Durham E-Theses

Targeting a promising new herbicide mode of action: chemical and genetic approaches to elucidate the role of IPC synthase in plants

PINNEH, ELIZABETH,CASSANDRA

Targeting a promising new herbicide mode of action: chemical and genetic approaches to elucidate the role of IPC synthase in plants

Elizabeth Cassandra Ebiere Pinneh

September 2017

Department of Chemistry

<u>Supervisors</u>

Dr. Paul Denny

Biophysical Sciences Institute, Chemistry department,

Durham University, Durham, DH1 3LE, UK,

p.w.denny@durham.ac.uk

+44 191 334 3983

Prof. Patrick Steel

Centre for Bioactive Chemistry, Chemistry Department,

Durham University, Durham, DH1 3LE, UK,

p.g.steel@durham.ac.uk

+44 191 334 2131

<u>Examiners</u>

Prof. Rob Fields

John Innes Centre, Norwich Research Park,

Norwich, NR4 7UH, UK,

rob.field@jic.ac.uk

+44 160 345 0720

Dr. Heather Knight

Department of Biosciences, Durham University,

Stockton Road, Durham, DH1 3LE

+44 191 334 3215

p.h.knight@durham.ac.uk

Targeting a promising new herbicide mode of action: chemical and genetic approaches to elucidate the role of IPC synthase in plants

ELIZABETH C. E. PINNEH

Abstract

Worldwide there are currently 479 recorded unique cases of herbicide resistant weeds, with 251 species (146 dicots and 105 monocots). According to current records, weeds have developed resistance to 22 of the 25 known herbicide sites of action and to 157 different herbicides. Research into sphingolipid synthesis has revealed a potential novel site of action involving the non-mammalian enzyme inositol phosphorylceramide synthase (IPCS). Two approaches to understanding the function of the enzyme in plants have been employed in this study. The genetic approach, which involves the overexpression and knockdown of the three IPCS orthologues in *Arabidopsis thaliana*, has given insights into the global function of the enzyme in plants at a phenotypic and transcriptomic level. Analyses of the RNASeq data show that *At*IPCS is a global negative regulator of plant defense, downregulating genes involved in plant defense against herbivory attack, nematodes, fungal and bacterial pathogens. RNASeq data in conjunction with phenotypic data, indicate that IPCS may play a role in floral transition from the vegetative phase to the reproductive phase, with the transgenic lines displaying an early flowering phenotype. *At*IPCS transcript levels is also shown to affect plant post-embryonic development. In parallel, high throughput screening facilitated the identification of chemical inhibitors of plant IPCS. Using a system of *Arabidopsis thaliana* as a model dicot and *Oryza sativa* as model monocot, two classes of inhibitors were identified and demonstrated differential activity: the phenylamidines and triazinones. In addition, *in vivo* screening of the phenylamidines demonstrated herbicidal activity, indicating that the necessary selectivity for herbicide development is achievable.

Table of contents

Abbreviations

IPCS	Inositol phosphorylceramide
IPCS	Inositol phosphorylceramide synthase
ALS	Actolactate synthase
ACCase	Acetyl CoA carboxylase
PS	Photosystem
PI	Phosophoinositol
EPSP	5-enolpyruvylshikimate-3-phosphate synthase
HPPD	p-Hydroxyphenylpyruvate dioxygenase
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
LCB	Long-chain base
GSL	Glycosphingolipids
VLCFAs	Very long chain fatty acids
GluCer	Glycosylceramide
GIPC	Glycosyl inositol phosphoceramide
AUR1p	AUR1 protein (fungal IPCS)
AtIPCS	*Arabidopsis thaliana* inositol phosphorylceramide synthase
OsIPCS	*Oryza sativa* inositol phosphorylceramide synthase
LmjIPCS	*Leishmania major* inositol phosphorylceramide synthase
FDGlu	Fluorescein di-(β-D-glucopyranoside)
AbA	Auerobasidin A
DMSO	Dimethyl sulfoxide
PDR	Pleotropic drug resistance
BLAST	Basic Local Alignment Search Tool
IC_{50}	Half-maximal Inhibitory Concentration
PCR	Polymerase chain reaction
HTS	High Throughput Screening
Monocot	Monocotyledons
Dicot	Dicotyledons
SOC1	Suppression of overexpression of *CONSTANS* 1
SM	Sphingomyelin
SMS	Sphingomyelin Synthase
5FOA	5-fluoroorotic acid
NBD ceramide	N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-ceramide
NBD-IPC	N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-inositol phosphorylceramide
NBD-phytoceramide	N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)-phytoceramide

Declaration

The work presented herein was carried out by the author in the Department of Chemistry and in the Department of Biosciences at the University of Durham from October 2013-September 2017. Some parts of the work were carried out by the author at Bayer Crop Science, Frankfurt, Germany, from February 2016-June2016 and in March 2017. No part of this work has been submitted for any other degree at this or any other University.

Statement of copyright

Acknowledgements

members) in labs 229 and 001 for your support and friendship, it does mean a lot to me. Last but certainly not least, special thanks to my family for their constant love and support, I certainly could not have done this without your help especially at times when I needed it most.

Dedication

I dedicate this work to my grandparents:

Lucy and Dumbiri Ogbu

Florence Omolabake and Anthony Abari Pinneh

CHAPTER 1

INTRODUCTION

1.1 Thesis synopsis

The aim of this work was to characterise the role of inositol phosphorylceramide synthase (IPCS) in plants, and to identify putative inhibitors of the enzyme, that can be further developed for use as herbicides. In Chapter 1, an introduction to the need for the identification of new herbicide mode of action is outlined, as is the reasoning for pursuing IPCS as an herbicidal target. In this chapter, sphingolipid structural variation and biosynthesis, in plants and mammals is reviewed and known inhibitors of IPCS orthologues in fungi and protozoa are examined. In addition, current understanding of the role of sphingolipids in relation to programmed cell death in plants is presented. In chapter 2, bioinformatics analysis is carried out on the plant IPCSs from different plant species to glean information on how similar or different they are from each other, and to find out biological pathways they could affect by analysis of transcription factor binding sites in the promoter region. In that same chapter, assay parameters for the screening campaign were developed and validated. In chapter 3 and 4, results from the primary and secondary screening, as is the work done to validate *At*IPCS1-3 over-expresser and RNAi transgenic lines in *A. thaliana* is shown. In chapter 5, RNASeq analysis of *At*IPCS1 t-DNA mutant and transgenic lines over-expressing *At*IPCS1, 2 and 3 are presented which reveal the role of plant *At*IPCS. Conclusions and future work are discussed in chapter 6 and in chapter 7 the methods and materials are covered.

1.2 Crop production and herbicide resistance

World food production is heavily reliant on the use of herbicides and pesticides to reduce the detrimental effects of pests and weeds on crop yield. With global population projected to increase by 2 billion by the year 2050[1], there comes the need for increased agricultural production of staple crops which the global community rely on to survive. It is currently estimated that 12.5% of the global population are undernourished[1], this could increase in the coming years due in part to the unfolding effects of climate change on agricultural output. It is also projected that by the 2080s agricultural yield in developing countries will decline by 9% due to high carbon emissions even when the benefit of carbon fertilization is taken into consideration, and on a global scale this decrease in yield will be 3%, with yield decrease being greatest in South East Asia and sub-Saharan Africa[2] (Figure 1-1).

A study conducted by the World Bank study forecasted 50 million more people could be at risk of undernourishment because of climate change by the year 2050[3]; without mitigating measures to combat climate change the number of malnourished people globally is set to grow to 321 million within the 2050 decade and by 391 million in the decade of 2080[4], with the most affected areas being developing countries.

Cereal crops are the most common food source for human consumption. According to the Food and Agricultural Organization of the United Nations (FAO) 2013 report approximately 2.3 billion tonnes of cereals are currently produced globally[1]; roughly 1 billion tonnes is destined for food use, 750 million tonnes will be employed as animal feed, and the remaining 500 million tonnes will be processed for industrial use as seed or wasted[1]. The growth rate of world cereal production fell to 1 percent per annum in the 1990s, down from 1.6 percent in the 1980s and almost 3 percent in the 1970s; between 2000 and 2003, growth was almost zero due to low prices and a surplus in stock[1].
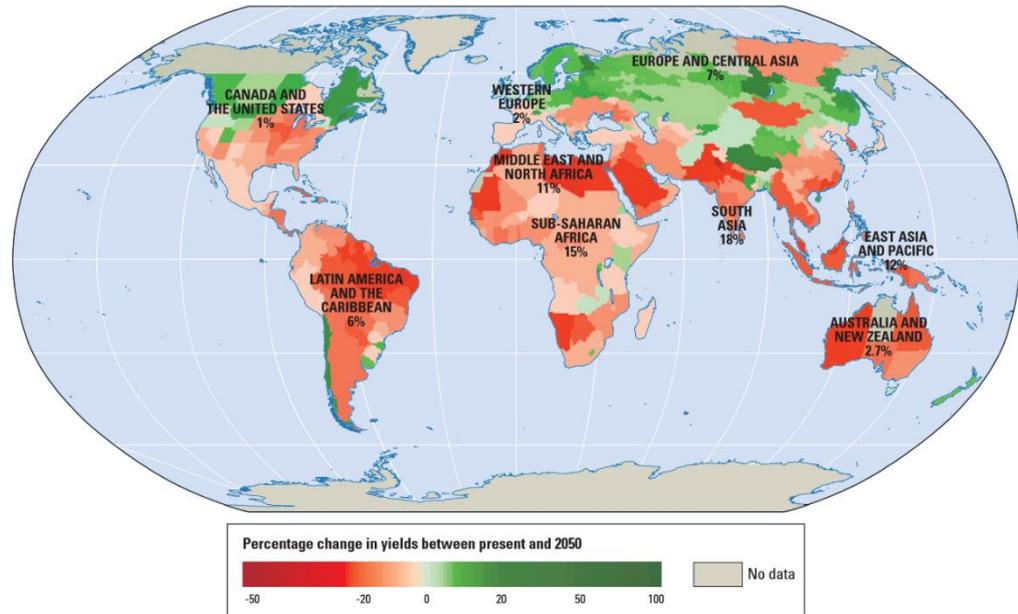
Figure 1-1: Global map of agricultural production projected for the years 2046-2055[2].

The challenge of increasing crop yield globally is further compounded by the decrease in available arable land; currently 40% of arable land is used to produce staple food which is expected to decrease to 30% by 2025 in order to provide land for renewable resources, and to feed companion and farm animals[5]. Another challenge is combating pesticide resistance which encompasses all plant protection agents, including but not restricted to fungicides, insecticides and herbicides. Global usage of pesticides is approximately 2 million tonnes, with 45% of this usage accounting for Europe, 25% in the US and the other 30% in the rest of the world[6]. The reliance on plant protection agents, in particular herbicides, which account for 47.5% of worldwide pesticide consumption[6], mirrors the changes in farming practices from mechanical to chemical solutions for combating weeds. This has resulted in widespread weed resistance to herbicides with deleterious effects on crop yield. The control of weeds up until the end of the 20th century was primarily achieved by physical and mechanical removal of weeds combined with crop rotation[7]. The research conducted in World War II for biological warfare and research into

chemicals that could control insects carrying malaria, paved way for the synthetic chemical industry and modern herbicides[8].

Herbicides are classified by activity, chemical family and site of action. The classification of herbicides by their site of action refers to the target enzyme and the subsequent effects on biochemical pathways as a result of inhibiting that enzyme. There are currently 21 known sites of action (Table 1) with some sites of action being unknown for certain herbicides such as Difenzoquat.

The first herbicide discovered was in 1940 by W. G. Templeman. 2,4-Dichlorophenoxyacetic acid (2,4-D)[9] a plant hormone, was found to selectively kill broad-leaved weeds in the midst of cereal crops[10]. This was followed by the discovery and development of atrazine (Figure 1-2) from the triazine class of herbicides used on broadleaf weeds amongst maize and sugarcane[9]. The introduction of the herbicide glyphosate (N-(phosphonomethyl)glycine) by Monsanto in 1974 under the trade name 'Roundup' revolutionized the use of herbicides for farming in America; this herbicide was found to be non-specific, killing a wide range of weeds including grasses, broadleaf weeds and woody plants[9]. The subsequent introduction of glyphosate resistant crops resulted in its widespread use in both crop and non-crop lands[11] for over three decades. Recently, it has been found that there is widespread glyphosate resistance across America, a consequence of the evolution of glyphosate resistant weed biotypes found in rigid and Italian ryegrass, marestail, goosegrass, common ragweed, waterhemp and velvet leaf[11]. This glyphosate resistance has been found in the majority of soybean, cotton and corn farms and has been reported at 69% of 144 waterhemp population sites sampled in Missouri and 64% of 500 in Iowa in the year 2011-2012[12].

Glyphosate resistance has affected crop yield and profits and has resulted in the use of alternative herbicides to kill resistant biotypes. The cost of using herbicides in the south

Table 1: Classification of herbicides based on site of action, adapted from Heap 2017[2]

Herbicide group	Site of action	Example of Herbicide
ALS inhibitors	Inhibits acetolactate synthase the key enzyme in the biosynthesis of branched amino acids isoleucine, leucine and valine	Chlorsulfuron
Photosystem II inhibitors	Inhibits photosynthesis by binding to the QB binding site on D1 protein of photosystem II complex in the thylakoid membrane	Atrazine
ACCase inhibitors	Inhibits acetyl CoA carboxylase which is involved in the de novo synthesis of fatty acids	Sethoxydim
PSI Electron Diverter	Herbicides accept electrons from photosystem II and are reduced to form herbicide radicals	Paraquat
Synthetic Auxins	Acts similar to endogenous auxins but the mechanism is not well understood; they acidify the cell wall by stimulating the activity of membrane bound ATPase pumps	2,4-D
EPSP synthase inhibitors	Inhibits 5-enolpyruvylshikimate-3-phosphate synthase resulting in depletion of aromatic amino acids needed in biosynthetic pathways involved in growth	Glyphosate
PSII inhibitor (Ureas and amides)	Urea and amides inhibits photosynthesis by binding to the QB binding site on D1 protein of photosystem II complex in the thylakoid membrane	Chlorotoluron
Microtubule inhibitors	Binds to tubulin thereby inhibiting microtubulin polymerization at the assembly end	Trifluralin
Lipid Inhibitors (thiocarbamates)	Inhibitors of several plant processes including biosynthesis of fatty acids, lipids, proteins, isoprenoids, flavanoids and gibberellin	Triallate
PPO inhibitors	Inhibits Protoporphyrinogen oxidase involved in chlorophyll and heme biosynthesis	Oxyfluorfen
Carotenoid biosynthesis (unknown target)	Targets enzymes involved in carotenoid biosynthesis	Amitrole
PSII inhibitors (Nitriles)	Nitriles inhibit photosynthesis by binding to the QB binding site on D1 protein of photosystem II complex in the thylakoid membrane	Bromoxynil
Long chain fatty acid inhibitors	Inhibit the synthesis of very long chain fatty acids (VLCFA)	Butachlor
Carotenoid biosynthesis inhibitors	Inhibition of phytoene desaturase, an enzyme involved in carotenoid biosynthesis	Diflufenican
Antimicrotubule mitotic disrupter	Unknown mode of action	Flamprop-methyl
HPPD inhibitors	Inhibits p-Hydroxyphenylpyruvate dioxygenase	Isoxaflutole
Glutamine synthase inhibitors	Inhibits glutamine synthase which converts ammonia and glutamate to glutamine.	Glufosinate-ammonium
Cellulose inhibitors	Inhibits enzymes involved in cell wall biosynthesis	Dichlobenil
Mitosis inhibitors	inhibits cell division, microtubule organisation and polymerization	Propham
Cell elongation inhibitors	Unknown mode of action	Difenzoquat
Nucleic acid inhibitors	Unknown mode of action	Monosodium methanearsonate (MSMA)

of America has increased from \$50-\$75 per hectare to about \$370 per hectare as a result of glyphosate resistance[12].
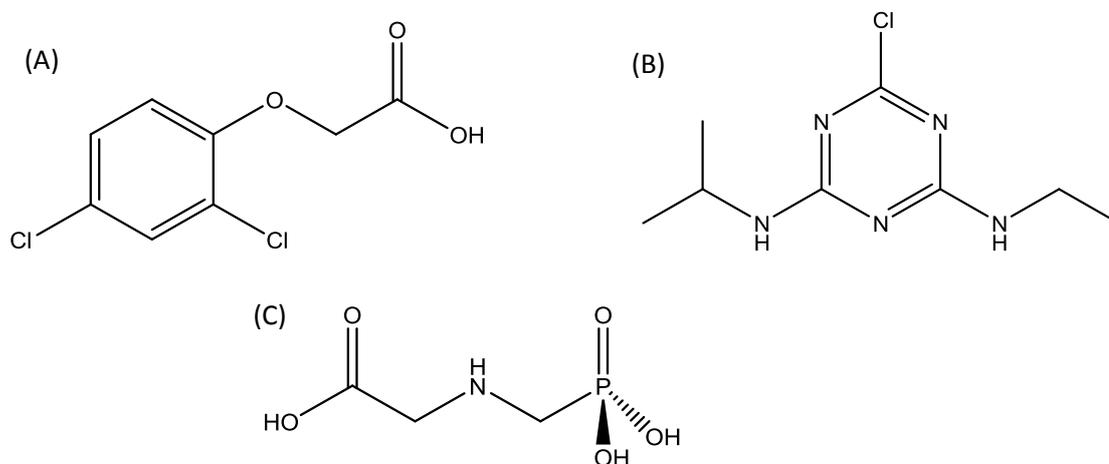
Figure 1-2: Chemical structures of first generation herbicides (A) 2,4-Dichlorophenoxyacetic acid (2,4-D) (B) Atrazine (C) Glyphosate.

Areas farming cotton have been heavily affected resulting in a shift away from growing cotton with a 60% and 70% decrease in area growing cotton in Tennessee and Arkansas respectively[12]. According to the International Survey of Herbicide-Resistant Weeds (http://www.weedscience.org), there are currently 430 unique cases of herbicide resistant weeds globally, with resistance to two or more herbicides found in 235 species (138 dicots and 97 monocots)[13]. It has also been recorded that weeds have evolved resistance to 22 of the 25 known herbicide sites of action and to 154 different herbicides. Herbicide resistant weeds have been reported in 82 crops in 65 countries[13]. The Poaceae family have 74 recorded cases of weed resistance with 21 cases reported in the Brassicaceae family to date. The yield of staple crops are affected by herbicide resistant weed species; wheat is affected the most with 65 herbicide resistance species, followed by corn, rice and soybeans which have 58, 50 and 46 herbicide resistant species respectively[13]. Globally there has been a rapid increase in the number of unique resistance cases between the years 1975-2014 and the highest recorded number of herbicide-resistant weeds are recorded in the US (145), followed by Australia (69), Canada (60), China (37) and France (35)[13].

The shift in farming methods in developing countries from mechanical removal of weeds to the use of herbicides means that herbicide resistance is a global problem. Countries such as China and Brazil now rely heavily on herbicides for weed control and are reporting a rapid increase in herbicide resistance cases with China currently ranked fourth in the world with 37 cases and Brazil ranked joint seventh with Germany on 31 cases. The full extent of herbicide resistance in most parts of Africa and Asia is not yet known, with no recorded cases in most countries on both continents.

1.2.1 Health and environmental concern of herbicide usage

In addition to herbicide resistance another major cause for concern in the use of herbicides are the health and environmental effects. The triazine herbicides have been linked to an increased risk of breast cancer[14], and phenoxy herbicides contaminated with dioxins have also been linked to mortality of workers exposed to the herbicide from diseases including neoplasms, soft-tissue sarcoma, and non-Hodgkin's lymphoma[15]. Some herbicides have been found to contaminate groundwater; a study carried out by Thurman and colleagues in the United States found that several herbicides including atrazine, alachlor and simazine exceeded the Environmental Protection Agency promulgated maximum contaminant levels for drinking water[16].

The health and environmental concerns of herbicide use combined with the growing cases of herbicide resistance on a global scale has led to a search for new chemically formulated herbicides with a specific site of action and minimal impact on the environment, animals and humans. The advent of research into sphingolipid synthesis has brought about a possible new site of action involving the non-mammalian enzyme, inositol phosphorylceramide synthases (IPCS). This enzyme catalyzes the transfer of phosphoinositol from phosphatidylinositol to the C-1 hydroxyl group of phytoceramide thereby generating the precursor inositol phosphorylceramide (Figure 1-3), for the

synthesis of the major sphingolipids in plants, glycosyl inositol phosphoceramides (GIPC). Exploitation of the divergence in sphingolipid biosynthesis in plants and animals can be used to create herbicides with minimal risks to mammals; fine-tuning of an IPCS inhibitor could result in the selective inhibition of monocots over dicots or vice versa, creating the pretext for the formation of herbicides that specifically kill certain types of weeds.
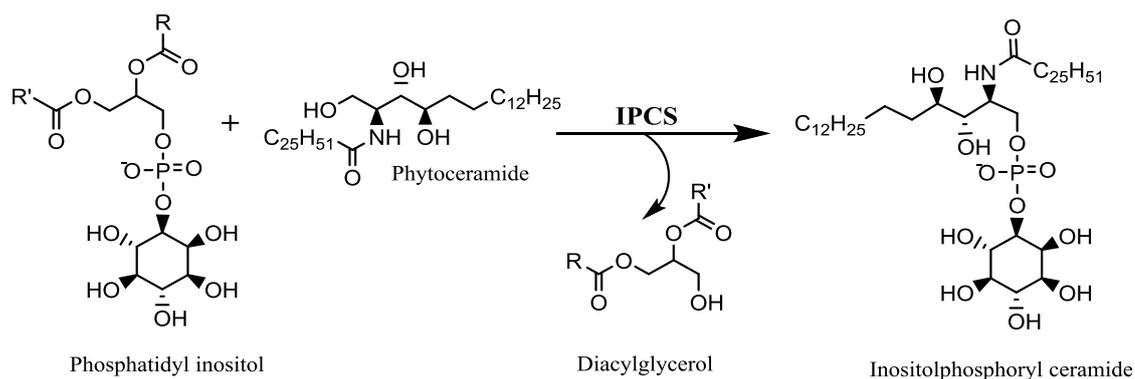
Figure 1-3: IPCS catalyses the transfer of phosphoinositol to phytoceramide thereby generating inositol phosphoryl ceramide and the by-product diacylglycerol.

1.3 Sphingolipids

Sphingolipids are a class of lipids that are essential components of all eukaryotic cell membranes and have been found to be crucial for survival across all kingdoms. In mammals they have been found to be involved in embryogenesis[17], differentiation[18], cell adhesion[19], and signal transduction[20, 21]. The model organism *Saccharomyces cerevisiae* has an abundance of sphingolipids comprising about 30% of the phospholipids found in the plasma membrane and these have been shown to be essential in growth regulation and cell integrity[22-24]. This class of lipids were first identified by Johann Tudichum during research into the effects of cholera on the brain[25]. The findings of his research published under the title 'a treatise on the chemical constitution of the brain' in 1884 identified sphingomyelin, cerebroside and sulfatide, which were isolated from fractional crystallization of ethanolic brain extracts[26].

