA	NA
A8JF87	NA	Cre16.g665250.t1.2	o	o	NA	NA	NA	NA
A8JF96	NA	Cre16.g664700.t1.2	o	o	NA	NA	NA	NA
A8JFA2	NA	Cre16.g664350.t1.2	o	-	NA	NA	NA	NA
A8JFA9	NA	Cre16.g664050.t2.1	o	-	NA	NA	NA	NA
A8JFA9	NA	Cre16.g664050.t1.2	o	-	NA	NA	NA	NA
A8JFB1	NA	Cre16.g663900.t1.2	o	+	NA	NA	NA	NA
A8JFD6	NA	Cre17.g732250.t1.1	o	+	NA	NA	NA	NA
A8JFE5	NA	Cre17.g731850.t1.2	o	-	NA	NA	NA	NA
A8JFJ1	NA	Cre06.g294750.t1.2	o	+	NA	NA	NA	NA
A8JFJ2	NA	Cre06.g294650.t1.2	o	o	NA	NA	NA	NA
A8JFJ4	NA	Cre06.g294450.t1.1	o	o	NA	NA	NA	NA
A8JFK4	NA	Cre06.g293950.t1.2	o	+	NA	NA	NA	NA
A8JFK6	NA	Cre06.g293850.t1.2	o	o	NA	NA	NA	NA
A8JFM9	NA	Cre06.g292700.t1.2	o	o	NA	NA	NA	NA
A8JFN4	NA	Cre06.g292500.t2.1	o	o	NA	NA	NA	NA
A8JFN4	NA	Cre06.g292500.t3.1	o	o	NA	NA	NA	NA
A8JFN4	NA	Cre06.g292500.t1.2	o	o	NA	NA	NA	NA
A8JFN8	NA	Cre11.g476200.t1.2	o	+	NA	NA	NA	NA
A8JFN8	NA	Cre11.g476250.t1.2	o	o	NA	NA	NA	NA
A8JFQ7	NA	NA	o	NA	NA	NA	NA	NA
A8JFR4	NA	Cre11.g474800.t1.2	o	o	NA	NA	NA	NA
A8JFT3	NA	Cre07.g354200.t1.2	o	+	NA	NA	NA	NA
A8JFW4	NA	Cre07.g356350.t1.1	o	o	NA	NA	NA	NA
A8JFW5	NA	Cre07.g356400.t1.2	o	+	NA	NA	NA	NA
A8JFX7	NA	Cre01.g005850.t1.2	o	o	NA	NA	NA	NA
A8JFZ0	NA	Cre01.g005150.t1.1	o	-	NA	NA	NA	NA
A8JFZ2	NA	Cre01.g005050.t1.1	o	o	NA	NA	NA	NA
A8JFZ4	NA	Cre01.g004950.t1.1	o	o	NA	NA	NA	NA
A8JFZ7	NA	Cre01.g004900.t1.2	o	o	NA	NA	NA	NA
A8JG03	NA	Cre01.g004500.t1.2	o	o	NA	NA	NA	NA
A8JG04	NA	Cre01.g004450.t1.2	o	o	NA	NA	NA	NA
A8JG06	NA	Cre01.g004350.t1.2	o	o	NA	NA	NA	NA
A8JG07	NA	Cre01.g004300.t1.2	o	o	NA	NA	NA	NA
A8JG49	NA	Cre05.g234300.t1.2	o	o	NA	NA	NA	NA
A8JG56	NA	Cre05.g233900.t1.2	o	o	NA	NA	NA	NA
A8JG70	NA	Cre05.g233305.t1.1	o	+	NA	NA	NA	NA
A8JG70	NA	Cre05.g233305.t2.1	o	o	NA	NA	NA	NA
A8JGB1	NA	Cre24.g755197.t1.1	o	o	NA	NA	NA	NA
A8JGB3	NA	Cre05.g230600.t1.1	o	-	NA	NA	NA	NA
A8JGC8	NA	Cre05.g232002.t2.1	o	-	NA	NA	NA	NA
A8JGC8	NA	Cre05.g232002.t1.1	o	o	NA	NA	NA	NA
A8JGD1	NA	Cre05.g232150.t1.2	o	-	NA	NA	NA	NA
A8JGF8	NA	Cre08.g359750.t2.1	o	+	NA	NA	NA	NA
A8JGF8	NA	Cre08.g359750.t1.2	o	o	NA	NA	NA	NA