This class of macromolecules consists of a sugar residue, a fatty acid and an organic base which Tudichum named `sphingosine' after the Greek sphinx, referring to the enigmatic structure of sphingolipids. Sphingosine was structurally characterized as 2*S*,3*R*,4*E*-2-aminooctadec-4-ene-1,3-diol in 1947 by Carter who also proposed the designation of lipids derived from sphingosines as sphingolipids. The basic unit of a sphingolipid is a sphingoid long-chain base (LCB), usually sphingosine, sphinganine (dihydrosphingosine) or 4-hydroxysphinganine[27] and a fatty acid component linked to carbon-2 of the LCB via an amide bond yielding ceramide, with different groups attached at the 1-hydroxyl position resulting in different types of complex sphingolipids (Figure 1-4). The presence of a phosphate group at the 1-hydroxy position of ceramide yields ceramide-1-phosphate, and an O-glycosidically linked glucose or galactose substituent results in the formation of the compounds: glucosylceramide and galactosylceramide (cerebroside)[28]. The higher glycosylated ceramide species, such as sulfatides[29] are derived from galactosylceramide by addition of a sulphate group to the 3-position of the sugar residue, whilst gangliosides contain a sialic acid (mostly N-acetyl neuraminic acid) moiety in the carbohydrate head group[30].
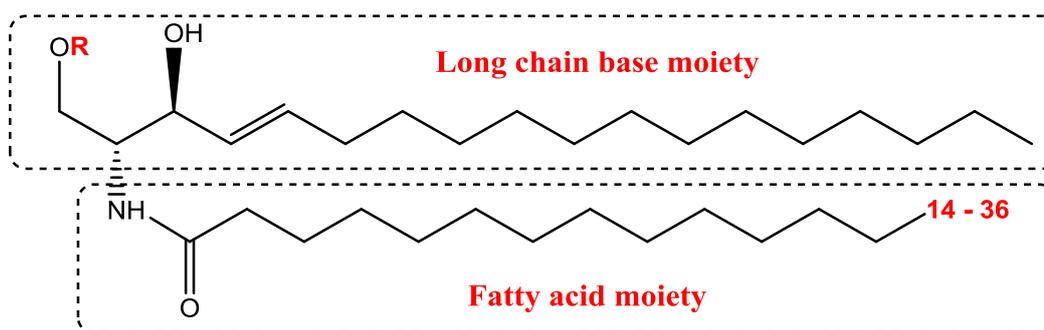
Figure 1-4: General structure of a sphingolipid which consists of an LCB moiety, a fatty acid moiety and different groups attached to the LCB component, where R can a variety of head groups such as glucose, phosphate, phosphocholine, phosphoinositol or the simplest head group H to generate ceramide.

Sphingolipid variation results from differences in the LCB moiety and the groups attached to the LCB component; this is seen in the variation found in hydroxylation patterns,

number and stereochemistry of the double bonds, chain length and the branched side chains.

1.3.1 LCB Variation

In mammals, the LCB moiety is mostly (*E*)-sphing-4-enine (sphingosine, d18:1[4]), whereas in yeast the predominant LCB is 4-hydroxysphinganine (phytosphingosine, t18:0) formed by desaturation or hydroxylation of sphinganine (d18:0) at C-4[30]. The predominance of 18 carbon sphingoid bases (d18:0, d18:1, and t18:0) in most mammalian sphingolipids is consistent with the preference of mammalian serine palmitoyltransferase (SPT) for saturated fatty acyl-CoAs composed of 16 carbons atoms in length[31], combined with the abundance of palmitoyl-CoA[32]. Sphingoid bases with chain lengths of 12 to 26 carbons have been reported in human skin[31] which belongs to a special class of ceramide containing ω-hydroxy fatty acid[30].

Plant sphingoid bases are composed of eight different C_{18}-sphingoid bases derived from sphinganine[33] (Figure 1.5). C*is*- or *trans*-desaturation at C-8 results in the unsaturated plant LCBs: (*E*/*Z*)-sphing-8-enine (d18:1[8]), (4*E*,8*E*/*Z*)-sphinga-4,8-dienine (d18:2[4,8]) and (8*E*/*Z*)-4-4-hydroxy-8-sphingenine (t18:1[8]); d18:1[4] is absent whilst the major LCBs in *Saccharomyces cerevisea,* d18:0 and t18:0 are present in minor proportions[34, 35]. Other LCBs are present in plants as minor components[36].

Plants are uniquely different from animals and yeast in that the major LCBs t18:1(8*E*/*Z*) and d18:2 (4*E*/(8*E*/*Z*)) are not found in these organisms. The proportion of sphingosine found in different species appears to segregate along taxonomic lines, with the *Solanacae* (tomato and tobacco) having large proportions and the *Fabacae* (pea and soybean) an intermediate amount, whilst *Brassicaceae* (*Arabidopsis*) have very low to non-existent levels of sphingosine[37]. The major difference between sphingoids found in

Sphinganine/dihydrosphingosine (d18:0)

(E)-Sping-4-enine/sphingosine (d18:1$^{\Delta 4E}$)

(E)-Sphing-8-enine (d18:1$^{\Delta 8E}$)

(Z)-Sphing-8-enine (d18:1$^{\Delta 8Z}$)

(4E,8E)-Sphinga-4,8-diene (d18:2$^{\Delta 4E, \Delta 8E}$)

(4E,8Z)-Sphinga-4,8-diene (d18:2$^{\Delta 4E, \Delta 8Z}$)

4-Hydroxysphinganine (t18:0)

(E)-4-Hydroxysphing-8-enine (t18:1$^{\Delta 8E}$)

(Z)-4-Hydroxysphing-8-enine (t18:1$^{\Delta 8Z}$)
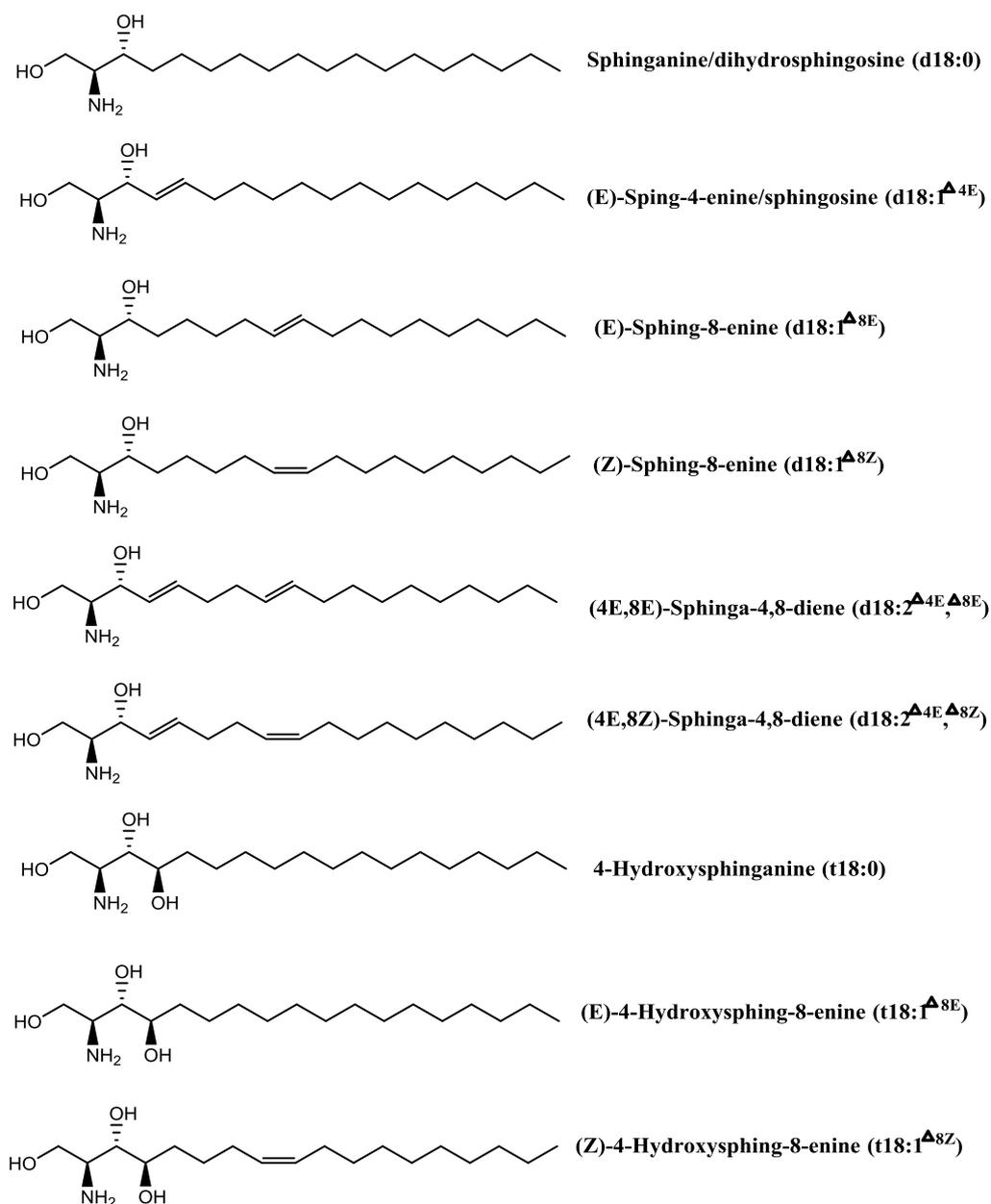
Figure 1-5: The structures and shorthand designation of the different sphingolipids found in plants, fungi and mammals whereby the LCB can be dihydroxy or trihydroxy; the number before the colon is the number of carbon atoms present in the compound and the number after the semi colon is the degree of unsaturation. Adapted from Pata *et. al*[33].

plants and fungi as compared to mammals is the lack of the C4-C5 double bond which shows a higher pattern of hydroxylation at C4, usually with a longer acyl chain[38]. Sphingolipid variation in other species is even more complex, with nematodes having both iso- branched and anteiso-branched (4*E*,6*E*-d13:2)[39] sphingoid bases (Figure 1-6A and 1-6B) . In *Drosophila* a sphingoid possessing a conjugated diene has been found[40]

(Figure 1-6C), whilst in aquatic organisms, unique sphingoids bearing a cyclopropane ring have been isolated from marine sponges (Figure1-6D)[41, 42].
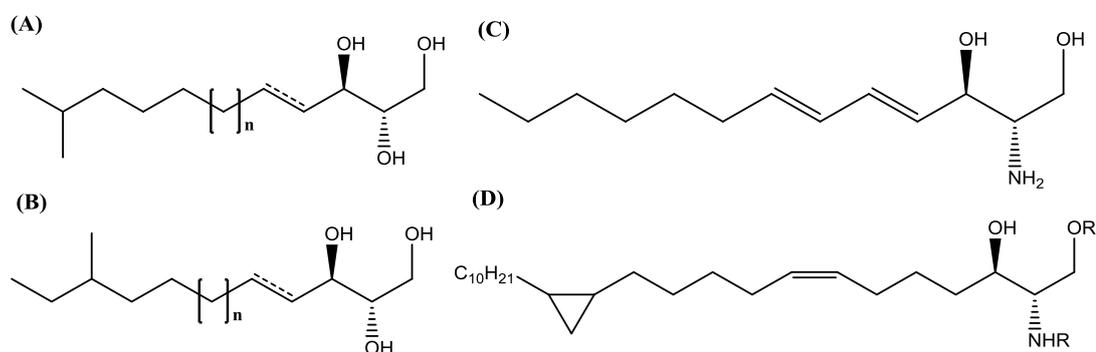
Figure 1-6: Structural variation in sphingoid bases in different species: (A, B) Nematode (C) Drosophila (D) Plakortis simplex. Figure adapted from Pruett *et. al*[42].

1.3.2 Sphingolipid diversity in plants

There are at least 500 (and perhaps thousands of) different molecular species of sphingolipids in eukaryotes[33], and in *Arabidopsis* alone it has been reported that there are at least 168 different sphingolipids[43, 44]. The structural variability found in plants arises from the diversity of the ceramide backbone and the high proportion of very long chain fatty acids (VLCFAs) found; whereby the fatty acid component can vary from C14-C26 and be saturated or monounsaturated. The fatty acid component can be an α hydroxylated VLCFA, which is predominantly found in maize root plasma membranes[45]. The most abundant α-hydroxylated FA composed of saturated C_{16}, C_{20}, C_{22} and C_{24} long chains, in contrast there is a sparse population of ω9-monounsaturated FAs composed of C_{22-26} carbon chains[46].

In plants there are four classes of sphingolipids: glycosyl inositol phosphoceramides (GIPCs), glycosylceramides, ceramides, and free long-chain bases (Figure 1-7)[47]. IPC is a precursor for the synthesis of GIPCs, but in fungi is identified as one of the predominant sphingolipids, which includes the two mannose-containing derivatives:

mannosyl inositolphosphorylceramide (MIPC) and mannosyl diinositolphosphorylceramide (M(IP)$_2$C)[48].
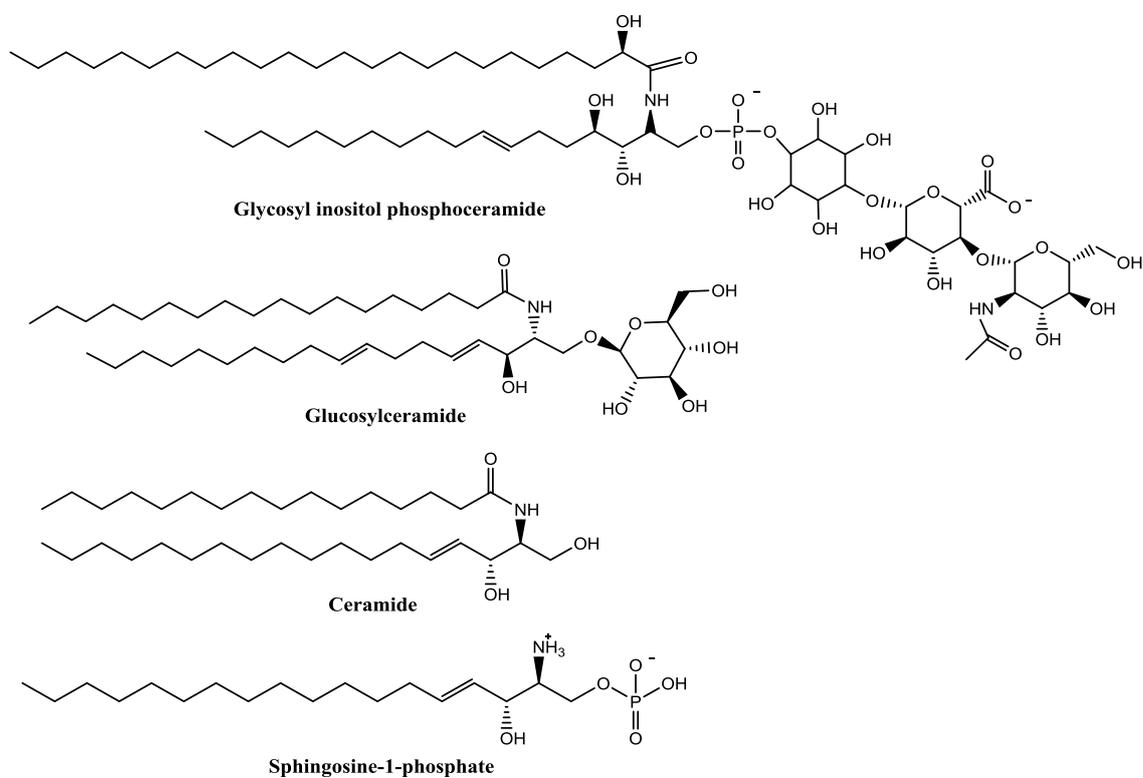
Figure 1-7: Structures of sphingolipids found in plants: Ceramide is the precursor to the more complex sphingolipids glucosylceramide (GlcCer) and glycosyl inositol phosphorylceramide (GIPC). Sphingosine-1-phosphate is a phosphorylated free LCB which has bioactive properties.

The deletion of inositol phosphorylceramide (IPC) synthase which synthesizes IPC has been found to be lethal in *Saccharomyces* and *Aspergillus* and therefore essential for survival[49.] To date sphingomyelin, which is the major sphingolipid in animal tissues, has not been detected in plants.

1.3.2.1 Glycosyl inositol phosphoceramides

Glycosyl inositol phosphoceramides (GIPCs) are acidic glycosphingolipids (GSLs) and are the major sphingolipids in plants, accounting for 64% of total sphingolipids[37]. IPC is the precursor for the synthesis of GIPCs, which consists of a polar head group linked to

C-1 of an *N*-acyl long-chain base that is attached to a fatty acid moiety[50] (Figure 1-8). The two GIPC core head groups, phytoglycolipid (PGL) and ceramide phosphate polysaccharide (CPPS), were identified by Carter et. al, and found to be attached to different sugars[51]. The oligosaccharide chains of PGL consist of glucosamine, hexuronic acid and inositol linked to mannose. The core head groups, phosphosphingolipid II and glycophosphoceramide consist of oligosaccharide chains attached to C-2 and/or to C-6 of *myo*-inositol and contain either *N*-acetylglucosamine or glucosamine and glucuronic acid with differing amounts of the additional sugars, arabinose, galactose, mannose and fuctose[52, 53] (Table 2).
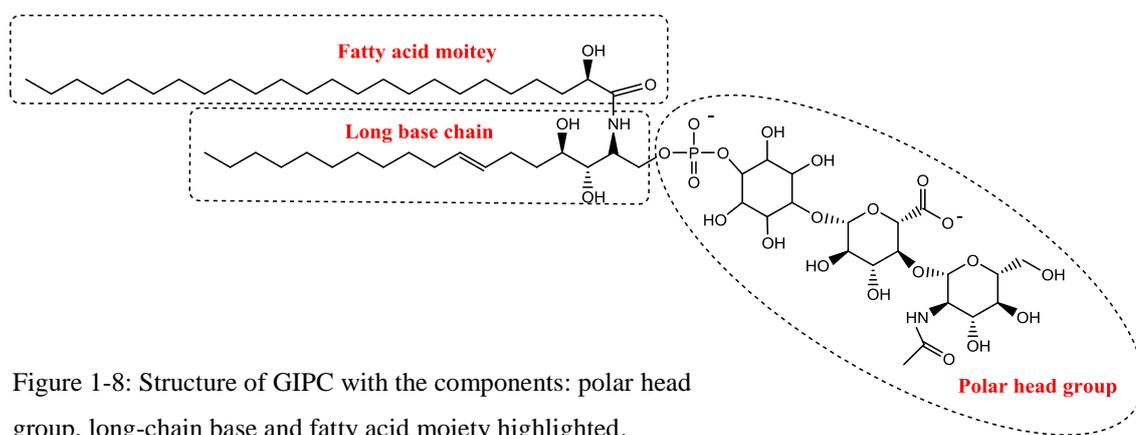
Figure 1-8: Structure of GIPC with the components: polar head group, long-chain base and fatty acid moiety highlighted.

Analysis of the LCB moiety of GIPCs located in leaf and endosperm tissue in the *Poaceae* family showed that t18:1Δ8(E/Z) and t18:0 LCBs are predominant[47, 54], with a high percentage of saturated and α hydroxylated very long fatty acid chains that are C24 or C25 long[55]; 2-hydroxylignoceric acid (h24:0) being the most common[49].

The enzyme, inositol phosphorylceramide glucuronosyltransferase (IPUT) from *Arabidopsis* when introduced into yeast has been shown to transfer glucuronic acid to IPC which is the first step in IPC glycosylation pathway to produce GIPCs, with the silencing of *IPUT* resulting in the accumulation of IPC (30% increase)[56].

This class of sphingolipids have been to interact with the oligosaccharide, rhamnogalacturonan II (RG-II) via boron, suggesting it has a role to play in the dimerization of RG-II which maintains the mechanical properties of the cell wall[57].

Table 2: Core head group structure of GIPCs in different plant species. Adapted from Pata *et. al* 2009[33]

Family	Species	Tissue	Core head group	Core head group structure	Additional sugars
Linaceae	*Linum usitatissimum*	Seed	PGL	Glucosamine-hexuronic acid-inositol-P-Man	Ara, Gal, Man, Fuc
			CPPS	Hexuronic acid-inositol	
Poaceae	*Zea mays*	Seed	PGL	Glucosamine-hexuronic acid-inositol-P-Man	Ara, Gal
			CPPS	Hexuronic acid-inositol	
	Triticum aestivum	Seed	(Inositol-P, GlcN, Ara,Gal, Man)	-	
Asteraceae	*Helianthus annus*	Seed	(Inositol-P, GlcN, Ara,Gal, Man)	-	
	Carthamus tinctorius	Seed	CPPS	Hexuronic acid-inositol	Ara, Gal, Fuc, Man
Solanaceae	*Nicotiana tabacum*	leaf	PSL-I	N-acetylglucosamine (α1-4)-glucuronic acid (α1-2)-myo-inositol-1-O-P	[Ara3Gal2] [Ara$_2$Gal$_2$] [Ara4Gal2]
			PSL-II	Glucosamine-glucuronic acid-inositol-P	[Ara$_3$Gal] [Ara2or3Gal2] [Ara2Gal2Man]
			GPC	N-acetylglucosamine(α1-4)-glucuronic acid(α1-2)-myo-inositol-1-O-P | Man	Ara, Gal

Table 2: The names of core head groups and their structures; PGL, phytoglycolipid; CPPS, ceramide phosphate polysaccharide; Ara, arabinose; Fuc, fucose; Gal, galactose; Man, mannose; CPPS, ceramide phosphate polysaccharide; GPC, glycophosphoceramide; PGL, phytoglycolipid; PSL, phosphosphingolipid. Adapted from Pata *et. al* (2009)[33].

In addition, GIPC has been found to be important for normal pollen development, with *iput1* mutants found to transmit the alleles through pollen at 1-2% compared to wild

type[56]. GIPCs are not present in mammals but have been found in protozoa, plants, fungi, and nematodes[58].

A variant of GIPC are the glycosylphosphatidylinositol (GPI)-anchors. GPI-anchors in plants are linked to C-terminus of a protein via an ethanolamine phosphate linkage and a conserved oligosaccharide core to an inositol phospholipid moiety[59]. Over 200 GPI-anchored proteins have been predicted to exist in the *A. thaliana* genome with functions including signaling, adhesion, response to stress, and cell-wall remodeling[60].

1.3.2.2 Glycosylceramides

Glycosylceramide is a ubiquitous sphingolipid found in animals, plants and fungi (but not *S. cerevisiae*). Glycosylceramide is often referred to as cerebroside due to its structural similarity with a compound found in the brain known as galactosylceramide[47]. GlcCers are a component of the plasma membrane and the tonoplast of plant cells, accounting for about 5-30 mol % of total lipid in the plasma membrane[61, 62] and constitute 34% of total sphingolipid content in plants[37]. Their abundance in the plant plasma and vacuolar membranes, combined with the ease of extraction and purification, has led to extensive structural characterization[63].

Composed of a β-glucose or β-mannose head group, GlcCers consist of α hydroxylated short or long fatty acid chains attached at the C1 position to a hexose sugar[64]. The exact nature of GlcCers structure varies for different plant species. GlcCers possessing VLCFAs are found to be predominant in the *Poaceae* family with a wide range of LCB variations; lower percentages of d18, $d18:1^{\Delta 4}E$, $d18:1^{\Delta 8}E$ *and* $d18:1^{\Delta 8}Z$ *LCBs* are observed in leaf and endosperm tissues, with much higher percentages of LCBs found to be $d18:2^{\Delta 4}E^{\Delta 8}E$, $d18:2^{\Delta 4}E^{\Delta 8}Z$, $t18:1^{\Delta 8}E$, $t18:1^{\Delta 8}E$ and *$t18:1\Delta 8Z^{64, 65}$*. In the *Brassicaceae*

family, saturated α hydroxylated VLCFAs content in *Arabidopsis* are found to around 5%, with >80% of GlcCers possessing α hydroxylated short fatty acids[66].

GlcCers identified in *Oryza sativa* endosperm and seed bran were found to be enriched with the *cis* and *trans* isomers of 8-sphingenine and 4,8-sphingadienine with C_{16}-C_{20} saturated hydroxyl fatty acids[67], contrasted with GlcCers identified in tobacco leaves which were enriched in trihydroxy long-chain bases (*cis* (*Z*) and *trans* (*E*) isomers of 4-hydroxy-8-sphingenine) with very long-chain (C_{20}- C_{26}) saturated[52]. The predominant fatty acid in both the leave and seed tissue in plants was found to be 2-hydroxypalmitic acid (h16:0)[34, 68]. Monoenoic hydroxy fatty acids, such as 2-hydroxynervonic acid (h24:1) are prevalent in *Arabidopsis* GlcCers[69] and cold hardy cereals. The variation of the fatty acid moiety in GlcCers has been implicated in chilling tolerances of plants; glycosylceramides containing *α*-OH monounsaturated VLCFAs have been detected mainly in chilling-resistant plants[66] and in chilling-sensitive plants, the predominant GlcCers are α-OH FA[70].

1.3.2.3 Ceramide

Ceramides are the third class of plant sphingolipids and are formed by the *N*-acylation of an LCB and a FA. These compounds are bioactive and are known to induce apoptosis in animals and programmed cell death in plants[71]. In comparison to GIPCs and glycosylceramides, plant ceramides are less documented in literature[47]; this can be attributed to their lower abundance in plant membranes making up 4-10% of plant glycosylceramides content and 1.7% of total sphingolipid content in *Arabidopsis*[37, 72].

Ceramide content in *Arabidopsis* is estimated to be 2-7 mol% and in rice leaf stems to be 6% [37,66]. The main FA moieties of plant ceramides are usually *α* hydroxylated VLCFAs especially in Poaceae, although high proportions of non-hydroxyl FAs have also been

found in rice and potato[67, 73]. The predominant LCB in plant ceramides are the trihydroxy-LCBs; the main LCB in *Oryza sativa* and *Arabidopsis thaliana* is t18:0[37].

1.3.2.4 LCB

The fourth class of plant sphingolipids is the free LCB. LCBs have been shown to be mediators in cellular responses; a study conducted by Dharmawardhane *et al.* showed that free LCBs such as 4-sphingenine (d18:1$^{\Delta 4}E$), sphinganine (d18:0) and 4-hydroxysphinganine (t18:0) modulated redox in oat mesophyll cells[74]. These sphingoid bases are known to be the most active inhibitors of protein kinase C in animal cells, inhibit electron transport in the transplasmalemma of the mesophyll cells in the dark, and stimulate redox transport in light[74]. Free LCBs such as 4-sphingenine (d18:1$^{\Delta 4}E$) regulate the vacuolar pyrophosphatase (V-PPase) proton pump in *Chenopodium rubrum* in order to maintain vacuolar and cellular acidity[75]. Work done by Brodersen *et al.* has shown that the accumulation of sphingosine in *accelerated-cell-death11 Arabidopsis* mutant (*acd11*), a sphingosine transfer protein, results in cell death characteristic of animal apoptosis[76] with an elevation in ceramide-1-phosphate and phytoceramide levels[77].

Interestingly, the relative proportions of E and Z isomers of 4-hydroxy-8-sphingenines was found to contribute to freezing tolerance, with higher levels of the Z isomers found in twelve different species of chilling resistant plants[70].

1.4 Sphingolipid biosynthetic pathway

Sphingolipids can be formed via two pathways: the *de novo* pathway, starting with the condensation of a serine with an acyl-CoA and the salvage pathway whereby ceramide and LCBs are released from more complex sphingolipids, followed by channelling of the metabolites formed into the synthetic pathway[27].

1.4.1 The *de novo* pathway

The core biosynthetic pathway of sphingolipid synthesis in plants and animals (Figure 1-9) is complex, involving five different enzymes and with synthesis occurring in the endoplasmic reticulum (ER) and the Golgi apparatus. The first four enzymes in the biosynthetic pathway are located in the ER, whilst enzyme involved in the formation of complex sphingolipids are located in the Golgi apparatus, with the active site of most facing the cytoplasm, but some facing the lumen implying significant transbilayer and interbilayer movement[27] involving vesicular and non-vesicular trafficking[78]. To date the known mode of transport between the ER and Golgi involves the cytoplasmic ceramide transport protein (CERT); this protein has a putative domain for catalysing lipid transfer, which enables it to specifically interact with ceramide in membranes where other complex sphingolipids are present[79,80].

Sphingolipid synthesis in plants and animals begins with the enzyme serine palmitoyl transferase (SPT), which was first identified in yeast and belonging to a family of pyridoxal 5'-phosphate-dependent *a*-oxoamine synthases. SPT converts serine and fatty acyl CoA into 3-ketosphinganine, CoA and CO_2[81]. Reduction of 3-ketosphinganine to dihydrosphingosine is accomplished by 3-ketosphinganine reductase (3KSR) using nicotinamide adenine dinucleotide phosphate (NADPH). After this reaction, there is a divergence in the sphingolipid biosynthetic pathway, whereby in mammals acylation of dihydrosphingosine by dihydroceramide synthase yields dihydroceramide[27], whereas in plants, dihydrosphingosine is hydroxylated to form phytosphingosine. Formation of varying acyl chain lengths results from the different substrate preference of dihydroceramide synthase isoforms.
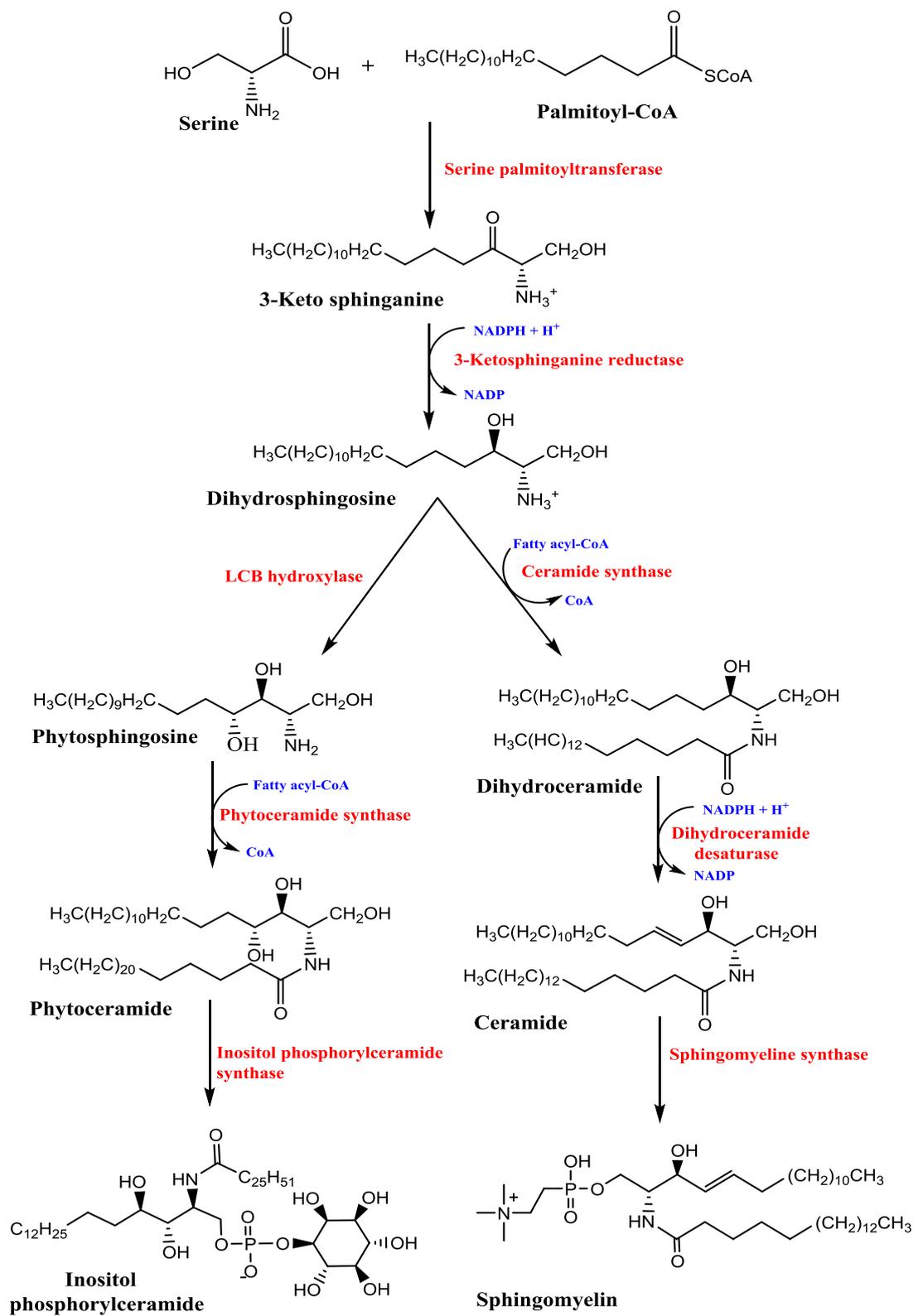
Figure 1-9: The biosynthetic pathway for the formation of IPC in plants and sphingomyelin in animals.

In mammals it has been found that there are six genes that encode ceramide synthase, known as the longevity-assurance gene (LAG1-6) [82,83] which are located in the

endoplasmic reticulum. The ceramide synthase isoforms show specific substrate preference in terms of the fatty acyl CoAs chain length thereby generating distinct ceramide acyl chains[84]. In plants, three ceramide synthase homologues termed LOH (LAG one homologue) have been identified, and like the mammalian orthologues show specific substrate preference, with LOH1 and 3 showing substrate specificity for trihydroxylated LCBs that have 20-26 carbon chains and LOH2 showing substrate specificity for LCB that are 16 carbon chain long and has a preference for dihydroxylated LCBs[85]. In animals insertion of a cis-4,5 double bond is achieved by dihydroceramide desaturase (DHCD) thereby generating ceramide, which is a precursor of many complex sphingolipids found in mammals, yeasts and plants[36]. In animals, ceramide is converted to sphingomyelin in the Golgi by sphingomyelin synthase (SMS), which involves transfer of the phosphorylcholine head group from phosphatidylcholine to ceramide yielding sphingomyelin and diacylglycerol[27].

Also, glucosylceramide synthase (GCS) converts ceramide to glycosylceramide, a precursor for the formation of complex GSLs; initial addition of glucose to ceramide occurs in the cytoplasmic side of the cis-Golgi and additional extension of the sugar chain occurs in the Golgi lumen as the GSLs are transported from the cis to the trans Golgi[86,87]. IPC, which has not been identified in animals, is the precursor for the generation of GPICs, which are the major sphingolipids in plants, and is generated by the transfer of phosphoinositol to phytoceramide by IPCS.

1.4.2 Salvage pathway

To date the salvage pathway of sphingolipid synthesis has only been studied in mammals and is yet to be investigated in plants and yeast. The salvage pathway of long-chain sphingoid bases leading to the regeneration of sphingolipids has been estimated to

contribute 50% to 90% of sphingolipid biosynthesis[88]. The degradation of sphingolipids occurs in the acidic subcellular compartments; the late endosomes and the lysosomes[89].

GSL degradation by exohydrolases results in the release of monosaccharide units from the end of the oligosaccharide chains thereby generating ceramide[89]. Experiments conducted using astrocytes showed that the major sphingolipid in mammals, sphingomyelin, is converted to ceramide and phosphocholine by acid sphingomyelinase (SMase) which can be stimulated in response to oxidative stress and bioactive compounds such as tumour necrosis factor and γ-interferon[20, 90]. Hydrolysis of ceramide by acid ceramidase forms sphingosine and free fatty acids in the lysosome[20]; once released from the lysosome, sphingosine can re-enter the de novo pathway for the regeneration of ceramide and sphingosine-1-phosphate (Figure 1.10), whilst long chain sphingoid bases are recycled in to the synthetic pathway by ceramide synthases.
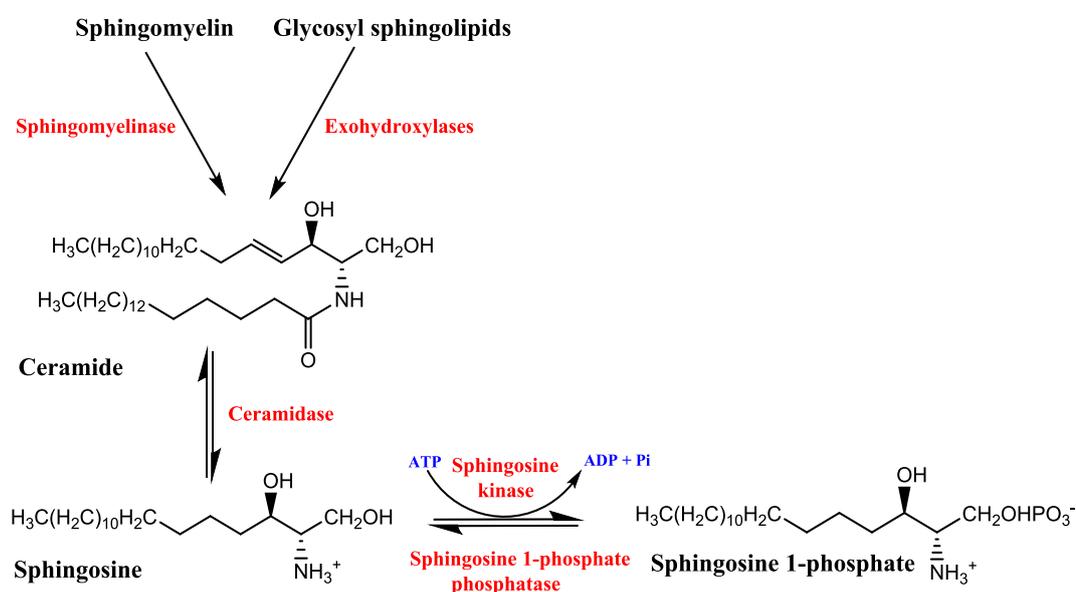
Figure 1-10: Schematic of the salvage pathway for the degradation of sphingolipids in animals.

1.5 Programmed cell death in plants

Programmed cell death (PCD) refers to the organised destruction of unwanted, infected and damaged cells[91]. This process is a genetically directed mechanism in response to

pathogen or abiotic stress. In addition it has a role in the maintenance of homeostasis in normal development such as senescence, vascular system formation and specification of unisexual floral organs[91]. The process known as apoptosis in animals, shares morphological characteristics with plant PCD such as condensation of nuclei, fragmentation of nuclear DNA, and cell shrinkage[92].

In plants, the biosynthesis of sphingolipids has been linked to PCD. As mentioned earlier, ceramide induces apoptosis/programmed cell death in animals/plants[72]. The first documented evidence of ceramide involvement in PCD was by Liang *et al.* which showed that in *Arabidopsis* the ceramide kinase mutant known as accelerated cell death 5 (acd5) accumulates lipid substrates similar to those utilized by wild type ceramide kinase, and exhibited disease like symptoms and apoptotic-like cell death during pathogen attack[93]. The mutant protein ceramidase activity was demonstrated in an assay which showed that the proteins had high specificity for synthetic C6 and C8 ceramides. The accumulation of ceramide in acd5 mutant cells resulted in apoptotic like features which had regions of condensed chromatin in the nuclei, possessing large numbers of DNA strand breaks generated by endonucleolytic cleavage[93]. The role of ceramide as an inducer of apoptosis was validated by treatment of wild type protoplasts with C2 ceramide which induced nuclear fragmentation characteristic of apoptotic cells[93]. Interestingly, phosphorylated derivatives of ceramide were able to partially block plant PCD indicating the role of phosphorylated ceramides in modulating cell death in plants[93].

In tobacco (*Nicotiana tabacum*) , cell death was found to be induced by palmitoleic acid (16:1), ceramide, and potassium cyanide; exhibiting features associated with PCD, including cell volume decrease, loss of membrane integrity, DNA damage, nuclear and plastid disorganization, and chromatin condensation[94]. Cell volume decrease was found to be caused by loss of intracellular K^+ channels; inhibition of K^+ channel with blockers

such as Ba^{2+} and quinine have been found to prevent cell shrinkage and the subsequent biochemical events leading to PCD[94].

Research carried out by Wang *et al.* identified the *Arabidopsis* functional homolog of yeast and protozoan IPCS which enhanced RPW8 hypersensitive-like cell death[95]. RPW8 is a resistance gene which confers broad-spectrum resistance to powdery mildew in *Arabidopsis*[96]. *RPW8* triggers the hypersensitive response (HR) to restrict powdery mildew infection via the salicylic acid–dependent signaling pathway; loss of function of *At*IPCS results in salicylic acid accumulation, *RPW8*-dependent spontaneous HR-like cell death (SHL) in leaf tissues, reduction in plant stature[95]. This study showed that T-DNA insertional mutants of *At*IPCS in plants unchallenged with powdery mildew, had an observable phenotype of chlorotic lesions and sporadic cell death, thereby revealing the role of IPCS as a negative regulator of PCD in plants[95].

1.6 IPC Synthase

IPCS belongs to a superfamily called the 'Sphingolipid Synthases' which includes animal sphingomyelin synthase (SMS)[97], protozoan IPCS[98], fungal IPCS[99] and plant IPCS[100]. Mutants *S. cerevisiae* strains defective in IPCS activity were found to accumulate ceramide which was accompanied by cell death, showing that IPCS is necessary for survival[101]; expression of a gene encoding AUR1 protein in these mutants was found to complement for the activity of IPCS[101]. Mutations in AUR1 have also been found to confer resistance to the antifungal IPCS inhibitor Aureobasidin A (AbA)[101, 102], showing that AUR1 is fungal IPCS.

The protozoan IPCS enzyme was identified in kinetoplastid *Leishmania major* by Denny *et al.* in 2006 after screening the genome database for the two conserved active site motifs (CGDX$_3$SG**H**T and **H**YTX**D**VX$_3$YX$_6$FX$_2$YH) shared by lipid phosphate phosphatase,

SMS and AUR1[98]. The conserved motifs contain the **HHD** catalytic triad (underlined and bold) involved in the proposed double displacement mechanism of IPCS. This protein was shown to rescue an *AUR1* yeast mutant[98] and subsequent analyses of the genomes of related kinetoplastids, *Trypanosoma brucei* and *Trypanosoma cruzi* genome, revealed four IPCS orthologues in *T. brucei* and two in *T. cruzi*[103].

Plant IPCS was first identified in *Arabidopsis thaliana* by Wang *et al.*[95] and the three encoded isoforms *At*IPCS1-3 were shown to be AbA resistant AUR1p orthologues differentially expressed in various tissues[104]. Across species, *At*IPCS1-3 and *Lmj*IPCS showed the highest sequence identity (~30 to 45%), especially in the conserved domains, D3 and D4 [104]. A Basic Local Alignment Search Tool (BLAST) search of the *Oryza sativa* (rice) genome using the *At*IPCS1 protein sequence revealed three orthologues designated *Os*IPCS1-3.

1.6.1 IPCS topology

The first proposed model for IPCS topology was by Denny *et al.* based on based on the single sequence analysis of *Lmj*IPCS[98]. This model proposed that IPCS consists of seven transmembrane helices with the catalytic site facing the lumen of the Golgi apparatus and lacked the conserved domain D2 present in mammals; in addition, there is an extra helix between TM3 and TM4 consisting of 11 residues with the inversion of TM1-3 which preserves the orientation of the active. Another model has been proposed by Bangs *et al.*[105] based on the alignment of single sequences from three kinetoplastids species, suggests the presence of the D2 domain which is conserved in animal SM synthase[97], and discounts the existence of the extra helix (Figure 1-11). There has been no experimental evidence for either model, although the latter model is more in keeping with data from mammalian systems[97]
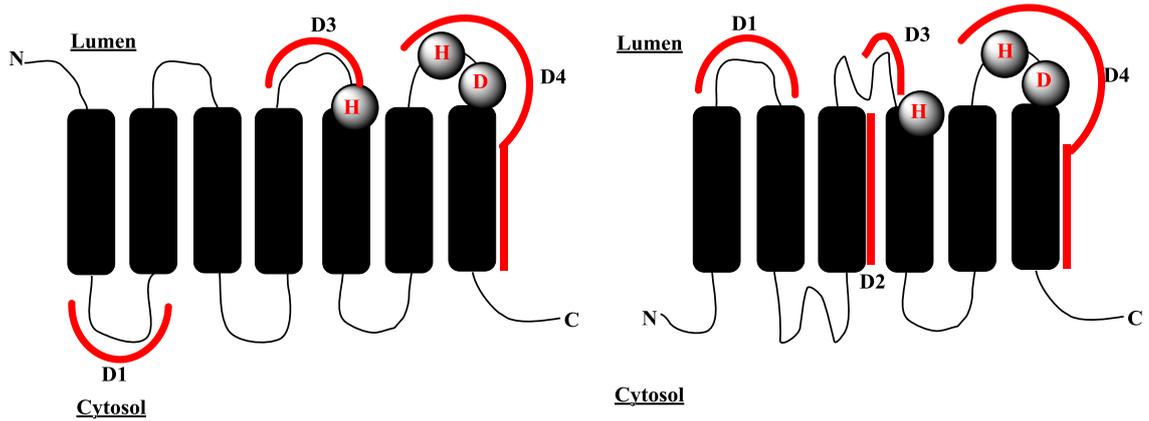
Figure 1-11: (A)Denny *et al.* model of *L. major* consisting of seven transmembrane helices and lacking D2 domain (B) Bang *et al.* model consisting of six helices and D2 domain.

1.6.2 Catalytic mechanism

The double displacement model is based on lipid phosphate phosphatase (LPP) hydrolysis of lipid phosphates. The domains designated D3 and D4 are similar to the C2 and C3 domains of LPPs; both enzymes have a catalytic triad consisting of two histidine residues and an aspartate residue[97]. Hence it has been proposed that IPCS catalysis follows a similar mechanism to the LLPs which involve the activation of one of the histidine by aspartate via a charge-relay system resulting in the deprotonation of the histidine, the formation of an inositol phosphoryl histidine intermediate and the release of diacyl glycerol (DAG)[99].

Heiddler and Radding proposed an IPCS mechanism based on the catalytic model of lipid phosphatases, whereby ceramide C_1-hydroxyl nucleophilicity is increased through hydrogen bonding with the histidine in D3, and the inositol phosphoryl histidine intermediate is protected from hydrolysis by the increased hydrophobicity of the catalytic site[99]. This sets up the conditions required for the nucleophilic attack on the inositol phosphoryl histidine intermediate by the C_1-hydroxyl of ceramide leading to the transfer of inositol phosphate to the C_1 of ceramide. Heiddler and Radding also postulate that the

role of domains D1 and D2 might be able to bind and stabilize ceramide, for a favourable positioning of the C_1-hydroxyl which facilitates the transfer of inositol phosphate at the surface of the membrane [99].

The catalytic phosphoinositol transfer has been proposed based on a catalytic mechanism of LPP activity by Sigal *et al.*, which involves the HHD catalytic triad, arginine, lysine, serine and glycine residues that stabilize the transition state[106]. Across different organisms, the LPP family of phosphatase have been found to consist of three conserved motifs in which one motif has the nucleophilic histidine in close proximity to an arginine residue[107]. This arginine residue has been proposed to be Arg262 in *Lmj*IPCS and found to be a conserved residue in all the identified orthologues of IPCs in *L. major*, *Trypanosoma brucei*, and *Trypanosoma cruzi* (Tritryp) genomes[108]. The amino acid residue Arg262 is hypothesised to be involved in the stabilisation of the transition state, with the protonation of the arginine group speculated to be affected by the protonation of the amino group of inhibitory sphingosines, resulting in inefficient stabilisation of the transition state and inhibition of IPCS activity[108].
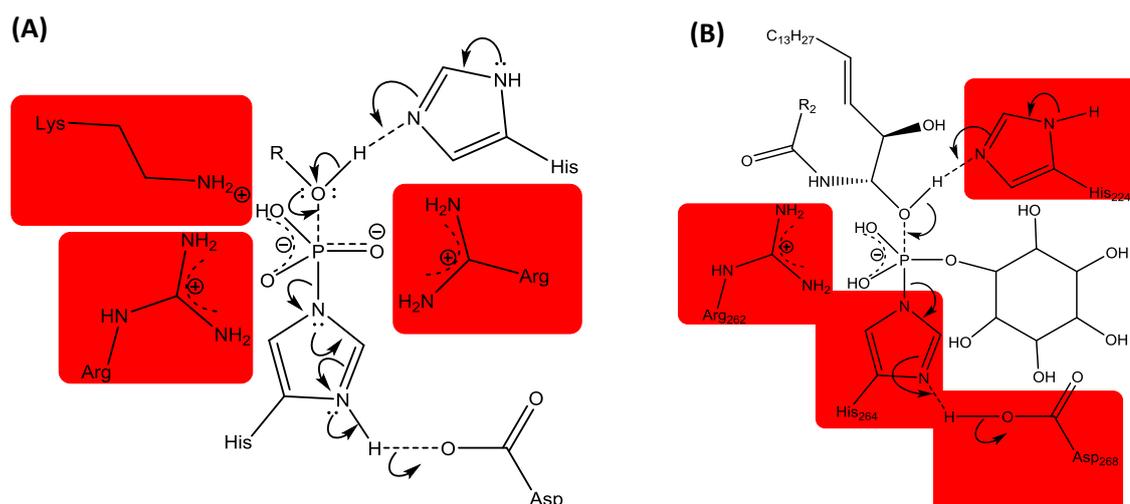
Figure 1-12: (A) Proposed mechanism of action of phosphoryl transferases (B) Proposed mechanism of action of *Lmj*IPCS, adapted from Sigal *et al*.