A8JGI8	NA	Cre12.g498950.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JGJ6	NA	Cre12.g498550.t1.1	o	o	NA	NA	NA	NA
A8JGK1	NA	Cre12.g498250.t1.2	o	o	NA	NA	NA	NA
A8JGK5	NA	Cre12.g498100.t1.2	o	o	NA	NA	NA	NA
A8JGL0	NA	Cre12.g497850.t1.2	o	o	NA	NA	NA	NA
A8JGR1	NA	Cre02.g119550.t1.2	o	+	NA	NA	NA	NA
A8JGS2	NA	Cre02.g118950.t1.2	o	o	NA	NA	NA	NA
A8JGS3	NA	Cre02.g118900.t1.2	o	o	NA	NA	NA	NA
A8JGS4	NA	Cre02.g118850.t1.2	o	o	NA	NA	NA	NA
A8JGS8	NA	Cre02.g118500.t1.1	o	o	NA	NA	NA	NA
A8JGT1	NA	Cre02.g118300.t1.2	o	o	NA	NA	NA	NA
A8JGU7	NA	Cre02.g117500.t1.2	o	o	NA	NA	NA	NA
A8JGV6	NA	Cre12.g559250.t1.2	o	o	NA	NA	NA	NA
A8JGW5	NA	Cre12.g558700.t1.2	o	+	NA	NA	NA	NA
A8JGW7	NA	Cre12.g558600.t1.2	o	o	NA	NA	NA	NA
A8JGX0	NA	Cre12.g558450.t1.2	o	o	NA	NA	NA	NA
A8JGX5	NA	Cre12.g558100.t1.2	o	+	NA	NA	NA	NA
A8JGY8	NA	Cre12.g557250.t1.2	o	o	NA	NA	NA	NA
A8JGY8	NA	Cre12.g557250.t2.1	o	+	NA	NA	NA	NA
A8JH01	NA	Cre07.g325748.t1.1	o	-	NA	NA	NA	NA
A8JH10	NA	Cre07.g325736.t1.1	o	o	NA	NA	NA	NA
A8JH12	NA	Cre07.g325734.t1.1	o	o	NA	NA	NA	NA
A8JH37	NA	Cre03.g180750.t1.2	o	+	NA	NA	NA	NA
A8JH39	NA	Cre03.g180850.t1.2	o	+	NA	NA	NA	NA
A8JH47	NA	Cre03.g181250.t1.1	o	-	NA	NA	NA	NA
A8JH48	NA	Cre03.g181300.t1.2	o	-	NA	NA	NA	NA
A8JH52	NA	Cre03.g181500.t1.2	o	o	NA	NA	NA	NA
A8JH58	NA	Cre03.g182050.t1.2	o	-	NA	NA	NA	NA
A8JH60	NA	Cre03.g182150.t1.2	o	o	NA	NA	NA	NA
A8JH65	NA	Cre03.g182450.t1.2	o	o	NA	NA	NA	NA
A8JH66	NA	Cre03.g182500.t1.2	o	+	NA	NA	NA	NA
A8JH68	NA	Cre03.g182551.t1.2	o	o	NA	NA	NA	NA
A8JH72	NA	Cre03.g182800.t1.1	o	o	NA	NA	NA	NA
A8JH75	NA	Cre03.g182950.t1.2	o	o	NA	NA	NA	NA
A8JH77	NA	Cre03.g183100.t1.2	o	+	NA	NA	NA	NA
A8JH81	NA	Cre03.g183300.t1.1	o	o	NA	NA	NA	NA
A8JH81	NA	Cre03.g183300.t2.1	o	o	NA	NA	NA	NA
A8JH97	NA	Cre12.g513150.t1.2	o	+	NA	NA	NA	NA
A8JHA6	NA	Cre12.g513600.t1.2	o	+	NA	NA	NA	NA
A8JHA9	NA	Cre12.g513750.t1.1	o	o	NA	NA	NA	NA
A8JHB2	NA	Cre12.g513950.t1.2	o	o	NA	NA	NA	NA
A8JHB7	NA	Cre12.g514200.t1.2	o	+	NA	NA	NA	NA
A8JHC3	NA	Cre12.g514500.t1.2	o	=	NA	NA	NA	NA
A8JHC9	NA	Cre12.g514750.t1.2	o	o	NA	NA	NA	NA
A8JHI5	NA	Cre15.g641800.t1.2	o	o	NA	NA	NA	NA