1.7 Targeting IPCS

As mentioned earlier, sphingolipids play an essential role in maintaining the integrity of the plasma membrane and are also involved in a plethora of biological pathways including signal transduction, cellular differentiation, and embryogenesis, as a result they are important for survival in organisms. The divergence in the sphingolipid biosynthetic pathway in animals, plants, fungi and protozoa post ceramide can be exploited in the development of herbicides, antifungal, and antiprotozoal drugs. (Figure 1-14).
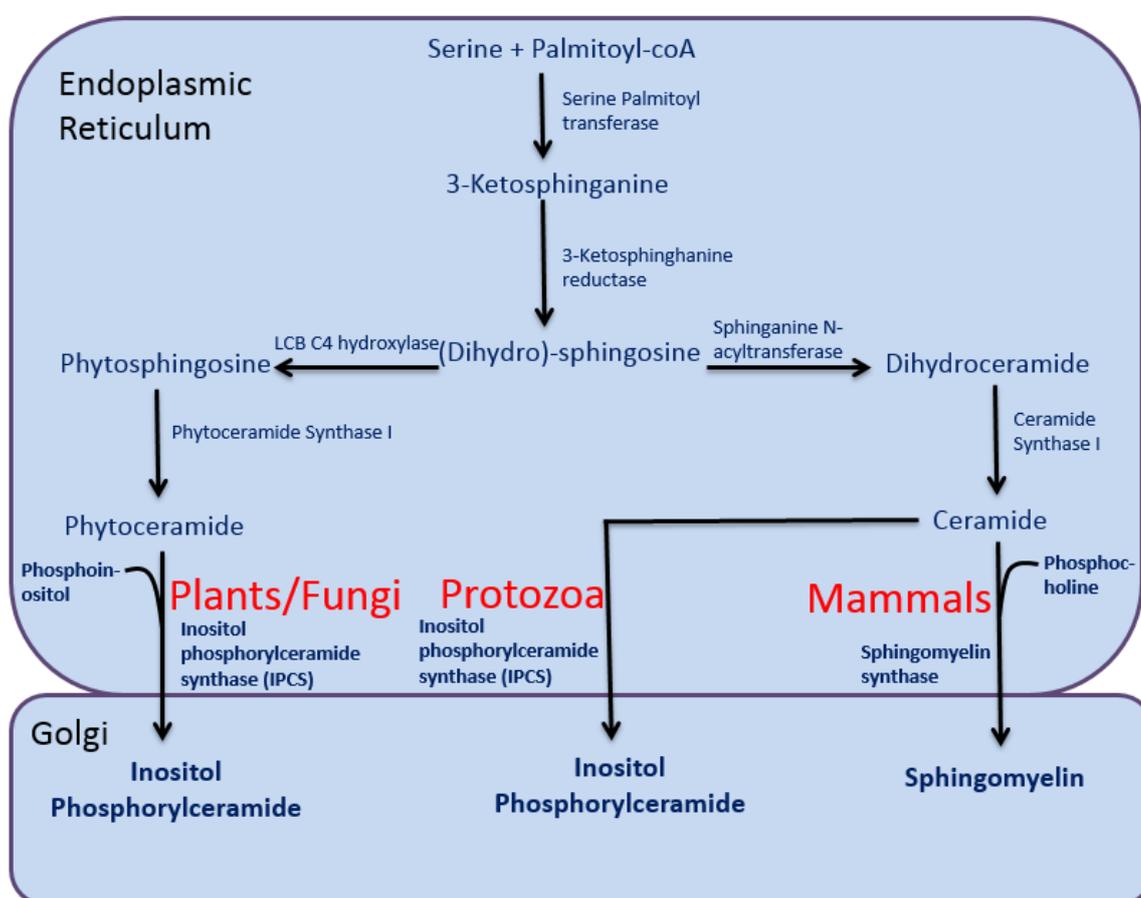
Figure 1-14: Schematic of the sphingolipid biosynthetic pathway: Fungi and higher plants IPCS converts phytoceramide into inositol phosphoryl ceramide, whilst kinetoplastids and mammals SMS converts ceramide to sphingomyelin (B) Inhibition of IPCS results in a decrease in mitogenic diacylglycerol, an accumulation of pro-apoptotic phytoceramide and a decrease in IPC, the precursor of the major plant sphingolipid GIPC.
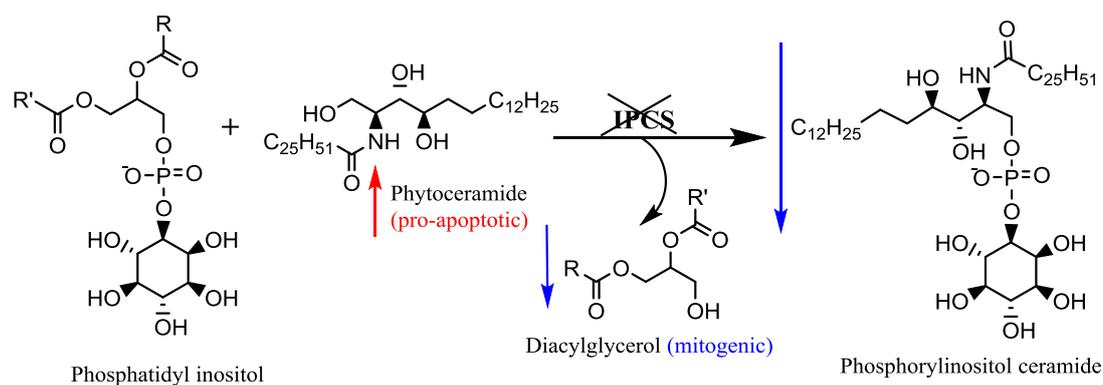
Figure 1-15: Inhibition of IPCS results in a decrease in mitogenic diacylglycerol, an accumulation of pro-apoptotic phytoceramide and a decrease in IPC, the precursor of the major plant sphingolipid GIPC.

The absence of sphingomyelin synthase in plants and the difference in enzyme substrate, i.e. phytoceramide and phosphatidylinositol for plant IPCS and ceramide and phosphatidylcholine for SMS, create the opportunity for the development of herbicides that specifically target plant enzyme with minimal toxicity to animals. Work done by Bi et al. showed that the loss of ceramide kinase results in ceramide accumulation and the formation of necrotic lesions associated with PCD, indicating that the accumulation of ceramide could trigger PCD[109]. Similarly, inhibition of IPCS should result in the accumulation of the pro-apoptotic phytoceramide, and a decrease in the turnover of the mitogenic compound diacylglycerol[110] (Figure 1-15) thereby initiating PCD.

1.7.1 Anti-protozoal and anti-fungal applications

The most common human diseases caused by kinetoplastids protozoa include human African trypanosomiasis (HAT), caused by *Trypanosoma brucei*; Chagas disease caused by infection with *Trypanosoma cruzi*, and leishmaniasis caused by *Leishmania* species[111]. Half a billion people are at risk of contracting these diseases, and it is estimated that more than 20 million people are infected with the pathogens that cause them, with more than 100,000 deaths recorded per year[111]. Current treatment for these diseases involves the use of drugs which are toxic with variable efficacy, difficult administration

and lengthy periods of treatment[112]. Research is currently being carried out to develop drugs against these diseases by the design of IPCS inhibitors.

The focus on IPCS inhibition has also been directed towards the development of anti-fungal drugs targeting phytopathogenic fungi which cause diseases in plants, animals and humans[113].

Saprophytic fungi from the *Aspergillus* genus are found worldwide in soil, food products, and decomposing matter[114]. However, the two species *Aspergillus flavus* and *Aspergillus parasiticus* are found to produce potent toxins (aflatoxins) on certain crops including oilseeds, corn, cottonseed and peanuts, which have been linked to death in humans and animals; in humans, inhaled aflatoxin B1 can cause inflammation and irreversible pulmonary interstitial fibrosis[115].

Four species, *A. fumigatus, A. flavus, A. terreus* and *A. niger*, are found to be the causative agents of fatal opportunistic infections such as invasive pulmonary aspergillosis[115]. Agricultural workers in developing countries whose eyes are injured with subsequent contamination by organic matter develop corneal infections by *Aspergillus* species[116]. Other phytopathgenic fungi species are part of the genus: *Fusariu, Alternaria, Curvularia*, *Cladosporium* and *Alternaria*; these produce toxins that affect humans and plant and create economic losses in agriculture and medicine.

There are a limited number of antifungals currently available for the treatment of life-threatening fungal infections[117]. These antifungal agents show some limitations, such as the significant nephrotoxicity of amphotericin B[118] and emerging resistance to the azoles[119]. The development of new antifungal agents is needed in medicine and agriculture, so the inhibition of essential fungal IPCS, which differs from the mammalian SMS, is a promising approach.

1.7.2 Herbicide application

A close look at herbicide weed resistance worldwide revealed that *Lolium rigidum* remains the world's most herbicide-resistant weed, with resistance to 11 sites of action and 95% of populations resistant to at least two herbicide modes of action (especially ACCase and ALS inhibition)[120]. *L. rigidum* affects 12 countries[11], especially southern Australia, and is found in crop fields growing wheat, barley, and oilseed rape. Other weeds that have a huge effect on agronomically important crops worldwide are the drought tolerant weed *Amaranthus palmeri* which affects cotton yields in north America, the weed *Avena fatua* which competes with wheat[121], and *Echinochloa crus-galli* a noxious weed affecting a wide variety of crops including wheat, rice, and potatoes[122].

Selectivity of herbicides for weeds over crops have been achieved in most cases because of differences in herbicide uptake, translocation and degradation. These differences have been found in some cases to be a result of structural differences between monocot and dicot plants. The classification of plants based on the monocot/dicot system is based on anatomical and developmental differences encompassing germination, root systems, leaf, stem, and vascular bundle architecture. The names are derived from the number of embryonic leaves found in the seed which is termed the cotyledon; monocots have one and dicots have two cotyledons. Difference in the germination process of seeds between the two classes refers to the location of the cotyledon below or above ground during plant development; monocots have the cotyledon below the ground whilst dicots have the cotyledon brought above the ground due to the elongation of hypocotyl[123]. A detailed look at the structural differences show that monocots have a fibrous root system which is shallow in depth compared to the tap root system found in dicots which possess lateral roots that branch out into tertiary roots, providing depth in growth[124]. In addition, the vascular architecture of the stem in dicots is arranged in the shape of a ring and the veins

of the leaves are reticulate[125], whilst in monocots the vascular bundles of the stem are highly complex[126] and the leaves have veins that are parallel in arrangement.

Discovering new herbicide modes of action is essential in the management of herbicide resistance worldwide. The study of IPCS as a new herbicide is promising because of the divergence in the latter stages of the sphingolipid biosynthesis in animals and plants, combined with the lack of IPCS in animals which would allow for the design of inhibitors/herbicides that should have no detrimental effect on animals and humans. The classification of plants based on the dicotyledon and monocotyledon system can be exploited to create IPCS inhibitors that would differentially inhibit monocotyledon weeds over dicotyledon crops or dicotyledon weeds over monocotyledon plants. For example it is well known that auxinic herbicides kill dicot weeds without killing acres of monocot crops such as corn, wheat, and barley[127].

1.7.3 Inhibitors of IPCS

To date there are five known potent inhibitors of fungal IPCS (Figure 1-16): AbA[102], khafrefungin[128], rustimicin[129], pleofungin A[130] and haplofungin A[131] , with IC_{50} values of 0.2 nM, 7nM, 20 nM, 40 µM and 1.5 µM against *Saccharomyces cerevisiae*.

The *T. brucei* IPCS which synthesises sphingomyelin and inositol phosphoryl ceramide is inhibited by AbA with an IC_{50} of 0.42 nM[132]. In contrast the *L. major* IPCS enzyme is relatively insensitive to AbA. *T. cruzi* IPC has been shown to be insensitive to rustmicin and Aba[133], although IPC formation was indicated to be inhibited by AbA suggesting the compound is active against a different enzyme in the sphingolipid biosynthetic pathway[133].

Differential selectivity between dicotyledon species has been observed between *Arabidopsis* and *Phaseolus.* All *Arabidopsis* IPCS are AbA insensitive[103], whereas in

Phaseolus vulgaris, AbA and rustmicin are found to be potent inhibitors of IPCS exhibiting IC$_{50}$ values of 0.4–0.8 and 16–20 nM respectively[100]. The screening of a library of compounds against *L. major* IPCS revealed a class of inhibitors which possess a structural motif consisting of an electronegatively-substituted biaryl unit separated from a structural amine by an alkyl chain[134], such as clemastine (Figure 1-17).
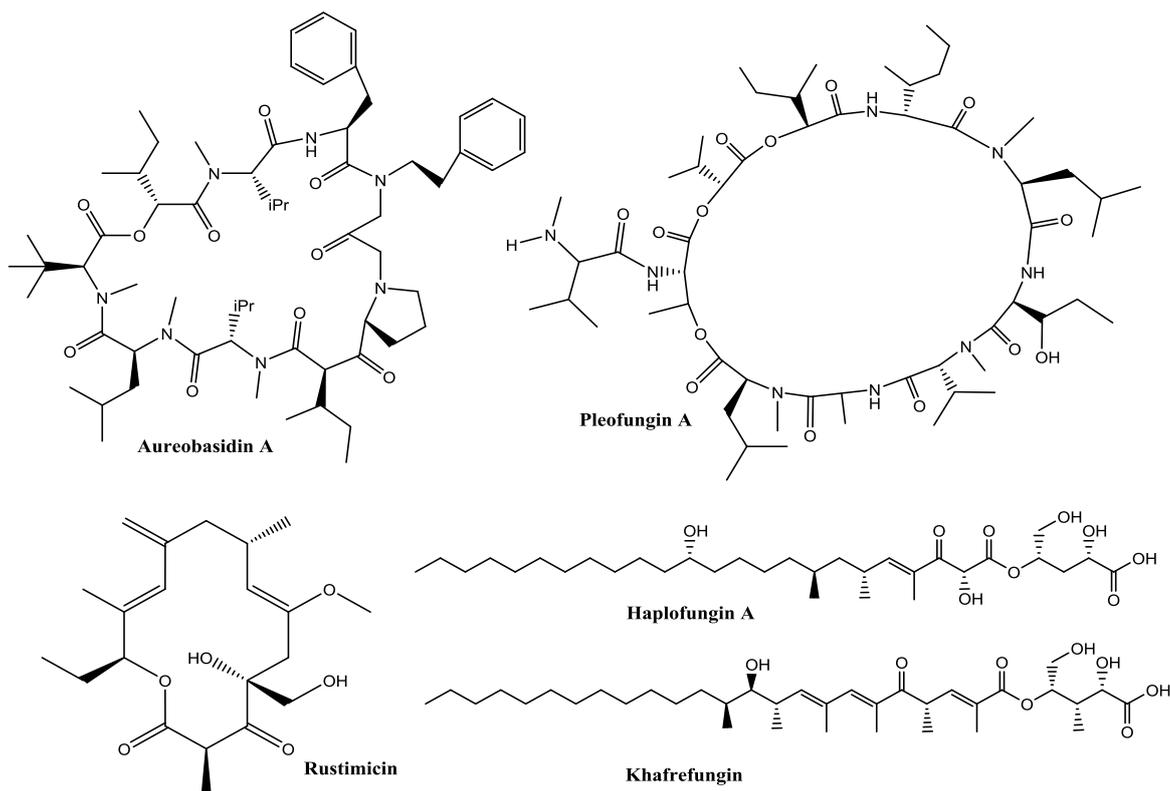
Figure 1-16: The structures of potent inhibitors of IPCS.
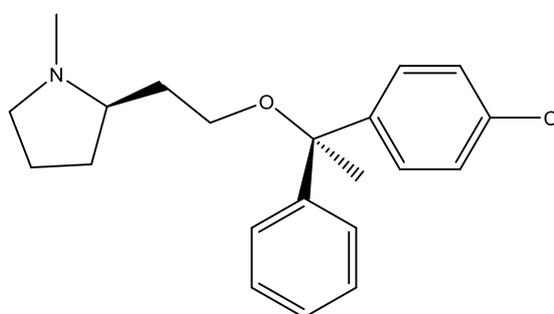
Figure 1-17: Structure of clemastine.

1.8 Project aims

The aim of this project was to identify and characterize IPCS orthologues from the agronomically important cereal crop, rice (*Oryza sativa*), and competing weeds, field poppy (*Papaver rhoeas*), a major persistent weed of cereal crops. However, due to lack of access to a weed genome to search for IPCS in *Papaver rhoeas* genome, *Arabidopsis thaliana* has been utilised as a model dicot weed in place of *Papaver rhoeas*. Mutant *S. cerevisiae* strains expressing functional plant IPCSs were formatted into an assay for the high throughput screening for plant IPCS inhibitors, that show selectivity for the monocot crop over the dicot weed (and vice versa). In tandem, transgenic lines of *Arabidopsis* with varying expression of IPCS were created and RNASeq carried out to identify genes/biological pathways that were altered in response to changes in IPCS expression; this gave an insight into the global function of IPCS in plants.

1.9 References

1.   UN, FAO statistical year book 2013: World food and agriculture. http://reliefweb.int/sites/reliefweb.int/files/resources/FAO_2013_stats_yrbook.pdf, 2013.

2.   Maxwell, S., World development report 2010: Development and climate change. Taylor & Francis Group: 2010.

3.   Hallegatte, S.; Fay, M.; Bangalore, M.; Kane, T.; Bonzanigo, L., *Shock waves: managing the impacts of climate change on poverty*. World Bank Publications: 2015.

4.   Arnell, N. W.; Cannell, M. G.; Hulme, M.; Kovats, R. S.; Mitchell, J. F.; Nicholls, R. J.; Parry, M. L.; Livermore, M. T.; White, A., The consequences of CO2 stabilisation for the impacts of climate change. *Climatic Change* **2002,** *53* (4), 413-446.

5.   Facts and figures 2013-2014 Bayer. *http://www.cropscience.bayer.com/en/Company/Facts-and-Figures.aspx*.

6.   De, A.; Bose, R.; Kumar, A.; Mozumdar, S., *Targeted delivery of pesticides using biodegradable polymeric nanoparticles*. Springer: 2014.

7.   Goldburg, R.; Rissler, J.; Shand, H.; Hassebrook, C., Biotechnology's bitter harvest: herbicide-tolerant crops and the threat to sustainable agriculture. *Biotechnology's bitter harvest: herbicide-tolerant crops and the threat to sustainable agriculture.* **1990**.

8.   Flint, M. L.; Van den Bosch, R., *Introduction to integrated pest management*. Springer Science & Business Media: 2012.

9.   Zimdahl, R. L., *A history of weed science in the United States*. Elsevier: 2010.

10.   Quastel, J., 2, 4-dichlorophenoxyacetic acid (2, 4-D) as a selective herbicide. ACS Publications: 1950.

11.   Heap, I., Global perspective of herbicide- resistant weeds. *Pest management science* **2014,** *70* (9), 1306-1315.

12.   AN, Y., What Happens When Weed Killers Stop Killing?

13.   Heap, I., The International Survey of Herbicide Resistant Weeds. *www.weedscience.org* **2017**.

14.   Kettles, M.; Browning, S. R.; Prince, T. S.; Horstman, S. W., Triazine herbicide exposure and breast cancer incidence: an ecologic study of Kentucky counties. *Environmental health perspectives* **1997,** *105* (11), 1222.

15.   Kogevinas, M.; Becher, H.; Benn, T.; Bertazzi, P. A.; Boffetta, P.; Bueno-de-Mesqurta, H. B.; Coggon, D.; Colin, D.; Flesch-Janys, D.; Fingerhut, M., Cancer mortality in workers exposed to phenoxy herbicides, chlorophenols, and dioxins an expanded and updated international cohort study. *American journal of epidemiology* **1997,** *145* (12), 1061-1075.

16.   Thurman, E. M.; Goolsby, D.; Meyer, M.; Kolpin, D., Herbicides in surface waters of the midwestern United States: The effect of spring flush. *Environmental Science and Technology* **1991,** *25* (10), 1794-1796.

17.   Bieberich, E., Ceramide in stem cell differentiation and embryo development: novel functions of a topological cell-signaling lipid and the concept of ceramide compartments. *Journal of lipids* **2011,** *2011*.

18.   Schengrund, C.-L., Gangliosides: glycosphingolipids essential for normal neural development and function. *Trends in biochemical sciences* **2015,** *40* (7), 397-406.

19.   Eich, C.; Manzo, C.; De Keijzer, S.; Bakker, G.-J.; Reinieren-Beeren, I.; García-Parajo, M. F.; Cambi, A., Changes in membrane sphingolipid composition modulate dynamics and adhesion of integrin nanoclusters. *Scientific reports* **2016,** *6*, 20693.

20.   Riboni, L.; Viani, P.; Bassi, R.; Prinetti, A.; Tettamanti, G., The role of sphingolipids in the process of signal transduction. *Progress in lipid research* **1997,** *36* (2-3), 153-195.

21.   Merrill Jr, A. H.; Sullards, M. C.; Wang, E.; Voss, K. A.; Riley, R. T., Sphingolipid metabolism: roles in signal transduction and disruption by fumonisins. *Environmental health perspectives* **2001,** *109* (Suppl 2), 283.

22.   Patton, J. L.; Lester, R. L., The phosphoinositol sphingolipids of Saccharomyces cerevisiae are highly localized in the plasma membrane. *Journal of Bacteriology* **1991,** *173* (10), 3101-3108.

23.   Skrzypek, M. S.; Nagiec, M. M.; Lester, R. L.; Dickson, R. C., Analysis of phosphorylated sphingolipid long-chain bases reveals potential roles in heat stress and growth control in Saccharomyces. *Journal of bacteriology* **1999,** *181* (4), 1134-1140.

24.   Schmelzle, T.; Helliwell, S. B.; Hall, M. N., Yeast protein kinases and the RHO1 exchange factor TUS1 are novel components of the cell integrity pathway in yeast. *Molecular and Cellular Biology* **2002,** *22* (5), 1329-1339.

25.   Thudichum, J. L. W., *A Treatise on the Chemical Constitution of the Brain.* Archon Books: 1962.

26.   Rao, R. P.; Acharya, J. K., Sphingolipids and membrane biology as determined from genetic models. *Prostaglandins & other lipid mediators* **2008,** *85* (1), 1-16.

27.   Futerman, A. H.; Riezman, H., The ins and outs of sphingolipid synthesis. *Trends in cell biology* **2005,** *15* (6), 312-318.

28.   Huwiler, A.; Kolter, T.; Pfeilschifter, J.; Sandhoff, K., Physiology and pathophysiology of sphingolipid metabolism and signaling. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2000,** *1485* (2), 63-99.

29.   Lochnit, G., Structural analysis and immunohistochemical localization of two acidic glycosphingolipids from the porcine, parasitic nematode, Ascaris suum. *Glycobiology* **1998,** *8* (9), 891-899.

30.   Berkey, R.; Bendigeri, D.; Xiao, S., Sphingolipids and plant defense/disease: the "death" connection and beyond. *Frontiers in plant science* **2012,** *3*, 68.

31.   Chalfant, C.; Del Poeta, M., *Sphingolipids as signaling and regulatory molecules*. Springer Science & Business Media: 2011; Vol. 688.

32.   Merrill, A.; Williams, R. D., Utilization of different fatty acyl-CoA thioesters by serine palmitoyltransferase from rat brain. *Journal of lipid research* **1984,** *25* (2), 185-188.

33.   Pata, M. O.; Hannun, Y. A.; Ng, C. K. Y., Plant sphingolipids: decoding the enigma of the Sphinx. *New Phytologist* **2010,** *185* (3), 611-630.

34.   Fujino, Y.; Ohnishi, M., Sphingolipids in wheat grain. *Journal of Cereal Science* **1983,** *1* (2), 159-168.

35.   Sperling, P.; Heinz, E., Plant sphingolipids: structural diversity, biosynthesis, first genes and functions. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2003,** *1632* (1), 1-15.

36.   Dunn, T. M.; Lynch, D. V.; Michaelson, L. V.; Napier, J. A., A Post- genomic Approach to Understanding Sphingolipid Metabolism in Arabidopsis thaliana. *Annals of Botany* **2004,** *93* (5), 483-497.

37.   Markham, J. E.; Li, J.; Cahoon, E. B.; Jaworski, J. G., Separation and identification of major plant sphingolipid classes from leaves. *Journal of Biological Chemistry* **2006,** *281* (32), 22684-22694.

38.   Gurr, M. I., Harwood, J. L., & Frayn, K. N., *Lipid biochemistry*. Blackwell Science: 2002.

39.   Chitwood, D. J.; Lusby, W. R.; Thompson, M. J.; Kochansky, J. P.; Howarth, O. W., The glycosylceramides of the nematodeCaenorhabditis elegans contain an unusual, branched-chain sphingoid base. *Lipids* **1995,** *30* (6), 567-573.

40.   Fyrst, H.; Zhang, X.; Herr, D. R.; Byun, H. S.; Bittman, R.; Phan, V. H.; Harris, G. L.; Saba, J. D., Identification and characterization by electrospray mass spectrometry

of endogenous Drosophila sphingadienes. *Journal of lipid research* **2008,** *49* (3), 597-606.

41.   Seki, M.; Mori, K., Synthesis of a Prenylated and Immunosuppressive Marine Galactosphingolipid with Cyclopropane- Containing Alkyl Chains:(2S, 3R, 11S, 12R, 2‴ R, 5‴ Z, 11‴ S, 12‴ R)- Plakoside A and Its (2S, 3R, 11R, 12S, 2‴ R, 5‴ Z, 11‴ R, 12‴ S) Isomer. *European Journal of Organic Chemistry* **2001,** *2001* (20), 3797-3809.

42.   Pruett, S. T.; Bushnev, A.; Hagedorn, K.; Adiga, M.; Haynes, C. A.; Sullards, M. C.; Liotta, D. C.; Merrill, A. H., Thematic Review Series: Sphingolipids. Biodiversity of sphingoid bases ("sphingosines") and related amino alcohols. *Journal of lipid research* **2008,** *49* (8), 1621-1639.

43.   Futerman, A. H.; Hannun, Y. A., The complex life of simple sphingolipids. *EMBO reports* **2004,** *5* (8), 777-782.

44.   Markham, J. E.; Jaworski, J. G., Rapid measurement of sphingolipids from Arabidopsis thaliana by reversed- phase high- performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry. *Rapid Communications in Mass Spectrometry* **2007,** *21* (7), 1304-1314.

45.   Bohn, M.; Heinz, E.; Lüthje, S., Lipid composition and fluidity of plasma membranes isolated from corn (Zea mays L.) roots. *Archives of Biochemistry and Biophysics* **2001,** *387* (1), 35-40.

46.   Imai, H.; Yamamoto, K.; Shibahara, A.; Miyatani, S.; Nakayama, T., Determining double-bond positions in monoenoic 2-hydroxy fatty acids of glucosylceramides by gas chromatography-mass spectrometry. *Lipids* **2000,** *35* (2), 233.

47.   Lynch, D. V.; Dunn, T. M., An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. *New phytologist* **2004,** *161* (3), 677-702.

48.   Wells, G. B.; Dickson, R. C.; Lester, R. L., Isolation and composition of inositolphosphorylceramide-type sphingolipids of hyphal forms of Candida albicans. *Journal of bacteriology* **1996,** *178* (21), 6223-6226.

49.   Kazuhiro, A.; Uchiyama, R.; Itonori, S.; Sugita, M.; Fang-Sik, C.; Isogai, A.; Noriyasu, H.; Junko, H.; Takeda, T.; Kumagai, H., Structural elucidation of novel phosphocholine-containing glycosylinositol-phosphoceramides in filamentous fungi and their induction of cell death of cultured rice cells. *Biochemical Journal* **2004,** *378* (2), 461-472.

50.   Buré, C.; Cacas, J.-L.; Mongrand, S.; Schmitter, J.-M., Characterization of glycosyl inositol phosphoryl ceramides from plants and fungi by mass spectrometry. *Analytical and bioanalytical chemistry* **2014,** *406* (4), 995-1010.

51.   Carter, H. E.; Celmer, W. D.; Galanos, D. S.; Gigg, R.; Lands, W. E.; Law, J. H.; Mueller, K. L.; Nakayama, T.; Tomizawa, H.; Weber, E., Biochemistry of the sphingolipides. X. Phytoglycolipide, a complex phytosphingosine-containing lipide from plant seeds. *Journal of the American Oil Chemists Society* **1958,** *35* (7), 335-343.

52.   Kaul, K.; Lester, R. L., Characterization of Inositol-containing Phosphosphingolipids from Tobacco Leaves Isolation and Identification of Two Novel, Major Lipids: N-Acetylglucosamidoglucuronidoinositol Phosphorylceramide and Glucosamidoglucuronidoinositol Phosphorylceramide. *Plant physiology* **1975,** *55* (1), 120-129.

53.   Kaul, K.; Lester, R. L., Isolation of six novel phosphoinositol-containing sphingolipids from tobacco leaves. *Biochemistry* **1978,** *17* (17), 3569-3575.

54.   Carter, H. E.; Brooks, S.; Gigg, R.; Strobach, D. R.; Suami, T., Biochemistry of the Sphingolipids XVI. STRUCTURE OF PHYTOGLYCOLIPID. *Journal of Biological Chemistry* **1964,** *239* (3), 743-746.

55.	Carter, H.; Kisic, A., Countercurrent distribution of inositol lipids of plant seeds. *Journal of lipid research* **1969,** *10* (4), 356-362.

56.	Rennie, E. A.; Ebert, B.; Miles, G. P.; Cahoon, R. E.; Christiansen, K. M.; Stonebloom, S.; Khatab, H.; Twell, D.; Petzold, C. J.; Adams, P. D., Identification of a sphingolipid α-glucuronosyltransferase that is essential for pollen function in Arabidopsis. *The Plant Cell* **2014,** *26* (8), 3314-3325.

57.	Voxeur, A.; Fry, S. C., Glycosylinositol phosphorylceramides from Rosa cell cultures are boron- bridged in the plasma membrane and form complexes with rhamnogalacturonan II. *The Plant Journal* **2014,** *79* (1), 139-149.

58.	Simenel, C.; Coddeville, B.; Delepierre, M.; Latgé, J.-P.; Fontaine, T., Glycosylinositolphosphoceramides in Aspergillus fumigatus. *Glycobiology* **2008,** *18* (1), 84-96.

59.	Spassieva, S.; Hille, J., Plant sphingolipids today-are they still enigmatic? *Plant Biology* **2003,** *5* (02), 125-136.

60.	Borner, G. H.; Sherrier, D. J.; Stevens, T. J.; Arkin, I. T.; Dupree, P., Prediction of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A genomic analysis. *Plant Physiology* **2002,** *129* (2), 486-499.

61.	Yoshida, S.; Uemura, M., Lipid composition of plasma membranes and tonoplasts isolated from etiolated seedlings of mung bean (Vigna radiata L.). *Plant physiology* **1986,** *82* (3), 807-812.

62.	Rochester, C. P.; Kjellbom, P.; Andersson, B.; Larsson, C., Lipid composition of plasma membranes isolated from light-grown barley (Hordeum vulgare) leaves: identification of cerebroside as a major component. *Archives of biochemistry and biophysics* **1987,** *255* (2), 385-391.

63.	Takakuwa, N.; Saito, K.; Ohnishi, M.; Oda, Y., Determination of glucosylceramide contents in crop tissues and by-products from their processing. *Bioresource technology* **2005,** *96* (9), 1089-1092.

64.	Sullards, M.; Lynch, D.; Merrill, A.; Adams, J., Structure determination of soybean and wheat glucosylceramides by tandem mass spectrometry. *Journal of Mass Spectrometry* **2000,** *35* (3), 347-353.

65.	Ohnishi, M.; Ito, S.; Fujino, Y., Structural characterization of sphingolipids in leafy stems of rice. *Agricultural and biological chemistry* **1985,** *49* (11), 3327-3329.

66.	Uemura, M.; Joseph, R. A.; Steponkus, P. L., Cold acclimation of Arabidopsis thaliana (effect on plasma membrane lipid composition and freeze-induced lesions). *Plant physiology* **1995,** *109* (1), 15-30.

67.	Fujino, Y.; Ohnishi, M.; Seisuke, I., Molecular Species of Ceramide and Mono-, Di-, Tri-, and Tetraglycosy leer amide in Bran and Endosperm of Rice Grains. *Agricultural and biological chemistry* **1985,** *49* (9), 2753-2762.

68.	Lester, R. L.; Wells, G. B.; Oxford, G.; Dickson, R. C., Mutant strains of Saccharomyces cerevisiae lacking sphingolipids synthesize novel inositol glycerophospholipids that mimic sphingolipid structures. *Journal of Biological Chemistry* **1993,** *268* (2), 845-856.

69.	Cahoon, E. B.; Lynch, D. V., Analysis of glucocerebrosides of rye (Secale cereale L. cv Puma) leaf and plasma membrane. *Plant Physiology* **1991,** *95* (1), 58-68.

70.	Imai, H.; Ohnishi, M.; Kinoshita, M.; Kojima, M.; Ito, S., Structure and distribution of cerebroside containing unsaturated hydroxy fatty acids in plant leaves. *Bioscience, biotechnology, and biochemistry* **1995,** *59* (7), 1309-1313.

71.	Hannun, Y. A.; Obeid, L. M., Principles of bioactive lipid signalling: lessons from sphingolipids. *Nature reviews Molecular cell biology* **2008,** *9* (2), 139-150.

72. Wang, L.; Wang, T.; Fehr, W. R., HPLC quantification of sphingolipids in soybeans with modified palmitate content. *Journal of agricultural and food chemistry* **2006,** *54* (20), 7422-7428.

73. Bartke, N.; Fischbeck, A.; Humpf, H. U., Analysis of sphingolipids in potatoes (Solanum tuberosum L.) and sweet potatoes (Ipomoea batatas (L.) Lam.) by reversed phase high- performance liquid chromatography electrospray ionization tandem mass spectrometry (HPLC- ESI- MS/MS). *Molecular nutrition & food research* **2006,** *50* (12), 1201-1211.

74. Dharmawardhane, S.; Rubinstein, B.; Stern, A. I., Regulation of transplasmalemma electron transport in oat mesophyll cells by sphingoid bases and blue light. *Plant physiology* **1989,** *89* (4), 1345-1350.

75. Bille, J.; Weiser, T.; Bentrup, F. W., The lysolipid sphingosine modulates pyrophosphatase activity in tonoplast vesicles and isolated vacuoles from a heterotrophic cell suspension culture of Chenopodium rubrum. *Physiologia Plantarum* **1992,** *84* (2), 250-254.

76. Brodersen, P.; Petersen, M.; Pike, H. M.; Olszak, B.; Skov, S.; Ødum, N.; Jørgensen, L. B.; Brown, R. E.; Mundy, J., Knockout of Arabidopsis accelerated-cell-death11 encoding a sphingosine transfer protein causes activation of programmed cell death and defense. *Genes & Development* **2002,** *16* (4), 490-502.

77. Simanshu, D. K.; Zhai, X.; Munch, D.; Hofius, D.; Markham, J. E.; Bielawski, J.; Bielawska, A.; Malinina, L.; Molotkovsky, J. G.; Mundy, J. W., Arabidopsis accelerated cell death 11, ACD11, is a ceramide-1-phosphate transfer protein and intermediary regulator of phytoceramide levels. *Cell reports* **2014,** *6* (2), 388-399.

78. Fukasawa, M.; Nishijima, M.; Hanada, K., Genetic evidence for ATP-dependent endoplasmic reticulum-to-Golgi apparatus trafficking of ceramide for sphingomyelin synthesis in Chinese hamster ovary cells. *The Journal of cell biology* **1999,** *144* (4), 673-685.

79. Kumagai, K.; Yasuda, S.; Okemoto, K.; Nishijima, M.; Kobayashi, S.; Hanada, K., CERT mediates intermembrane transfer of various molecular species of ceramides. *Journal of Biological Chemistry* **2005,** *280* (8), 6488-6495.

80. Hanada, K.; Kumagai, K.; Yasuda, S.; Miura, Y.; Kawano, M.; Fukasawa, M.; Nishijima, M., Molecular machinery for non-vesicular trafficking of ceramide. *Nature* **2003,** *426* (6968), 803-809.

81. Hanada, K., Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2003,** *1632* (1), 16-30.

82. Mizutani, Y.; Kihara, A.; Igarashi, Y., LASS3 (longevity assurance homologue 3) is a mainly testis-specific (dihydro) ceramide synthase with relatively broad substrate specificity. *Biochemical Journal* **2006,** *398* (3), 531-538.

83. Mizutani, Y.; Kihara, A.; Igarashi, Y., Mammalian Lass6 and its related family members regulate synthesis of specific ceramides. *Biochemical Journal* **2005,** *390* (1), 263-271.

84. Pewzner-Jung, Y.; Ben-Dor, S.; Futerman, A. H., When do Lasses (longevity assurance genes) become CerS (ceramide synthases)? Insights into the regulation of ceramide synthesis. *Journal of Biological Chemistry* **2006,** *281* (35), 25001-25005.

85. Markham, J. E.; Molino, D.; Gissot, L.; Bellec, Y.; Hématy, K.; Marion, J.; Belcram, K.; Palauqui, J.-C.; Satiat-JeuneMaître, B.; Faure, J.-D., Sphingolipids containing very-long-chain fatty acids define a secretory pathway for specific polar plasma membrane protein targeting in Arabidopsis. *The Plant Cell* **2011,** *23* (6), 2362-2378.

86.   Futerman, A. H.; Pagano, R. E., Determination of the intracellular sites and topology of glucosylceramide synthesis in rat liver. *Biochemical Journal* **1991,** *280* (2), 295-302.

87.   Gillard, B. K.; Clement, R. G.; Marcus, D. M., Variations among cell lines in the synthesis of sphingolipids in de novo and recycling pathways. *Glycobiology* **1998,** *8* (9), 885-890.

88.   Tettamanti, G.; Bassi, R.; Viani, P.; Riboni, L., Salvage pathways in glycosphingolipid metabolism. *Biochimie* **2003,** *85* (3), 423-437.

89.   Kitatani, K.; Idkowiak-Baldys, J.; Hannun, Y. A., The sphingolipid salvage pathway in ceramide metabolism and signaling. *Cellular signalling* **2008,** *20* (6), 1010-1018.

90.   Riboni, L.; Prinetti, A.; Bassi, R.; Tettamanti, G., Formation of bioactive sphingoid molecules from exogenous sphingomyelin in primary cultures of neurons and astrocytes. *FEBS letters* **1994,** *352* (3), 323-326.

91.   Pennell, R. I.; Lamb, C., Programmed cell death in plants. *The Plant Cell* **1997,** *9* (7), 1157.

92.   Collazo, C.; Chacón, O.; Borrás, O., Programmed cell death in plants resembles apoptosis of animals. *Biotecnología Aplicada* **2006,** *23* (1), 1-10.

93.   Liang, H.; Yao, N.; Song, J. T.; Luo, S.; Lu, H.; Greenberg, J. T., Ceramides modulate programmed cell death in plants. *Genes & development* **2003,** *17* (21), 2636-2641.

94.   Peters, J.; Chin, C.-k., Potassium loss is involved in tobacco cell death induced by palmitoleic acid and ceramide. *Archives of biochemistry and biophysics* **2007,** *465* (1), 180-186.

95.   Wang, W.; Yang, X.; Tangchaiburana, S.; Ndeh, R.; Markham, J. E.; Tsegaye, Y.; Dunn, T. M.; Wang, G.-L.; Bellizzi, M.; Parsons, J. F., An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in Arabidopsis. *The Plant Cell* **2008,** *20* (11), 3163-3179.

96.   Xiao, S.; Ellwood, S.; Findlay, K.; Oliver, R. P.; Turner, J. G., Characterization of three loci controlling resistance of Arabidopsis thaliana accession Ms- 0 to two powdery mildew diseases. *The Plant Journal* **1997,** *12* (4), 757-768.

97.   Huitema, K.; van den Dikkenberg, J.; Brouwers, J. F.; Holthuis, J. C., Identification of a family of animal sphingomyelin synthases. *The EMBO journal* **2004,** *23* (1), 33-44.

98.   Denny, P. W.; Shams-Eldin, H.; Price, H. P.; Smith, D. F.; Schwarz, R. T., The protozoan inositol phosphorylceramide synthase a novel drug target that defines a new class of sphingolipid synthase. *Journal of Biological Chemistry* **2006,** *281* (38), 28200-28209.

99.   Heidler, S. A.; Radding, J. A., Inositol phosphoryl transferases from human pathogenic fungi. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2000,** *1500* (1), 147-152.

100.   Bromley, P. E.; Li, Y. O.; Murphy, S. M.; Sumner, C. M.; Lynch, D. V., Complex sphingolipid synthesis in plants: characterization of inositolphosphorylceramide synthase activity in bean microsomes. *Archives of Biochemistry and Biophysics* **2003,** *417* (2), 219-226.

101.   Nagiec, M. M.; Nagiec, E. E.; Baltisberger, J. A.; Wells, G. B.; Lester, R. L.; Dickson, R. C., Sphingolipid synthesis as a target for antifungal drugs complementation of the inositol phosphorylceramide synthase defect in a mutant strain of Saccharomyces cerevisiae by the AUR1 gene. *Journal of Biological Chemistry* **1997,** *272* (15), 9809-9817.

102.   Heidler, S. A.; Radding, J. A., The AUR1 gene in Saccharomyces cerevisiae encodes dominant resistance to the antifungal agent aureobasidin A (LY295337). *Antimicrobial agents and chemotherapy* **1995,** *39* (12), 2765-2769.

103.   Mina, J., Thesis: Exploring IPC synthase: a potential anti-leischmanial drug target. **2010**.

104.   Mina, J.; Okada, Y.; Wansadhipathi-Kannangara, N.; Pratt, S.; Shams-Eldin, H.; Schwarz, R.; Steel, P.; Fawcett, T.; Denny, P., Functional analyses of differentially expressed isoforms of the Arabidopsis inositol phosphorylceramide synthase. *Plant molecular biology* **2010,** *73* (4-5), 399-407.

105.   Sutterwala, S. S.; Hsu, F. F.; Sevova, E. S.; Schwartz, K. J.; Zhang, K.; Key, P.; Turk, J.; Beverley, S. M.; Bangs, J. D., Developmentally regulated sphingolipid synthesis in African trypanosomes. *Molecular microbiology* **2008,** *70* (2), 281-296.

106.   Sigal, Y. J.; McDERMOTT, M. I.; Morris, A. J., Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions. *Biochemical Journal* **2005,** *387* (2), 281-293.

107.   Stukey, J.; Carman, G. M., Identification of a novel phosphatase sequence motif. *Protein science* **1997,** *6* (2), 469-472.

108.   Mina, J. G.; Mosely, J. A.; Ali, H. Z.; Denny, P. W.; Steel, P. G., Exploring Leishmania major Inositol Phosphorylceramide Synthase (LmjIPCS): Insights into the ceramide binding domain. *Organic & Biomolecular Chemistry* **2011,** *9* (6), 1823-1830.

109.   Bi, F.-C.; Liu, Z.; Wu, J.-X.; Liang, H.; Xi, X.-L.; Fang, C.; Sun, T.-J.; Yin, J.; Dai, G.-Y.; Rong, C., Loss of ceramide kinase in Arabidopsis impairs defenses and promotes ceramide accumulation and mitochondrial H2O2 bursts. *The Plant Cell* **2014,** *26* (8), 3449-3467.

110.   Tafesse, F. G.; Ternes, P.; Holthuis, J. C., The multigenic sphingomyelin synthase family. *Journal of Biological Chemistry* **2006,** *281* (40), 29421-29425.

111.   Stuart, K.; Brun, R.; Croft, S.; Fairlamb, A.; Gürtler, R. E.; McKerrow, J.; Reed, S.; Tarleton, R., Kinetoplastids: related protozoan pathogens, different diseases. *The Journal of clinical investigation* **2008,** *118* (4), 1301-1310.

112.   Croft, S. L.; Barrett, M. P.; Urbina, J. A., Chemotherapy of trypanosomiases and leishmaniasis. *Trends in parasitology* **2005,** *21* (11), 508-512.

113.   De Lucca, A. J., Harmful fungi in both agriculture and medicine. *Revista iberoamericana de micología* **2007,** *24* (1), 3.

114.   Gugnani, H. C., Ecology and taxonomy of pathogenic aspergilli. *Front Biosci* **2003,** *8* (1-3), 346.

115.   Dvoráčková, I.; Pichova, V., Pulmonary interstitial fibrosis with evidence of aflatoxin B1 in lung tissue. **1986**.

116.   Thomas, P., Fungal infections of the cornea. *Eye* **2003,** *17* (8), 852-862.

117.   Vicente, M.; Basilio, A.; Cabello, A.; Peláez, F., Microbial natural products as a source of antifungals. *Clinical Microbiology and Infection* **2003,** *9* (1), 15-32.

118.   Deray, G., Amphotericin B nephrotoxicity. *Journal of Antimicrobial Chemotherapy* **2002,** *49* (suppl 1), 37-41.

119.   Snelders, E.; van der Lee, H. A.; Kuijpers, J.; Rijs, A. J. M.; Varga, J.; Samson, R. A.; Mellado, E.; Donders, A. R. T.; Melchers, W. J.; Verweij, P. E., Emergence of azole resistance in Aspergillus fumigatus and spread of a single resistance mechanism. *PLoS Med* **2008,** *5* (11), e219.

120.   Han, H.; Yu, Q.; Owen, M. J.; Cawthray, G. R.; Powles, S. B., Widespread occurrence of both metabolic and target- site herbicide resistance mechanisms in Lolium rigidum populations. *Pest management science* **2016,** *72* (2), 255-263.

121.   Carlson, H. L.; Hill, J. E., Wild oat (Avena fatua) competition with spring wheat: plant density effects. *Weed Science* **1985,** *33* (2), 176-181.

122.   Bajwa, A. A.; Jabran, K.; Shahid, M.; Ali, H. H.; Chauhan, B. S., Eco-biology and management of Echinochloa crus-galli. *Crop Protection* **2015,** *75*, 151-162.

123.   Chandler, J. W., Cotyledon organogenesis. *Journal of Experimental Botany* **2008,** *59* (11), 2917-2931.

124.   Taiz, L.; Zeiger, E.; Møller, I. M.; Murphy, A., *Plant physiology and development*. Sinauer Associates, Incorporated: 2015.