A8JHJ5	NA	Cre15.g641200.t1.2	o	-	NA	NA	NA	NA
A8JHJ9	NA	Cre15.g641000.t2.1	o	o	NA	NA	NA	NA
A8JHJ9	NA	Cre15.g641000.t1.2	o	o	NA	NA	NA	NA
A8JHL1	NA	Cre04.g216600.t1.2	o	o	NA	NA	NA	NA
A8JHL7	NA	Cre04.g216950.t1.2	o	o	NA	NA	NA	NA
A8JHM2	NA	Cre04.g217550.t1.1	o	o	NA	NA	NA	NA
A8JHN9	NA	Cre12.g560950.t1.2	o	o	NA	NA	NA	NA
A8JHQ3	NA	Cre12.g560300.t3.1	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
A8JHQ3	NA	Cre12.g560300.t2.1	o	o	NA	NA	NA	NA
A8JHQ3	NA	Cre12.g560300.t1.2	o	+	NA	NA	NA	NA
A8JHQ7	NA	Cre12.g560150.t1.2	o	o	NA	NA	NA	NA
A8JHU2	NA	Cre12.g484050.t1.2	o	=	NA	NA	NA	NA
A8JHU6	NA	Cre12.g484200.t1.2	o	+	NA	NA	NA	NA
A8JHW7	NA	Cre12.g515650.t1.2	o	+	NA	NA	NA	NA
A8JHY4	NA	Cre12.g516450.t1.2	o	o	NA	NA	NA	NA
A8JI60	NA	Cre19.g750547.t4.1	o	o	NA	NA	NA	NA
A8JI60	NA	Cre19.g750547.t1.1	o	o	NA	NA	NA	NA
A8JI60	NA	Cre19.g750547.t3.1	o	o	NA	NA	NA	NA
A8JI60	NA	Cre19.g750547.t2.1	o	o	NA	NA	NA	NA
A8JI83	NA	Cre07.g357350.t1.2	o	+	NA	NA	NA	NA
A8JI94	NA	Cre07.g357850.t1.2	o	=	NA	NA	NA	NA
A8JIB5	NA	Cre04.g231418.t1.1	o	o	NA	NA	NA	NA
A8JIB7	NA	Cre04.g231222.t1.1	o	o	NA	NA	NA	NA
A8JIB7	NA	Cre04.g231222.t2.1	o	o	NA	NA	NA	NA
A8JIB8	NA	Cre04.g231124.t1.1	o	o	NA	NA	NA	NA
A8JIB9	NA	Cre04.g231026.t1.1	o	o	NA	NA	NA	NA
A8JIC4	NA	Cre08.g358540.t1.1	o	o	NA	NA	NA	NA
A8JID6	NA	Cre08.g358526.t1.1	o	o	NA	NA	NA	NA
A8JIE3	NA	Cre08.g358559.t1.1	o	o	NA	NA	NA	NA
A8JIE3	NA	Cre08.g358558.t1.1	o	-	NA	NA	NA	NA
A8JIE5	NA	Cre08.g358556.t1.1	o	o	NA	NA	NA	NA
A8JIE5	NA	Cre08.g358556.t2.1	o	o	NA	NA	NA	NA
A8JIJ0	NA	Cre17.g718950.t1.2	o	o	NA	NA	NA	NA
A8JIJ1	NA	Cre17.g719000.t1.2	o	+	NA	NA	NA	NA
A8JIN6	NA	Cre06.g276550.t1.2	o	o	NA	NA	NA	NA
A8JIR0	NA	Cre08.g358580.t1.1	o	o	NA	NA	NA	NA
A8JJ55	NA	NA	o	NA	NA	NA	NA	NA
A8JJ80	NA	Cre14.g627200.t1.1	o	o	NA	NA	NA	NA
A8JJE0	NA	Cre17.g713400.t1.2	o	o	NA	NA	NA	NA
A8JJE1	NA	Cre17.g713500.t1.2	o	o	NA	NA	NA	NA
A8JJH8	NA	Cre05.g242300.t1.2	o	o	NA	NA	NA	NA
A8JJS2	NA	Cre02.g127237.t1.1	o	-	NA	NA	NA	NA
A8JJY5	NA	Cre06.g261200.t1.2	o	o	NA	NA	NA	NA
A8JK20	NA	Cre43.g760497.t1.1	o	o	NA	NA	NA	NA
A8JK39	NA	Cre06.g265250.t1.2	o	o	NA	NA	NA	NA