125.   Nelson, T.; Dengler, N., Leaf vascular pattern formation. *The Plant Cell* **1997,** *9* (7), 1121.

126.   Zimmermann, M. H.; Tomlinson, P. B., The vascular system of monocotyledonous stems. *Botanical Gazette* **1972,** *133* (2), 141-155.

127.   McSteen, P., Auxin and monocot development. *Cold Spring Harbor perspectives in biology* **2010,** *2* (3), a001479.

128.   Mandala, S. M.; Thornton, R. A.; Rosenbach, M.; Milligan, J.; Garcia-Calvo, M.; Bull, H. G.; Kurtz, M. B., Khafrefungin, a novel inhibitor of sphingolipid synthesis. *Journal of Biological Chemistry* **1997,** *272* (51), 32709-32714.

129.   Mandala, S. M.; Thornton, R. A.; Milligan, J.; Rosenbach, M.; Garcia-Calvo, M.; Bull, H. G.; Harris, G.; Abruzzo, G. K.; Flattery, A. M.; Gill, C. J., Rustmicin, a potent antifungal agent, inhibits sphingolipid synthesis at inositol phosphoceramide synthase. *Journal of Biological Chemistry* **1998,** *273* (24), 14942-14949.

130.   Yano, T.; Aoyagi, A.; Kozuma, S.; Kawamura, Y.; Tanaka, I.; Suzuki, Y.; Takamatsu, Y.; Takatsu, T.; Inukai, M., Pleofungins, novel inositol phosphorylceramide synthase inhibitors, from Phoma sp. SANK 13899. *Journal of Antibiotics* **2007,** *60* (2), 136.

131.   Ohnuki, T.; Yano, T.; Takatsu, T., Haplofungins, new inositol phosphorylceramide synthase inhibitors, from Lauriomyces bellulus SANK 26899 II. Structure elucidation. *Journal of Antibiotics* **2009,** *62* (10), 551-557.

132.   Mina, J. G.; Pan, S.-Y.; Wansadhipathi, N. K.; Bruce, C. R.; Shams-Eldin, H.; Schwarz, R. T.; Steel, P. G.; Denny, P. W., The Trypanosoma brucei sphingolipid synthase, an essential enzyme and drug target. *Molecular and biochemical parasitology* **2009,** *168* (1), 16-23.

133.   Figueiredo, J. M.; MENDONÇA-PREVIATO, L.; PREVIATO, J. O.; HEISE, N., Characterization of the inositol phosphorylceramide synthase activity from Trypanosoma cruzi. *Biochemical Journal* **2005,** *387* (2), 519-529.

134.   Brown, C. Inhibitors of Leishmania major Inositol Phosphorylceramide Synthase: New Therapies for Leishmaniasis. Durham University, 2016.

CHAPTER 2

CONSTRUCTION OF PLANT IPCS COMPLEMENTED YEAST STRAINS AND ASSAY VALIDATION

2.1 Chapter synopsis

In this chapter, bioinformatics analysis of the plant IPCS was carried out. Phylogenetic analyses of IPCS revealed a divide in clustering between monocots and dicots. Conserved domains previously identified in animal sphingomyelin synthases and protozoan IPCSs were also identified among the plant IPCSs. Analyses of the promoter region of the three IPCS isoforms in *Arabidopsis thaliana* (*At*IPCS1-3) and *Oryza sativa* (*Os*IPCS1-3) were carried out to identify binding sites of transcription factors. The three IPCS isoforms in *A. thaliana* and *O. sativa* were cloned into an expression vector in *Saccharomyces cerevisiae* and their activities characterized. These yeast strains were formatted into an assay and the assay conditions were validated to screen for inhibitors of plant IPCSs.

2.2 Bioinformatics analysis of plant IPCS

Identification of inhibitors that show differential selectivity for *At*IPCSs over *Os*IPCSs (vice versa) is possible if there are structural differences between the orthologues which arise from differences in the amino acid sequence. To find out these differences, structural analysis of *At*IPCS1-3 and *Os*IPCS1-3 was carried out.

The identification of functional IPCS orthologues in *Oryza sativa* was achieved by a BLAST search of phytozyme database using the coding sequence of *At*IPCS1 which yielded three homologues. Analyses of the coding sequence for the three isoforms found in *Arabidopsis thaliana* and *Oryza sativa*, showed a difference in the number and length of exons and introns (Figure 2-1). *At*IPCS1 possesses 11 exons and 10 introns, whilst *At*IPCS2 has 12 exons and 11 introns and *At*IPCS3 possesses 11 exons and 10 introns. The analogous orthologues in *O. sativa* are approximately two times longer and are composed of 12 exons and 11 introns.
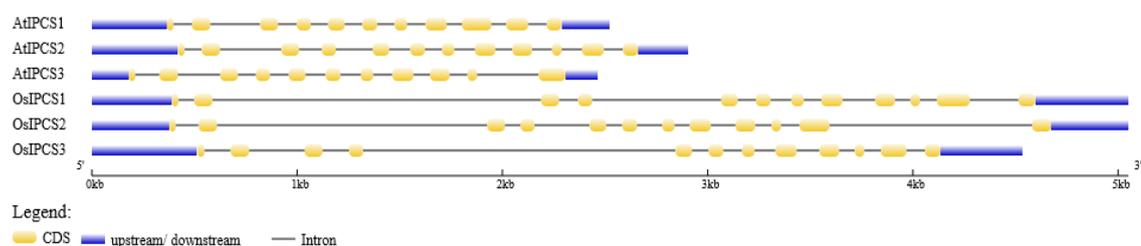
Figure 2-1: Schematic representation of exon and intron organisation of IPCS in *O. sativa* and *A. thaliana*; black line and yellow boxes indicates introns and exon respectively, with the blue boxes representing the 5' and 3' UTRs.

From work done by Mina *et al.* looking at the differential expression of the three homologues of IPCS in *A. thaliana*, it is known that the second orthologue of IPCS is the most highly expressed by a log 2-fold difference per µg RNA extracted[1]. It is highly expressed in stem, flower and siliques and to an even greater extent (one log fold increase per µg of RNA extracted) in roots, cauline leaf, and rosette leaf[2], pointing to a key role in these tissues.

2.2.1 Phylogenetic analysis

With IPCS2 begin the most highly expressed isoform in *A. thaliana*, a phylogenetic analyses of the evolutionary distance of this particular isoform in different plant species, protozoa and fungi was carried out to gain an insight into how closely related the orthologues were to each other. Unsurprisingly, the result shows that the plant IPCS are clustered together and are independent of the protozoan and fungi IPCS (Figure 2-2). Interestingly, the clustering of IPCS2 in plants predominantly shows a divide between monocots and dicots, which is promising in the aim to identify inhibitors that show differential selectivity for either dicots over monocots or monocots over dicots.

An examination of the phylogenetic relationship between the three isoforms in *A. thaliana* and *O. sativa* show that IPCS1 and IPCS2 in both plants are more closely related compared to each other than IPCS3. Access was given by Bayer Crop Science to the genome of the noxius weed *Echinochloa crus-galli,* known to cost billions of dollars of losses to rice production, but the sequence was not complete and attempts to reconstruct the sequence by primer walking failed. However, analyses were carried out to see how closely related it is to the IPCS found in *O. sativa* and *A. thaliana* (Figure 2-3). From the phylogenetic tree it was clear that the *E. crus-galli* IPCS is distantly related and form a cluster separate from *At*IPCS1-3, *Os*IPCS1 and 2. Interestingly, *Os*IPCS3 is placed on a separate evolutionary branch from *Os*IPCS1 and 2, which indicates that it is very different from these two isoforms and is shown to be more closely related to *Ech*PCS1. The reason for this is not immediately clear from differences in the amino acid sequence between *Os*IPCS1, 2 and 3. However, the evolutionary distance between *O. sativa* isoforms and *Ech*IPCS1 is particularly promising, as it indicated that in the search for new herbicide mode of action, selective IPCS inhibitors targeting *E. crus-galli* over *O. sativa* IPCS could potentially be developed.
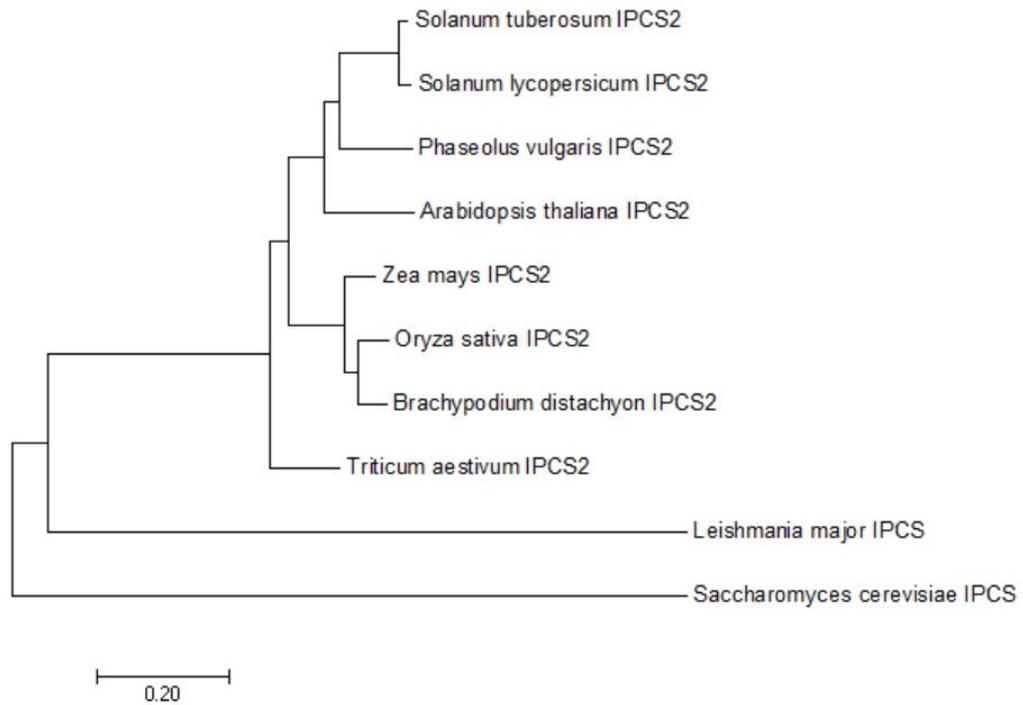
Figure 2-2: Phylogenetic tree generated from the amino acid sequence of IPCS from plant, protozoa and fungi. The evolutionary distances were calculated using the neighbourhood joining method and analysed in MEGA7; drawn to scale with branch lengths in the same units as those of the evolutionary distances used to construct the phylogenetic tree in units of the number of amino acid substitution per site.
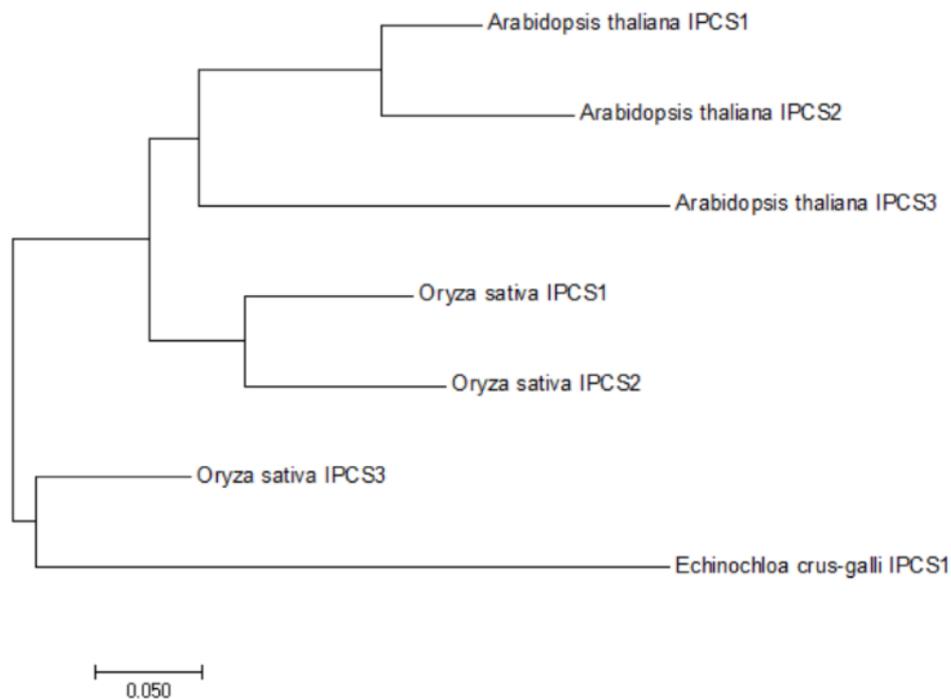
Figure 2-3: Phylogenetic tree constructed using IPCS amino acid sequence from *A. thaliana, O. sativa* and *E. crus-galli.* The evolutionary distances were calculated using the neighbourhood joining method and analysed in MEGA7; drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to construct the phylogenetic tree. The unit of measurement quantifies the number of amino acid substitutions per site.

2.2.2 Prediction of transmembrane topology

Using the transmembrane hidden Markov model[3] (TMHMH) server, the number, location of the transmembrane regions and protein topology were identified for *Os*IPCS1-3, *At*PCS1-3 and *Ech*IPCS. It was found that *At*IPCS1, *At*IPCS2, *At*IPCS3 and *Os*IPCS3 were predicted to have five transmembrane regions as opposed to six, because the probability of the first transmembrane region was below 0.5 so was not significant enough to use as a reliable indication of its presence (Figure 2-4). *Ech*IPCS was predicted to have three transmembrane regions because of the low probability of the other three transmembrane regions (data not shown). However, *Os*IPCS1 and *Os*IPCS2 were both predicted to have six transmembrane regions with a probability of the prediction being between 0.8 and 1 (Figure 2-3). The defined transmembrane regions are also in agreement with the predicted transmembrane region and topology carried out using predictprotein[4] server, with an inside-inside orientation of the N and C terminal which is in agreement with Bang's model and the lumenal topology of the active site in the Golgi[5].

From the structural analysis, it seems that there is a structural difference between the plant IPCS from *A. thaliana*, *O. sativa* and *E. crus- gali,* with *At*IPCS1-3 and *Os*IPCS3 predicted to have five transmembranes and an extracellular domain, whilst *Os*IPCS1 and 2 have six transmembrane regions. The differences in the structural predictions particularly in regard to *At*IPCS2 and *Os*IPCS2 arise due to the presence of hydrophobic resides such as phenylalanine (Phe[16]) and tyrosine (Tyr[22] and Tyr[33]) at the N-terminus of *At*IPCS2 in place of valine (Val[16] and Val[22]) and leucine (Leu[33]) at the N-terminus of *Os*IPCS2. Additional variation in amino acid sequence at the N-terminus include the presence of serine (Ser[9]), cysteine (Cys[17]) and asparagine Asn[30] (*At*IPCS2) in place of alanine (Ala[9]), tyrosine (Tyr[17]) and lysine (Lys[30]) respectively (*Os*IPCS2). These differences affect topology prediction using TMHHM because the prediction is based on

the positive-inside rule with positively charged residues positioned in the cytosol and hydrophobic residues located in the transmembrane region. For AtIPCS2 the existence of hydrophobic residues makes it difficult for the THMMH programme to ascribe the location of the N-terminus and subsequently affects the topology prediction of the first transmembrane. Topology prediction using MEMSAT3 a program that relies on a neural network to determine the location of amino acid residues in a protein structure[6]. Utilising MEMSAT3, *Ech*IPCS, *At*IPCS2 and *Os*IPCS2 (data not shown) were predicted to be six transmembrane spanning proteins, however, with no experimental evidence of IPCS structure available at the moment, there is no way of knowing if the variation in amino acid sequence results in a difference in structure between the orthologues. If the structural differences predicted by TMHMM are correct, then the possibility of the identification of selective inhibitors is much more likely.

Using the structural prediction of *Os*IPCS1 and 2 which is in agreement with Bang's model[7], and the identified conserved domains identified in yeast[5], and human sphingomyelin synthase[7], the transmembrane regions, domains, and catalytic triad (His, His and Asp) were identified from the multiple sequence alignment of plant IPCSs amino acid coding sequence (Figure 2-4).

From the alignment, the plant IPCSs are shown to be have highly conserved regions at the amino acid level in the transmembrane regions and domains, however, it is clear that at the N-terminal, there is variation at the amino acid level with fewer amino acids observed to be identical but do share similar chemical properties. This difference could account for the observed difference in the structural prediction. *Ech*IPCS does not align with the other IPCSs, and the catalytic triad is unidentified upon alignment, which is probably more likely to be due to the sequence being incomplete, than to a real structural difference at the active site.

Figure 2-3: TMHMM posterior probabilities for (A) *Os*IPCS2 and (B) *At*IPCS2 showing the six predicted transmembranes and their topologies.

```
EchIPCS1   1 MSPFYLARSASKLVRRITSEVSVELKILSEKWRLLLAGLVFQYIHGLAARGVHYLHRPGP
AtIPCS3    1 -MPVYVDREAPKLWRRIYSEATLEASLLAEKWKLVLAGLVFQYIHGLAAHGVHYLHRPGP
OsIPCS3    1 MSALYLARGASKVVRRITSETSVELKILTEKWQLLLAGLVFQYIHGLAARGVHYLHRPGP
AtIPCS1    1 -MTLYIRREASKLWRRFCSEITTEIGLIAENWKYLLAGLICQYIHGLAARGVHYIHRPGP
AtIPCS2    1 -MTLYIRRESSKLWKRFCSEISTEIGLIAENWKYLLAGLICQYIHGLAAKGVHYIHRPGP
OsIPCS1    1 -MAVYIAREATKLWRKVCAEIAVELQLLFEKWRLLLAGLVFQYIHGLAARGVHYLHRPGP
OsIPCS2    1 -MTIYIAREASKVWRKVTTETSVEISLLREKWGLLLAGLIVFQYIHGLAARGVHYLHRPGP
```

```
EchIPCS1  61 TLQDLGFMILPELGKERGYISETLFTFIFLSFVLWTFHPFILHIKRFYTVLIWRRVLAFL
AtIPCS3   60 TLQDAGFFILPALGQDKAFFSETVFVTIFGSFILWTFHPFVSHSKKICTVLIWCRVFVYL
OsIPCS3   61 TLQDLGFMILPELGKERGYISETLFTFIFLSFVLWTFHPFILQIKRFYTVLIWRRVLAFL
AtIPCS1   60 TLQDSGFFVLPELGQDKGFISETVFTCVFLSFFLWTFHPFIVKSKKIYTVLIWCRVLAFL
AtIPCS2   60 TLQDLGFFILPELGQERSYISETVFTSVFLSFFLWTFHPFILKTKKIYTVLIWCRVLAFL
OsIPCS1   60 LLQDLGFMALPELGQDKGYVSESVFTFIFTSFLWSFHPFIYHSKRFYTVLLWRRVLAFL
OsIPCS2   60 LLQDLGFMALPEFGQDKGYLSESIFASIFASFVLWTFHPFIYHSKRFYTVLIWRRVLAFL
```

```
EchIPCS1 121 CASQFLRIITFYSTQLPGPNYHCREGSPLARLPPPKNAAEVLL-DFPKGGEY-MAVVTYF
AtIPCS3  120 AASQSLRIITFFATQLPGPNYHCREGSKLAKIPPPKNVLEVLLINFPDGVIYGCGDLIFS
OsIPCS3  121 CASQFLRIVTFYSTQLPGPNYHCREGSALARLPHPQNVAEVLLINFPRGVIYGCGDLIFS
AtIPCS1  120 VACQFLRVITFYSTQLPGPNYHCREGSELARLPRPHNVLEVLLLNFPRGVIYGCGDLIFS
AtIPCS2  120 VACQFLRVITFYSTQLPGPNYHCREGSKVSRLPWPKSALEVLEIN-PHGVMYGCGDLIFS
OsIPCS1  120 SASQFLRIITFYSTQLPGPNYHCREGSKMATLPPPHNVLEVLLINFPRGVLFGCGDLIFS
OsIPCS2  120 VASQVLRIITFYSTQLPGPNYHCREGSKLATLPPPNNVFEVLLINFPRGVLFGCGDLIFS
```

```
EchIPCS1 179 HPTFSLLVFVITYQKYGSIRFVKYASNGALLLQSSLIISSRKLLQCLLLLHGIRVVIFVF
AtIPCS3  180 SHTIFTLVFVRTYQRYGTRRWIKHLAWLMAVIQSILIIASRKHYTVDIVVAWYTVNLVMF
OsIPCS3  181 SHMIFTIVFVVTYQKYGNIRFIKMLAWCIAIAQSLLIIASRKHYSVDVVVAWYTVNLVVF
AtIPCS1  180 SHMIFTLVFVRTYQKYGSKRFIKLLGWVIAILQSLLIIASRKHYTVDVVVAWYTVNLVVF
AtIPCS2  179 SHMIFTLVFVRTYQKYGTKRFIKLFGWLTAIVQSLLIIASRKHYSVDVVVAWYTVNLVVF
OsIPCS1  180 SHMIFTLVFVRTYHKYGSKRLIKILAWLMAIQSLLIIASRKHYSVDVVVAWYTVNLVVF
OsIPCS2  180 SHMIFTLVFVRTYHKYGSKRFVKLLAWFMAIVQSLLIIASRKHYSVDVVVAWYTVNLVVF
```

```
EchIPCS1 239 VCGNRNLQL-------------PIDQQD----------------------LLYNYSPVS
AtIPCS3  240 YVDSKLPEMAERSSGPSPTPLLPLSTKD-------------SKNKSKEDHQRLLNENNVA
OsIPCS3  241 FVDKKLTELPDRSAG--STSVLPVSIKE------------KDSKLKEDKTRMLNGNSVD
AtIPCS1  240 FIDKKLPELPDRT-----TALLPVISKD----------------RTKEESHKLLNGNGVD
AtIPCS2  239 CLDKKLPELPDRT-----AVLLPVISKD----------------RTKEENHKLLNGNGVD
OsIPCS1  240 FIDNKLPEMPDRTNG---SSLLPVTAKDKDGRTKEELHKLEKDCKMKEEFHKLLNGNTVD
OsIPCS2  240 FVDNKLPEMPDRTNG---VPLLPLSTREKDGRLKE-----EKDSRLKEEFHKLLNGNHGD
```

```
EchIPCS1 263 DEGQRLT--------------------------
AtIPCS3  287 D----------------DH--------------
OsIPCS3  286 SADWRPWTQMNGKHIENGNH-----LDTETTKT-
AtIPCS1  279 PADRRPRAQVNGK-DSNGGH-----TDNATNGT-
AtIPCS2  278 PADWRPRAQVNGKIDSNGVH-----TDNTMNGA-
OsIPCS1  297 STDRRQRVQMNGKHGEDINH---TLSDATPNGT-
OsIPCS2  292 PTDRRQRAQMNGRHDEDINHAHSTLSDAAVNGGT
```

Figure 2-4: Alignment of *At*IPCS1-3, *Os*IPCS1-3 and *Ech*IPCS with transmembrane regions and conserved domains highlighted; catalytic residues are also highlighted in red.

2.2.3 Identification of cis-elements and transcription factors upstream of *At*IPCS promoter

To get an insight into the biological processes that IPCS might be involved in, an investigation of the promoter region was carried out to identify cis-regulatory elements to which transcription factors can bind, eliciting a response in plants. The IPCS promotor region (1000 bp upstream of transcription initiation) was examined using PlantPAN2 to identify motifs and transcription factors involved in the response to biotic and abiotic stress (Figure 2-5). A number of WRKY transcriptional factors binding sites were identified in the promoter region of *At*IPCS1-3, and of particular interest was WRKY4 which is involved in plant pathogen response, having a negative effect on plant response to biotrophic pathogens, and a positive effect on plant resistance to necrotrophic pathogens[8]. Also identified were WRKY18 and WRKY60, which are known to enhance plant sensitivity to abscisic acid in relation to root development, whilst also enhancing sensitivity to salt and osmotic stress[9].

The binding site of the *cis*-regulatory elements LS5ATPR1 and LS7ATPR1 were identified and are well characterized positive salicylic acid-inducible elements recognized by TGA factors[10] (Figure 2-5). Binding sites for TGA factors are frequently found in promoters of stress-regulated plant genes[11]; TGA2 and TGA3 interaction with LS7 is enhanced by binding with non-expresser of pathogenesis related genes protein (NPR1)[12, 13], a defence response gene associated with the induction of systematic acquired resistance in plant defence response to pathogen[14].

The same analysis was carried out for *Os*IPCS1-3 and no WRKY transcriptional factor binding sites were found, but the search did yield different basic region leucine zipper binding sites with unknown targets. Expansion of the region upstream the transcription start site to 1500 bp and subsequent analyses using the PlantCARE database, generated

several motifs which are shared between *A. thaliana* and *O. sativa* IPCSs; *cis* acting elements involved in stress and pathogen response were identified, as were *cis*-regulatory elements inducible in response to the hormones abscisic acid and gibberellin (Table 2-1) indicating that the modulation of plant IPCSs expression is involved in plant response to abiotic and biotic stress.

Figure 2-5: Transcription factor binding motif sequence located 1000bp upstream of the transcription start site; the transcription factor binding sites highlighted are conserved in *At*IPCS1-3; shown above is the promoter region of *At*IPCS1; table of motifs can be found in Appendix A.

It is clear from the analyses that the IPCS from *A. thaliana* and *O. sativa* share the same transcriptional binding elements, and at the amino acid level the sequence is highly conserved. *Os*IPCS1 shares 70% sequence identity and 81% similarity to *At*IPCS1; *At*IPCS1sequence identity and similarity was between 69-70% and 79-80% for *Os*IPCS

2 and 3 respectively. Sequence identity between *Os*IPCS1 and *Os*IPCS2 is 82% and the sequence similarity is 89%; for *Os*IPCS3 the sequence identity and similarity to *Os*IPCS1 is 70% and 84% respectively. The sequences identify between *At*IPCS1 and *At*IPCS2 is 86% and the similarity is 93%, whilst *At*IPCS3 shares a lower sequence identity with *At*IPCS1 of 69% and a sequence similarity of 82%.

Mutagenesis studies of fungal IPC (AUR1p) have shown that a single amino acid mutation is enough to alter the sensitivity of the enzyme to Aureobasidin A (AbA). Mutation of Leu-137 to phenylalanine and His-157 to tyrosine in the *AUR1* gene sequence resulted in AbA resistance in *S. cerevisiae*[15]. This is corroborated by a similar study carried out by Okado *et al.* which revealed the ability of AUR1 to become insensitive to AbA upon the mutation of Phe-158 to a tyrosine[16]. In addition, a mutation changing glycine-240 to a cysteine residue in *the* AUR1 homologue in *Schizosaccharomyces pombe* has been found to be responsible for the acquisition of AbA resistance in both fungal species[17, 18]. Likewise, the corresponding residue 275 in the aurA gene of *Aspergillus nidulans*, a homolog of yeast AUR1 has been found to confer a high level of AbA resistance[19]. These studies show that IPCS sensitivity to an inhibitor can be altered by a single change of an amino acid residue and stimulates the consideration of an inhibitor being selective for different orthologues of IPCS in the same plant, which could open up the possibility of creating a cocktail of herbicides that targets two out of the three orthologues, thereby reducing the likelihood of herbicide resistance. The possibility of selective inhibition between the very closely related IPCS isoforms is possible, and has been shown to be the case by the selective inhibition of MAP Kinase p38α over p38β, which share a sequence identity of 75% at the amino acid level[20].

In the next section, work carried out to express *At*IPCS1-3 and *Os*IPCS1-3 and characterise the activity of plant IPCS in a *S. cerevisiae* mutant line is presented, with the

aim of utilising the transgenic yeast lines in a high throughput screen for inhibitors. This will enable the identification of compounds which show differential selectivity for either IPCS from *A. thaliana* or *O. sativa.*

Table 2-1: Identification of conserved cis-elements located upstream of the transcription start site of *At*IPCS1-3 and *Os*IPCS1-3

Motif name	Sequence function
ABRE	*Cis*-acting element involved in abscisic acid responsiveness
Box W1	Fungal elicitor responsive element
MBS	MYB binding site involved in drought-inducibility
TCA element	*Cis*-acting regulatory element involved in salicylic acid responsiveness
Circadian (CBS)	*Cis*-acting regulatory element involved in circadian control
ARE	*Cis*-acting regulatory element essential for anaerobic induction
P-Box	Gibberellin responsive element
TC-rich repeats	*Cis*-acting element involved in defence and stress responsiveness
RY-element	*Cis*-acting regulatory element involved in seed-specific regulation
HSE	Heat responsive elements

2.3 Cloning and expression of plant IPCS to complement a mutant yeast strain

To analyse the multi-transmembrane plant IPCS and format a high throughput screening (HTS) assay, the coding sequences were cloned into suitable expression vectors to complement a yeast mutant as described below.

Saccharomyces cerevisiae pdr genes encode extrusion pumps responsible for pleiotropic drug resistance[21]. Fungi invoke resistance by the up-regulation of these genes to facilitate the efflux of cytotoxic drugs, thus leading to decreased drug accumulation and diminished concentrations[22]. In order to maximize the possibility of identifying IPCS inhibitors, it was necessary to prepare a suitable *S. cerevisiae* strain lacking these efflux pumps.

A mutant yeast strain lacking fungal IPCS, AUR1 (MATa his3Δ1leu2Δ0ura3Δ0 aur1::HIS3) and four genes conferring drug hypersensitivity (*pdr1Δ::KanMX4, pdr3Δ::KanMX4, pdr16Δ::KanMX4 and pdr17Δ::KanMX4*) were created by Prof. Michael Stark (University of Dundee). This strain can survive by the introduction of the pRS316-AUR1 plasmid, (Prof. Michael Stark), resulting in the expression AUR1 under a URA3 selectable marker. This yeast strain (the parental line) can be maintained on media lacking histidine and uracil.

For the cloning of *At*IPCS1-3, *Os*IPCS1-3 into the pESC-LEU expression vector at multiple cloning site 2, primers were designed for each IPCS insert and amplified by proofreading PCR from the pRS426MET plasmid. Subsequently, infusion® cloning was carried out, followed by transformation in D-H5α *E. coli* to isolate high yields of the purified plasmid. These plasmids were then used to transform the parental yeast strain using the lithium acetate/single carrier DNA/polyethylene glycol method[23, 24].

To obtain the yeast strains containing only the pESC-LEU vector with the IPCS of interest, and lacking pRS316-AUR1, the mutant strains were grown on media containing 5-fluoroorotic acid (5-FOA). 5-FOA is an extremely useful reagent for the selection of ura$^-$ cells amid a population of ura$^+$ cells; this selection is effective in studies where loss of *URA3* is desired[25]. URA3 encodes for orotidine-5′-monophosphate decarboxylase (ODcase), an enzyme responsible for catalysing the last step in the de novo synthesis of

pyrimidines by the decarboxylation of orotidine monophosphate to generate uridine monophosphate[26]. In cells maintaining the pRS316-AUR1 URA3 plasmid, ODcase is expressed and converts 5-FOA to fluorouracil (Figure 2-6), a toxic analogue of uracil, resulting in cell death. This allows selection of yeast strain lacking this plasmid, a process termed 'plasmid shuffle' (Figure 2-7)[25].
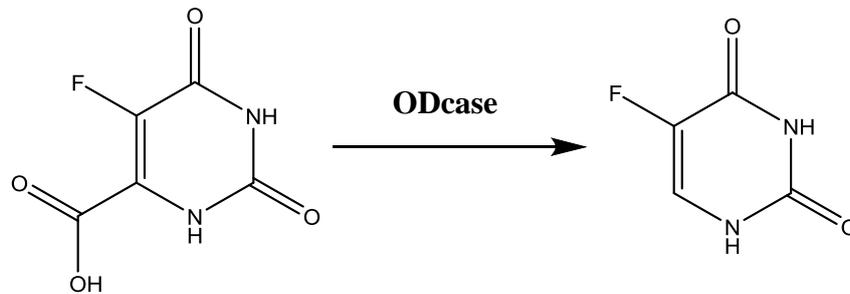
Figure 2-6: Conversion of 5-fluoroorotic acid to fluorouracil is used to select against yeast harbouring the AUR1 plasmid under the URA3 marker.

Yeast cells without either the pRS316-AUR1 URA3 plasmid or an introduced plasmid encoding a complementing enzyme, have no IPCS, and so cannot synthesise essential sphingolipids. The plant IPCSs cloned into the pESC-LEU vector are under the influence of the GAL1 promoter, therefore yeast dependent on IPCS expression from these vectors can only grow in media/agar containing galactose (SGR-H-L).
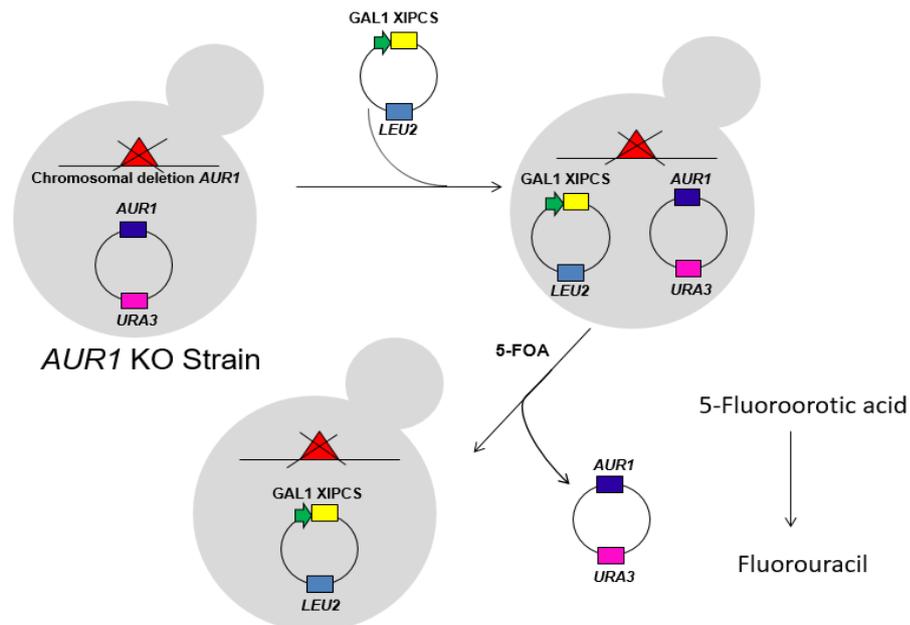
Figure 2-7: Plasmid shuffle; transformation with pESC-LEU plant IPCS plasmid and elimination of pRS426-AUR1 plasmid in yeast.

*At*IPCS1-3 and *Os*IPCS1-3 complemented yeast strains were taken through two rounds of 5FOA selection. Using PCR, yeast colonies that grew on SGR-H-L+FOA agar were verified for the presence of the pESC-LEU IPCS plasmid and the absence of the pRS316 AUR1 plasmid; these colonies were then used in subsequent experiments. It was noted that *Os*IPCS3 complemented strains had few colonies following FOA selection and required an extended period of incubation at 30°C (10 days) for colony formation.

2.3.1 Complementation Assay

All six complemented strains were grown on SGR-H-L agar and on glucose agar without histidine and leucine (SD-H-L) agar respectively. As expected, all the strains grew only on SGR-H-L due to the IPCS being under the influence of the galactose promoter (Figure 2-8). This indicated that the encoded plant enzymes complemented for the absence of the yeast AUR1p. The parental strain can grow in glucose and galactose because AUR1, which is in the pRS316 plasmid is not under the influence of the GAL1 promoter.
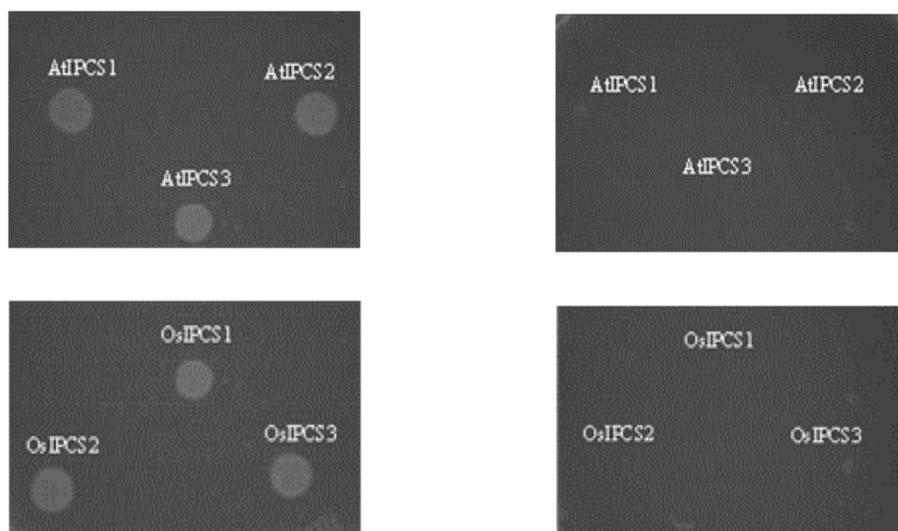
Figure 2-8: Complementation Assay - SGR-H-L plate (left) shows that all yeast strains grew but in SD-H-L plate (right) none of the strains grew because all plant IPCSs are under the influence of the GAL1 promoter.

2.3.2 Diffusion Assay

A diffusion assay is a simple and low-cost technique that has been previously used to characterise the selectivity and sensitivity of AUR1p to AbA[26, 27]. Diffusion assays involve embedding the yeast strain in the required media with 0.8% agarose. The inhibitor is then applied to the surface and diffuses through the agarose; a zone of clearance around the point of application is formed, where the yeasts are unable to grow. The zone of clearance is dictated by how active the compound is against the yeast strains.

As mentioned previously, AbA is a cyclic peptide that is a well-known inhibitor of fungal AUR1 and is toxic against various *Leishmania* species[28]. However, the concentration of AbA needed to inhibit the growth of yeast strains harbouring *Lmj*IPCS was found to be 100 µM, which is so high that the yeast strains can be considered as insensitive to the compound[27]. The sensitivity to AbA of all six yeast strains rescued with the respective plant IPCSs, *At*IPCS1-3 and *Os*IPCS1-3 were tested in a diffusion assay as was the sphingolipid biosynthesis inhibitor myriocin[29]. Myriocin inhibits the enzyme SPT thereby inhibiting sphingolipid formation and hence, membrane formation, therefore myriocin should inhibit all strains of yeast in this experiment. The sensitivity of the complemented, pdr/aur1 mutant strain to AbA and myriocin was tested, with the vehicle dimethyl sulphoxide (DMSO) as the negative control (Figure 2-9).

All plant IPCS complemented yeast strains showed a zone of clearance for myriocin (1 mM), and, interestingly, all but *Os*IPCS3 showed insensitivity to AbA at a concentration of 100 µM (Figure 2-9) implying differential selectivity of the compound for *Os*IPCS3. The same assay carried out for *Lmj*IPCS and AUR1 showed contrary to results from Denny et el. 2006, yeast strains complemented with *Lmj*IPCS being insensitive to AbA at 100 µM (Figure 2-10) which indicated differential selectivity between protozoan and

fungal IPCS. However, it is possible that insensitivity of *Lmj*IPCS complemented yeast at 100 µM might be due to a higher number of cell used in the assay. Based on the findings of Denny *et al.*, LmjIPCS and OsIPCS3 are sensitive to AbA at 100 µM which indicates that they share a structural similarity. Analysis of the amino acid sequence shows that two IPCs share a percentage identity of 25% which is quite low and alignment of the two sequences revealed no identifiably conserved region.
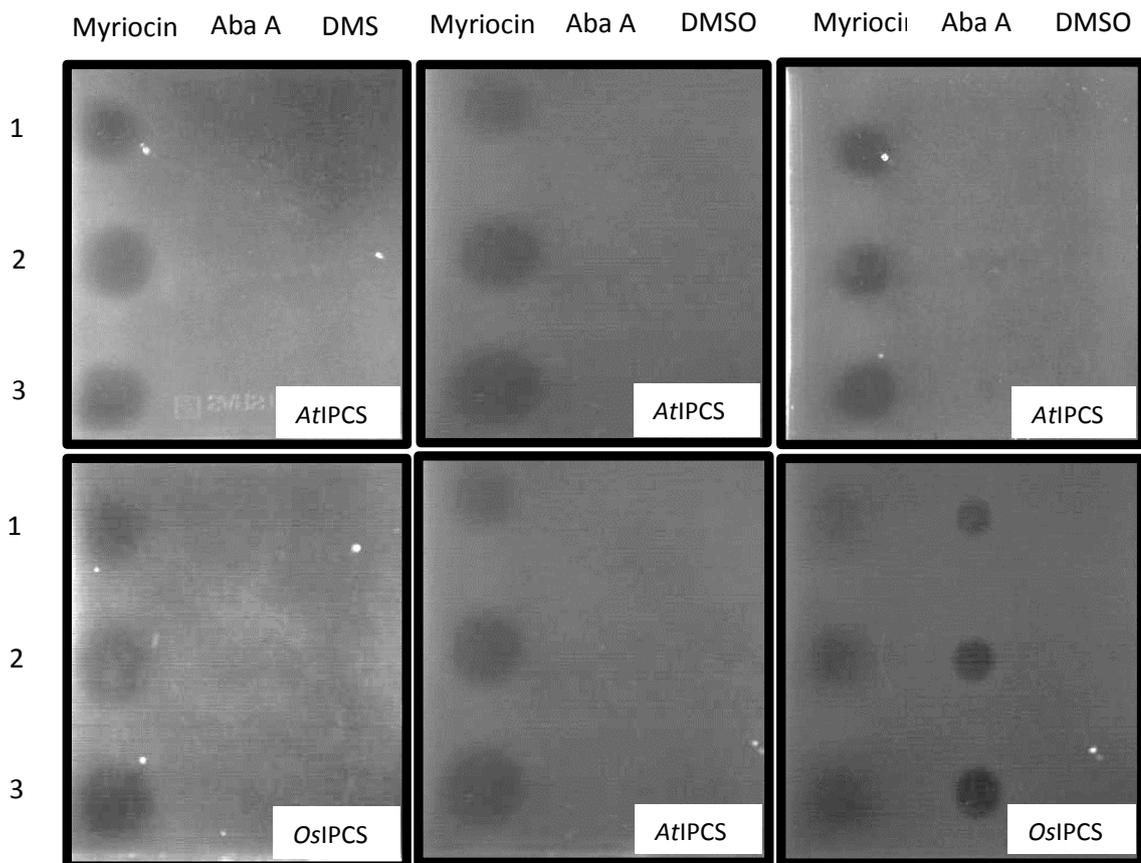
Figure 2-9: Diffusion assay plates for (top, left to right) *At*IPCS31-3 and (bottom, left to right) *Os*IPCS1-3 mutant yeast strains embedded in SGR-H-L + 0.8% agarose media. 1 mM Myriocin and 100 µM AbA in DMSO and the negative control DMSO, were pipetted in 1 µl, 2 µl and 3 µl aliquots onto the plates and left to grow at 30ºC for 5 days.

To confirm the loss of pRS426 AUR1 and the gain of pESC-LEU IPCS, specific oligonucleotide primers were designed to amplify a region of about 300bp of each plant IPCS and 500bp from AUR1 by PCR. Two yeast colony PCR reactions were carried out for each yeast strain; one with the plant IPCS and the other with the AUR1 primers. The

results demonstrated that each complemented strain contained the expected *At*IPCS or *Os*IPCS orthologue in the absence of AUR1 (Figure 2-11A and Figure 211B); confirming that the yeast strains lacked AUR1 and were expressing the respective plant IPCS isoforms.
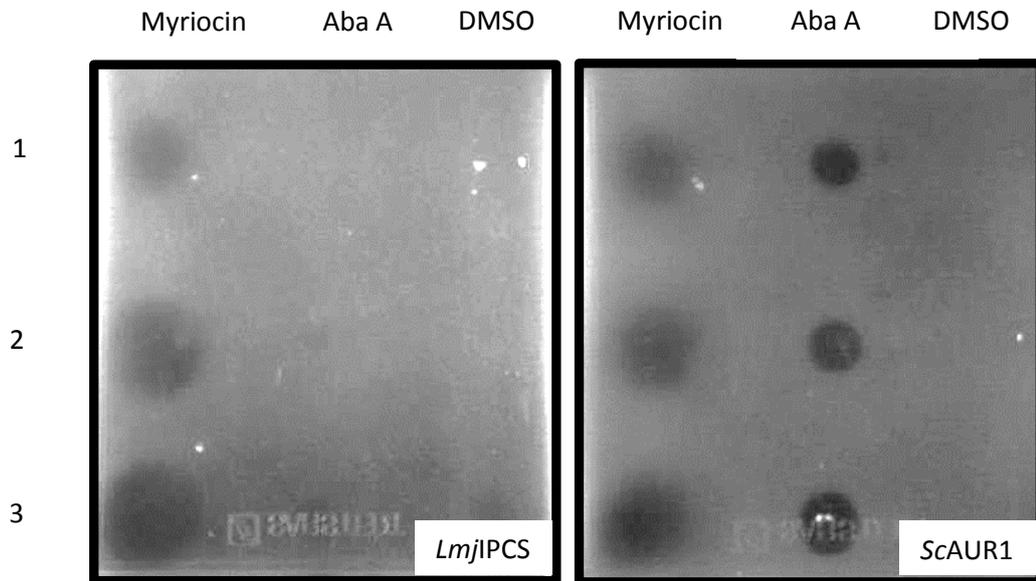
Figure 2-10: Diffusion assay plates for (left to right) *Lmj*IPCS and *Sc*AUR1 mutant yeast strains embedded in SGR-H-L + 0.8% agarose media. 1 mM Myriocin, 100 μM Ab A in DMSO, and the negative control DMSO, were pipetted in 1 μl, 2 μl and 3 μl aliquots onto the plates and left to grow at 30ºC for 5 days.

2.3.3 Verification of *At*IPCS and *Os*IPCS3 activity

Following verification of the successful complementation of the plant IPCS yeast strain, microsomal preparation was undertaken to isolate microsomes containing *At*IPCS1-3, *Os*IPCS1-3 and AUR1. To measure the activity of IPCS, the turnover of NBD-C6-ceramide into the fluorescent product NBD-C6-IPC (Figure 2-12) was quantified by fluorescence scanning following high-performance thin layer chromatography (HPTLC) fractionation. NBD-C_6-ceramide has an R_f value of 0.96 and NBD-C_6-IPC has an R_f value of 0.57[30]. This results to a clear resolution between the substrate and the product bands on the TLC plate, allowing for the quantification of IPC turnover.