A8JKF2	NA	Cre13.g566650.t2.1	o	o	NA	NA	NA	NA
A8JKF2	NA	Cre13.g566650.t1.2	o	o	NA	NA	NA	NA
B8XSL8	NA	NA	o	NA	NA	NA	NA	NA
C0SPI7	NA	Cre12.g509000.t1.1	o	o	NA	NA	NA	NA
C5HJ22	NA	N/A	o	NA	NA	NA	NA	NA
C5HJB7	NA	NA	o	NA	NA	NA	NA	NA
C5HJD6	NA	NA	o	o	NA	NA	NA	NA
D5LAT9	NA	Cre06.g251450.t1.1	o	o	NA	NA	NA	NA
D5LAZ0	NA	Cre06.g252000.t1.2	o	o	NA	NA	NA	NA
D5LB05	NA	NA	o	NA	NA	NA	NA	NA
E3SC57	NA	Cre10.g417700.t1.2	o	=	NA	NA	NA	NA
H6V961	NA	Cre02.g106400.t1.1	o	o	NA	NA	NA	NA
O20032	NA	NA	o	NA	NA	NA	NA	NA
O22448	NA	Cre10.g458450.t1.1	o	o	NA	NA	NA	NA
O47027	NA	NA	o	NA	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
O48949	NA	Cre02.g088200.t1.2	o	+	NA	NA	NA	NA
O49822	NA	Cre02.g087700.t1.2	o	o	NA	NA	NA	NA
O63075	NA	NA	o	NA	NA	NA	NA	NA
O64985	NA	Cre01.g039300.t1.2	o	o	NA	NA	NA	NA
O65104	NA	Cre01.g002300.t1.2	o	o	NA	NA	NA	NA
O81523	NA	Cre16.g685550.t1.2	o	o	NA	NA	NA	NA
O81648	NA	Cre16.g673001.t2.1	o	o	NA	NA	NA	NA
P05726	NA	NA	o	NA	NA	NA	NA	NA
P07753	NA	NA	o	NA	NA	NA	NA	NA
P09144	NA	NA	o	NA	NA	NA	NA	NA
P11094	NA	NA	o	NA	NA	NA	NA	NA
P14149	NA	NA	o	NA	NA	NA	NA	NA
P22675	NA	Cre01.g021251.t1.1	o	o	NA	NA	NA	NA
P23489	NA	Cre04.g229300.t1.1	o	o	NA	NA	NA	NA
P26565	NA	NA	o	NA	NA	NA	NA	NA
P32974	NA	NA	o	NA	NA	NA	NA	NA
P36495	NA	NA	o	NA	NA	NA	NA	NA
P37255	NA	NA	o	NA	NA	NA	NA	NA
P48267	NA	NA	o	NA	NA	NA	NA	NA
P48268	NA	NA	o	NA	NA	NA	NA	NA
P48270	NA	NA	o	o	NA	NA	NA	NA
P49644	NA	Cre12.g485150.t1.2	o	o	NA	NA	NA	NA
P59775	NA	NA	o	NA	NA	NA	NA	NA
P59776	NA	NA	o	NA	NA	NA	NA	NA
P81831	NA	Cre16.g673852.t1.1	o	o	NA	NA	NA	NA
P83564	NA	Cre02.g078300.t1.1	o	o	NA	NA	NA	NA
Q00471	NA	NA	o	NA	NA	NA	NA	NA
Q06824	NA	NA	o	NA	NA	NA	NA	NA
Q08363	NA	NA	o	NA	NA	NA	NA	NA
Q19VH4	NA	Cre10.g456050.t2.1	o	+	NA	NA	NA	NA
Q19VH4	NA	Cre10.g456050.t1.2	o	o	NA	NA	NA	NA
Q1G2Y1	NA	Cre04.g214501.t1.1	o	+	NA	NA	NA	NA
Q24K71	NA	NA	o	NA	NA	NA	NA	NA
Q27YU5	NA	Cre07.g330200.t1.2	o	+	NA	NA	NA	NA
Q2VA41	NA	Cre12.g552200.t1.2	o	+	NA	NA	NA	NA
Q39576	NA	Cre06.g275900.t1.2	o	o	NA	NA	NA	NA
Q39595	NA	Cre03.g203850.t1.2	o	o	NA	NA	NA	NA
Q3Y8L7	NA	Cre04.g216902.t1.1	o	o	NA	NA	NA	NA
Q3Y8L7	NA	Cre04.g216902.t2.1	o	o	NA	NA	NA	NA
Q42683	NA	Cre10.g436050.t1.2	o	o	NA	NA	NA	NA
Q42685	NA	Cre10.g451950.t1.2	o	o	NA	NA	NA	NA
Q42703	NA	Cre01.g051900.t1.2	o	o	NA	NA	NA	NA
Q4U0V8	NA	Cre02.g092600.t1.2	o	+	NA	NA	NA	NA
Q4U1D9	NA	Cre06.g282000.t1.1	o	-	NA	NA	NA	NA
Q5QEB2	NA	Cre12.g519180.t2.1	o	o	NA	NA	NA	NA
Q5QEB2	NA	Cre12.g519180.t1.1	o	-	NA	NA	NA	NA
Q5S7Y5	NA	Cre01.g029300.t1.2	o	+	NA	NA	NA	NA
Q5VLJ9	NA	Cre12.g549300.t1.2	o	o	NA	NA	NA	NA
Q5W9T2	NA	Cre03.g199535.t1.1	o	+	NA	NA	NA	NA
Q65Z18	NA	Cre07.g339050.t1.2	o	o	NA	NA	NA	NA
Q66T67	NA	Cre12.g514850.t1.2	o	o	NA	NA	NA	NA
Q66YD0	NA	Cre13.g583550.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
Q66YD3	NA	Cre12.g507650.t1.2	o	o	NA	NA	NA	NA
Q66YD3	NA	Cre12.g507650.t2.1	o	+	NA	NA	NA	NA
Q68RJ5	NA	Cre17.g723600.t1.2	o	+	NA	NA	NA	NA
Q68UI8	NA	Cre06.g279900.t1.1	o	o	NA	NA	NA	NA
Q6EMK7	NA	Cre12.g483700.t2.1	o	o	NA	NA	NA	NA
Q6EMK7	NA	Cre12.g483700.t1.2	o	+	NA	NA	NA	NA
Q6J213	NA	Cre12.g509650.t1.2	o	o	NA	NA	NA	NA
Q6QAY1	NA	Cre02.g100200.t1.2	o	o	NA	NA	NA	NA
Q6QAY3	NA	Cre09.g416150.t1.2	o	+	NA	NA	NA	NA
Q6QAY4	NA	Cre14.g617826.t1.1	o	o	NA	NA	NA	NA
Q6QJE1	NA	Cre06.g273750.t1.1	o	o	NA	NA	NA	NA
Q6R2V6	NA	Cre03.g171950.t1.1	o	o	NA	NA	NA	NA
Q6RBZ0	NA	Cre03.g177200.t1.2	o	+	NA	NA	NA	NA
Q6S7R7	NA	Cre09.g415850.t1.2	o	o	NA	NA	NA	NA
Q6UKY5	NA	Cre13.g577100.t1.2	o	=	NA	NA	NA	NA
Q6UP29	NA	Cre03.g146247.t1.1	o	o	NA	NA	NA	NA
Q6V506	NA	Cre10.g434450.t1.2	o	o	NA	NA	NA	NA
Q6V507	NA	Cre07.g327400.t1.1	o	o	NA	NA	NA	NA
Q6V8K6	NA	Cre16.g665364.t1.1	o	o	NA	NA	NA	NA
Q6V9B0	NA	Cre12.g492300.t1.2	o	o	NA	NA	NA	NA
Q6V9B2	NA	Cre10.g422600.t1.1	o	o	NA	NA	NA	NA
Q6VTH1	NA	Cre12.g561550.t1.2	o	o	NA	NA	NA	NA
Q6WEE4	NA	Cre14.g624350.t1.2	o	o	NA	NA	NA	NA
Q6Y682	NA	Cre10.g435800.t1.2	o	o	NA	NA	NA	NA
Q70DX8	NA	Cre09.g394750.t1.2	o	o	NA	NA	NA	NA
Q75NZ1	NA	Cre06.g307500.t1.1	o	o	NA	NA	NA	NA
Q75NZ3	NA	Cre06.g309000.t1.2	o	o	NA	NA	NA	NA
Q75VY7	NA	Cre06.g272650.t1.2	o	o	NA	NA	NA	NA
Q75VY9	NA	Cre11.g467573.t1.1	o	o	NA	NA	NA	NA
Q75VZ0	NA	Cre10.g452050.t1.2	o	+	NA	NA	NA	NA
Q763T6	NA	Cre16.g656400.t1.2	o	-	NA	NA	NA	NA
Q7DM26	NA	Cre06.g283050.t1.2	o	+	NA	NA	NA	NA
Q7Y258	NA	Cre16.g659950.t1.1	o	o	NA	NA	NA	NA
Q7YKX3	NA	NA	o	NA	NA	NA	NA	NA
Q84JV7	NA	Cre12.g539000.t1.2	o	o	NA	NA	NA	NA
Q84UB2	NA	Cre02.g092900.t1.2	o	o	NA	NA	NA	NA
Q84X75	NA	Cre03.g172000.t1.2	o	o	NA	NA	NA	NA
Q84X77	NA	Cre03.g170800.t1.2	o	+	NA	NA	NA	NA
Q84X82	NA	Cre18.g750047.t1.1	o	+	NA	NA	NA	NA
Q8GV23	NA	Cre06.g268600.t1.2	o	o	NA	NA	NA	NA
Q8HTL1	NA	NA	o	NA	NA	NA	NA	NA
Q8HTL2	NA	NA	o	NA	NA	NA	NA	NA
Q8HTL3	NA	NA	o	NA	NA	NA	NA	NA
Q8HTL5	NA	NA	o	NA	NA	NA	NA	NA
Q8HUG9	NA	NA	o	NA	NA	NA	NA	NA
Q8HUH1	NA	NA	o	NA	NA	NA	NA	NA
Q8LKI3	NA	Cre17.g729800.t1.2	o	o	NA	NA	NA	NA
Q8LL91	NA	Cre09.g393150.t1.2	o	o	NA	NA	NA	NA
Q8LPD9	NA	Cre03.g199000.t1.2	o	o	NA	NA	NA	NA
Q8LRU1	NA	Cre09.g387800.t1.2	o	+	NA	NA	NA	NA
Q8S2V8	NA	Cre12.g517150.t1.1	o	o	NA	NA	NA	NA
Q8S4W5	NA	Cre06.g290100.t1.2	o	o	NA	NA	NA	NA

Appendix 2.

UniProt ID	JGI v4 protein ID	JGI v5.5 (Augustus u11.6) ID	LC-MS/MS	WT vs hybrids	Bruegg. C-effect	Fang overall DE	Fang C-effect	Fang S-effect
Q8SAQ6	NA	Cre12.g526800.t1.2	o	o	NA	NA	NA	NA
Q8W4V3	NA	Cre16.g664550.t1.2	o	o	NA	NA	NA	NA
Q93VE0	NA	Cre01.g066917.t1.1	o	o	NA	NA	NA	NA
Q93WL4	NA	Cre04.g232104.t1.1	o	o	NA	NA	NA	NA
Q93Y49	NA	Cre09.g394436.t1.1	o	o	NA	NA	NA	NA
Q93Z22	NA	Cre10.g459200.t1.2	o	o	NA	NA	NA	NA
Q945T2	NA	Cre08.g373100.t1.1	o	o	NA	NA	NA	NA
Q94CJ2	NA	Cre06.g284100.t1.2	o	o	NA	NA	NA	NA
Q94KS3	NA	Cre10.g427250.t1.2	o	o	NA	NA	NA	NA
Q94KV1	NA	Cre06.g272050.t1.2	o	o	NA	NA	NA	NA
Q9AU03	NA	Cre01.g049500.t1.2	o	o	NA	NA	NA	NA
Q9AU05	NA	Cre03.g154350.t1.2	o	o	NA	NA	NA	NA
Q9AXF6	NA	Cre12.g548950.t1.2	o	o	NA	NA	NA	NA
Q9FEH5	NA	Cre01.g002500.t1.2	o	o	NA	NA	NA	NA
Q9FEK6	NA	Cre16.g673650.t1.1	o	o	NA	NA	NA	NA
Q9FXQ1	NA	Cre16.g663200.t1.1	o	o	NA	NA	NA	NA
Q9LLL6	NA	Cre03.g188250.t1.2	o	o	NA	NA	NA	NA
Q9S7V1	NA	Cre02.g085450.t1.2	o	+	NA	NA	NA	NA
Q9SW75	NA	Cre02.g101350.t1.2	o	o	NA	NA	NA	NA
Q9ZSJ4	NA	Cre03.g156900.t1.2	o	+	NA	NA	NA	NA
Q9ZTA7	NA	Cre09.g396300.t1.2	o	+	NA	NA	NA	NA