The activity of *At*IPCS1-3 and *Os*IPCS1-3 were tested in the presence and absence of the known fungal IPCS inhibitor AbA at a concentration of 5 µM. The result demonstrated that *At*IPCS1-3 and *Os*IPCS1-2 were insensitive to AbA, whilst *Os*IPCS3 was sensitive to the inhibitor (Figure 2-13 and 2-14) in agreement with the diffusion assay data (Figure 2-9). The results for *At*IPCS1-3 and AUR1 are as expected; work previously done by Mina *et al.* had shown that the three *A. thaliana* IPCS isoforms are insensitive to Aba[1] and earlier work by Heidler *et al*, as mention in chapter 1 demonstrated that AUR1 is sensitive to the inhibitor[15].
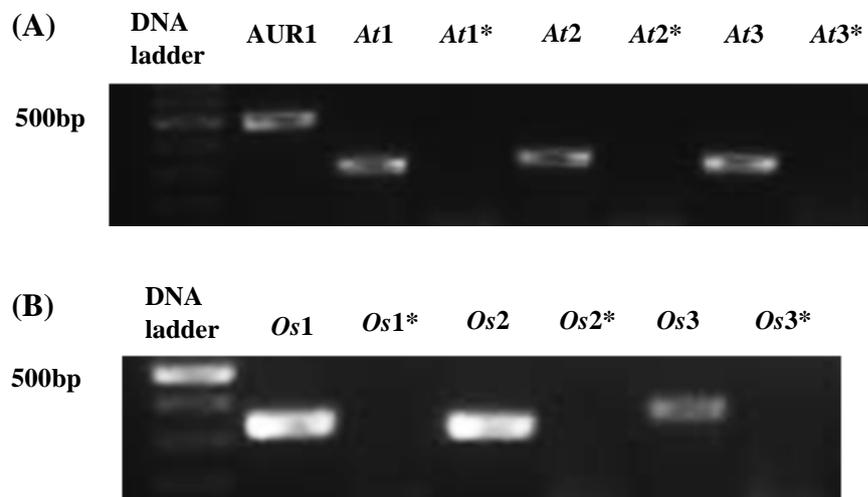
Figure 2-11: Transformed yeast colony PCR for (A) AUR1 and *At*IPCS1-3 (B) *Os*IPCS1-3 with specific primers for AUR1 (*) and each insert.
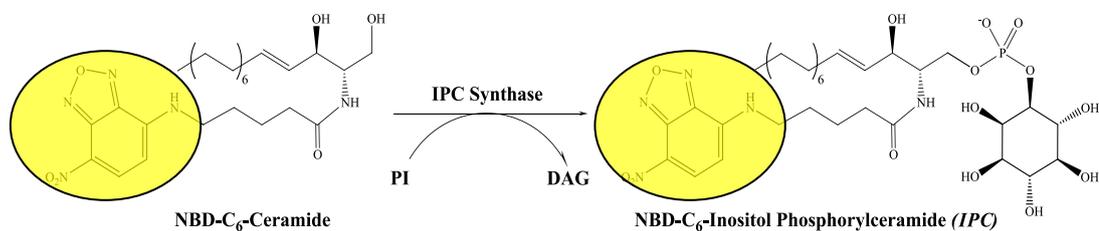
Figure 2-12: Conversion of NBD-C_6-ceramide to NBD-C_6-IPC.

To facilitate further analysis, the enzyme activity for each IPCS of interest was deduced using an NBD-C_6-ceramide standard curve. It was assumed that the fluorescence of the

substrate NBD-ceramide is equal to the fluorescence of NBD-IPC, which allowed for the quantification of IPC turnover in a reaction and was subsequently used to deduce the concentration of active IPCS in the microsome mixture. Active IPCS concentration was quantified in 'Unit/µl', where one Unit is defined as the amount of enzyme required to catalyse the formation of 1 picomol of product in 1 minute[31].
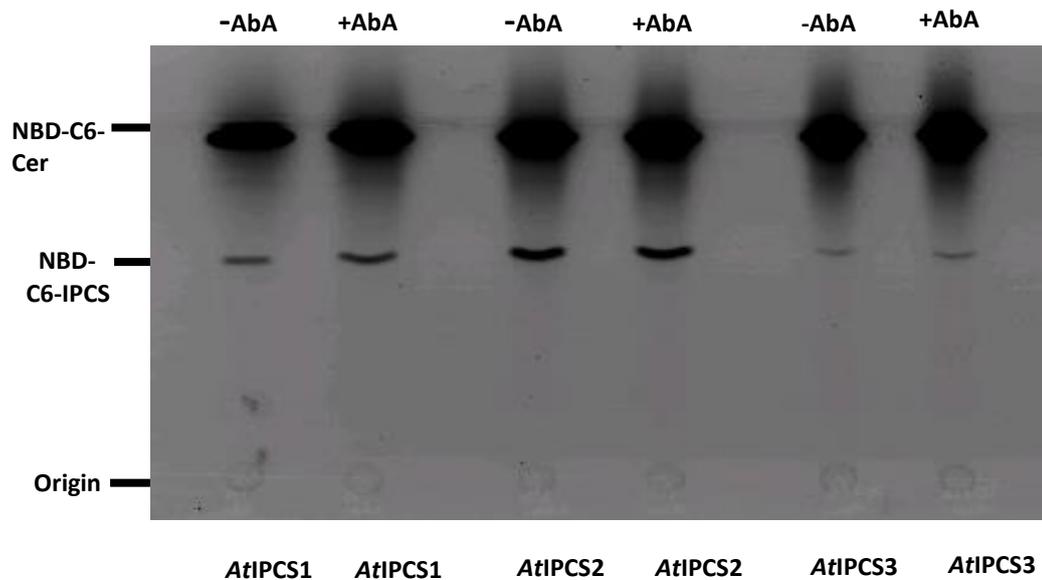
Figure 2-13: HPTLC assay *At*IPCS1-3 in the absence and presence of AbA
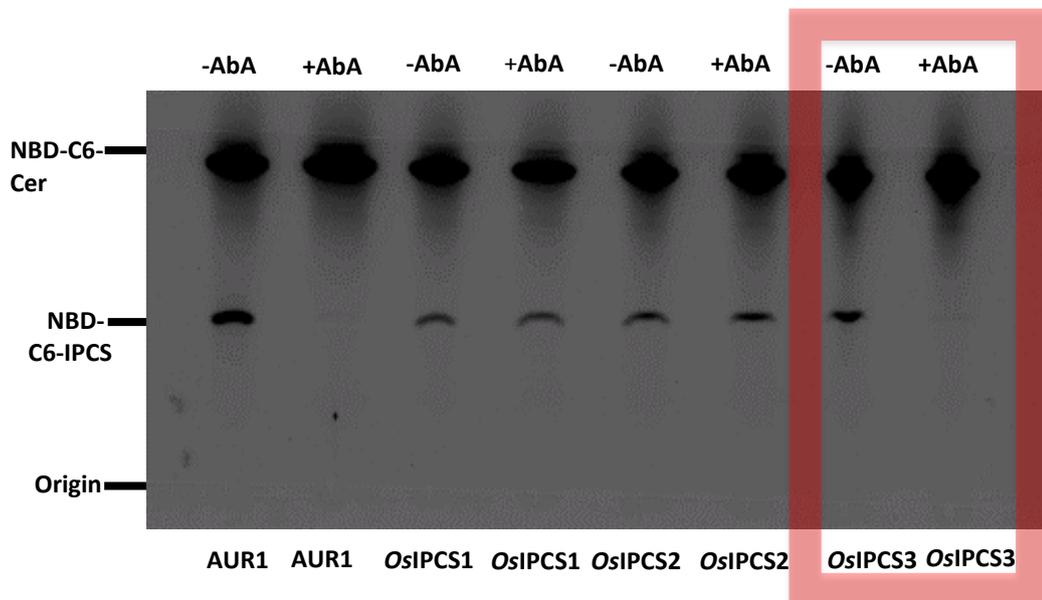
Figure 2-14: HPTLC assay *Os*IPCS1-3 in the absence and presence of AbA with the inhibition of *Os*IPCS3 highlighted in the red box showing the lack of IPC turnover in the presence of the compound.

To establish the IC$_{50}$ of AbA against AUR1 and *Os*IPCS3, it was necessary to quantify

the enzyme activity; the activity of the *Os*IPCS3 and AUR1 samples were deduced to be 5.32 Unit/µl and 14.68 Unit/µl respectively (Table 2-2). By quantifying the enzyme activity, it was possible to ensure that the enzyme concentration of *Os*IPCS3 and AUR1 used in the assay was the same.

Table 2-2: IPCS activity for AUR1, *At*IPCS1-3 and *Os*IPCS1-3 quantified

Microsomes	Fluorescence at 473nm	pmol	pmol/min (Unit)	Unit/ul
AUR1	96150	11000	183.3	14.68
*At*IPCS1	18123	1400	23.3	1.88
*At*IPCS2	38985	3200	53.3	4.28
*At*IPCS3	6902	480	8.0	0.64
*Os*IPCS1	16934	1240	20.7	1.64
*Os*IPCS2	16057	1200	20.0	1.6
*Os*IPCS3	46541	4000	66.7	5.32

Equivalent active enzyme concentration of AUR1 and *Os*IPCS3 was assayed in the presence of AbA at concentrations ranging from 0.05-5 nM. The IC_{50} of AUR1 was found to be 0.6 nM compared to the literature value of 0.89 nM[26] and the IC_{50} of *Os*IPCS3 was found to be 2.62 nM. The IC_{50} values obtained show that there is a more than 2-fold difference in sensitivity to AbA between AUR1 and *Os*IPCS3; AUR1 is much more sensitive to Aureobasidin A than *Os*IPCS3 (Figure 2-15).

To ensure that all complemented strains were still fully dependent on the specific plant IPCSs, another colony PCR experiment was carried out five months later and it was found that all yeast strains had retained their specific IPCS apart from yeast strains complemented with *Os*IPCS3 (Figure 2-16) which showed the presence of the fungal IPCS signifying reversion to dependence on the fungal orthologue. There is currently no known mechanism for how this type of gene reversion occurs but work done by Schiestel *et al.*[32] show that reversion of a deleted gene results from the excision of plasmid

harbouring the dominant marker (usually HIS3) which replaces the target gene and integrated into the genome.

(A)
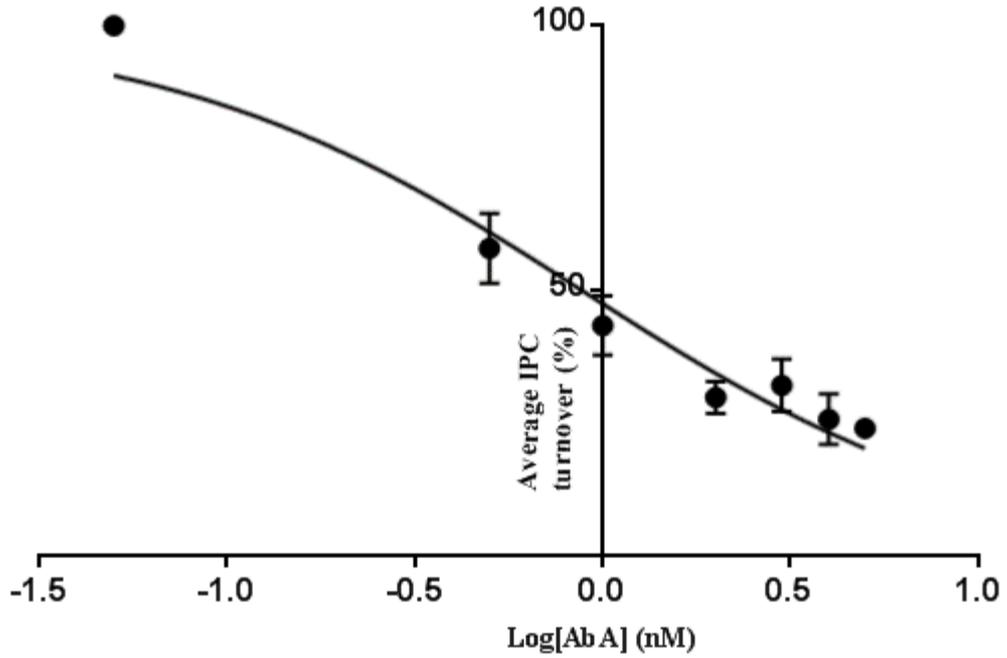
(B)
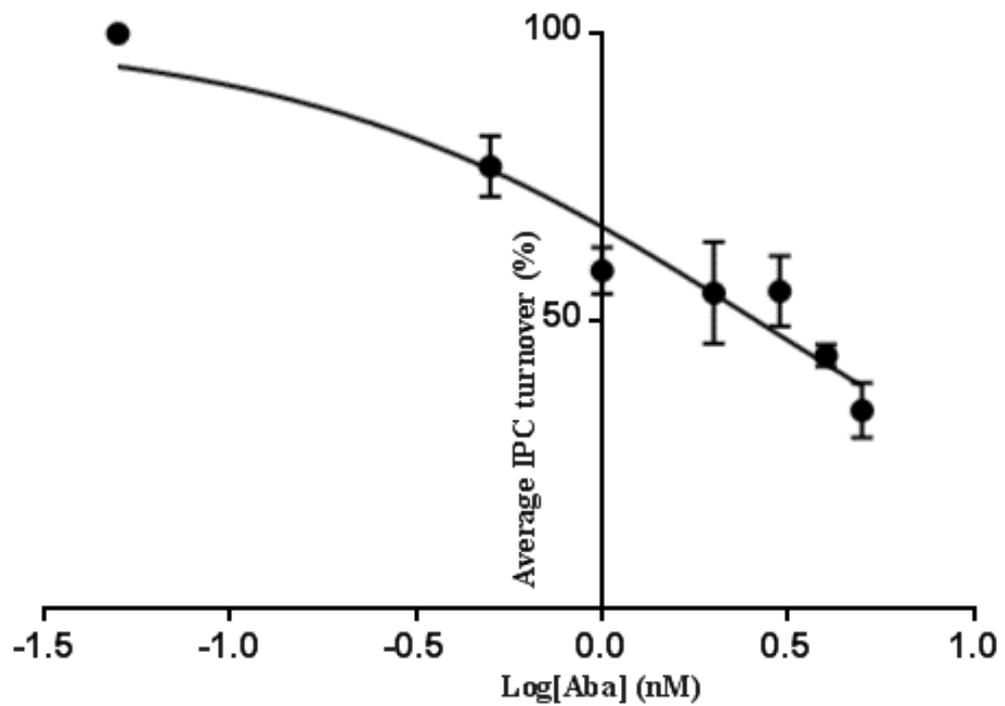
Figure 2-15: IC_{50} graph showing the percentage turnover of IPC upon addition of different concentration of AbA for (A) AUR1 (B) *Os*IPCS3.

For successful integration into the genome, the plasmid usually has ends flanked with areas homologous to the target gene; when reversion occurs, this is because these ends are retained upon excision of the plasmid, which is repaired by sister chromatid conversion, resulting in the reconstruction of the wildtype gene that is fully functional. In this case, although *Os*IPCS3 is expressed in the yeast strains, the deleted AUR1 gene in the genome has undergone reversion, resulting in the expression of the fungal IPCS.
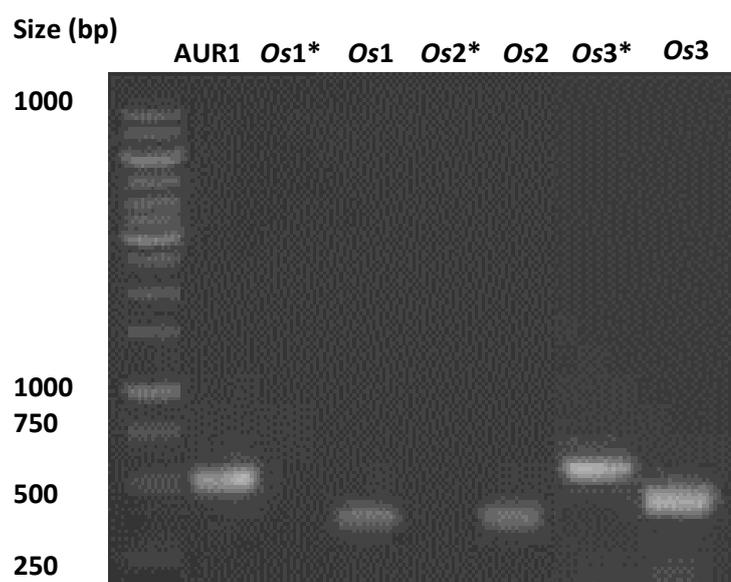
Figure 2-16: Agarose gel (0.8%) showing bands amplified for *Os*IPCS1-3 with respective specific primers and AUR1 primers as control (* denotes PCR carried out with AUR1 primers); *Os*IPCS3 shows a product for *Os*IPCS3 and AUR1 primers.

2.4 Assay development

Following the verification of *At*IPCS1-3 and *Os*IPCS1-2 yeast complemented strains, an assay to test compounds against the different yeast strains was developed and formatted in a 96 well plate. To set the assay parameters for HTS, the growth of all yeast strains with the IPCS of interest was monitored by optical density (OD_{600}) at different starting concentrations, over 72 hours at 24-hour intervals. From the results, it was clear that yeast growth plateaus after 48 hours. At the highest starting concentration in some cases, there

was a decrease in the OD_{600}, presumably, due to cell death as nutrients become a limiting factor.
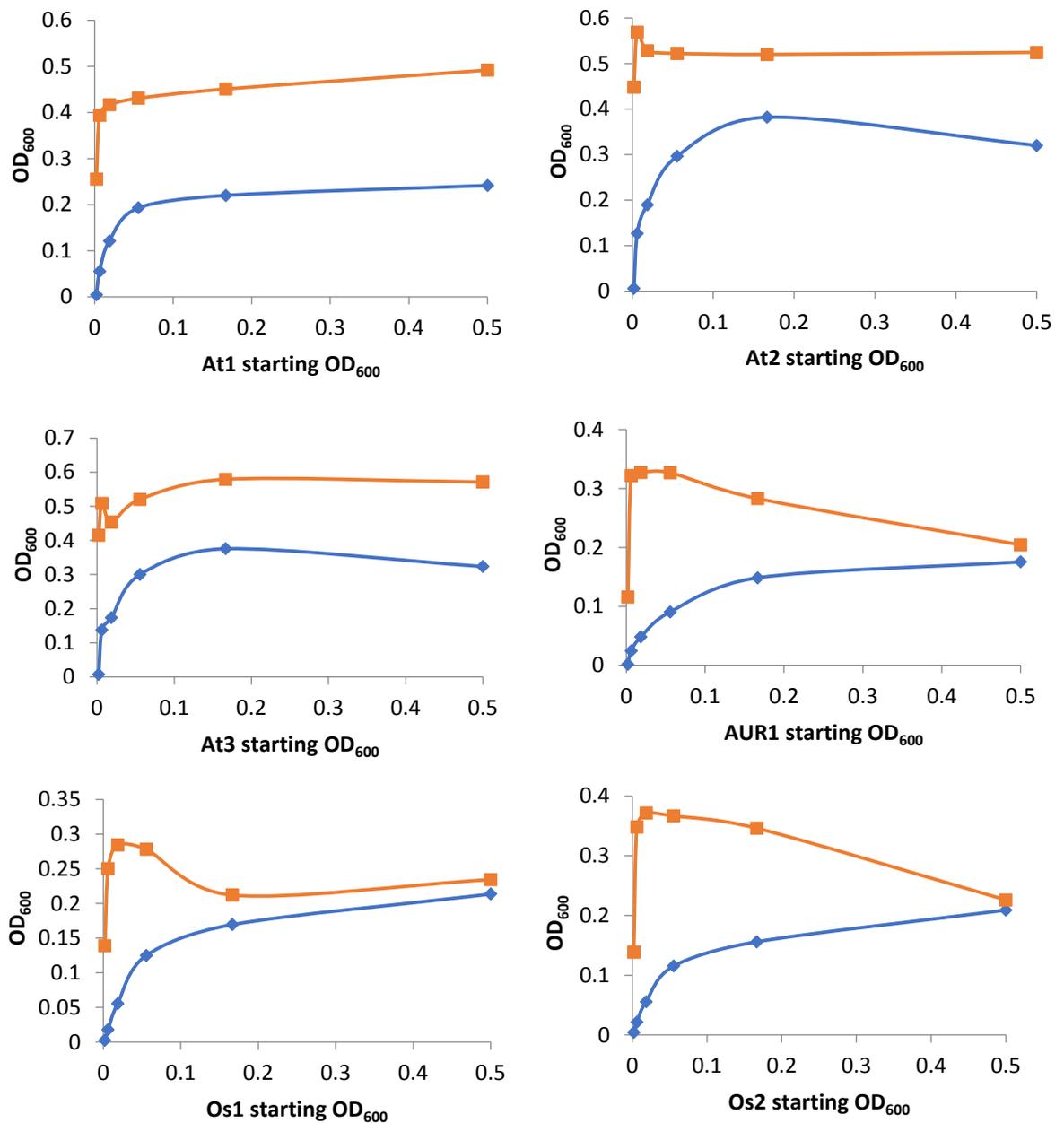
Figure 2-17: Graphs showing the relationship between OD_{600} starting cell concentrations and AUR1, *At*IPCS1-3, *Os*IPCS1 and 2 cell concentrations measured after 24 (blue line) and 48 hours (orange line) growth respectively (starting OD_{600} from highest to lowest corresponds to dilutions of 0, 1:3, 1:6,1:9, 1:12 and 1:16 respectively).

For an assay (e.g. HTS) the conditions must ensure a direct linear correlation between OD_{600} read out and cell number over a defined time period. This creates well defined

assay parameters whereby the efficacy of a compound can be directly measured by monitoring yeast growth. Similar assay parameters have been applied for the testing of antimicrobial peptides against *Leishmania*[33]. As the results show, a consistent linear correlation was observed at 24 hours for 1:9 (OD_{600}= 0.018) and 1:12 (OD_{600}=0.006) dilutions (Figure 2-17). To increase the signal to noise ratio, the assay starting cell concentration was set at an OD_{600} = 0.018 corresponding to a 1:9 dilution.

HTS usually involve the screening of compounds once or twice for the identification of hit compounds and requires that the assay has high sensitivity and accuracy. Due to the perturbations introduced in the assay resulting from human and instrumental error, assay measurements can contain high degrees of variation which need to be considered. For the identification of 'real' hits, it is necessary that variations are low which can be measured by standard deviation from the mean and by the coefficient of variation[33].

The Z'-factor[34] is a quantitative assessment of the quality of the assay that can be calculated using the standard deviation and the mean of the positive and negative control (Equation 1.1). An ideal assay would have a Z' factor value equal to 1, whilst an 'excellent assay' has a Z' factor value < 1 and ≥ 0.5.

Z-factor = 1- (3(σp + σn)/ (|μp - μn))   **[Equation 1.1]**

σ is the standard deviation and μ is the mean of the positive (p) and negative (n) control

To test the suitability of the assay in identifying hit compounds, Z' factor values were calculated for *At*IPCS1-3 and AUR1 (Table 2-3). DMSO was used as the negative control and cycloheximide (1 mM), a known inhibitor of cytoplasmic protein synthesis in yeast and other eukaryotes[35] was used as the positive control in the assay. The assay was set up in 96 well plates, incubated with shaking over 24 hours and yeast cell viability quantified

by measuring cell density at OD_{600}.

A trial of the assay's reproducibility and hit identification was carried out by testing 15 compounds against *At*IPCS1 complemented yeast strain. The 15 compounds included Aba, clemastine analogues[36], khafrefungin analogues and the known sphingolipid biosynthesis inhibitor, myriocin. Results obtained from the assay were reproducible and identified myriocin, khafrefungin and the khafrefungin analogues: W2345 and ND8 as hits (Figure 2-18, structures of compounds tested are in Appendix D). Interestingly, the khafrefungin analogue, W2345, has been shown to have no activity against *S. cerevisiae* (work done by Kentaro Hanada, unpublished), indicating differential selectivity for the plant IPCS over fungal IPCS. However, the khafrefungin analogues are not attractive as inhibitors of IPCS because these compounds are not easily synthesised and required several synthetic steps to produce the final product. In this study, the aim is to find small molecule inhibitors that can be easily modified for use as herbicides, so the selectivity of the khafrefungin analogue, W2345, although interesting was not pursued further.

Table 2-3: Calculated Z' factor value to validate assay

	DMSO OD_{600} average	Cycloheximide OD_{600} average	Z' factor
*At*IPCS1	1.54	0.07	0.9
*At*IPCS2	1.47	0.15	0.7
*At*IPCS3	1.46	0.14	0.6
AUR1	1.26	0.21	0.6

Having established that the assay was suitable for the identification of hit compounds, yeast strains were grown to log phase (OD_{600} = 0.5-0.7) and frozen to screen the 11,440

compound library for inhibitors of plant IPCSs which was carried out at Bayer Crop Sciences, Frankfurt Germany.
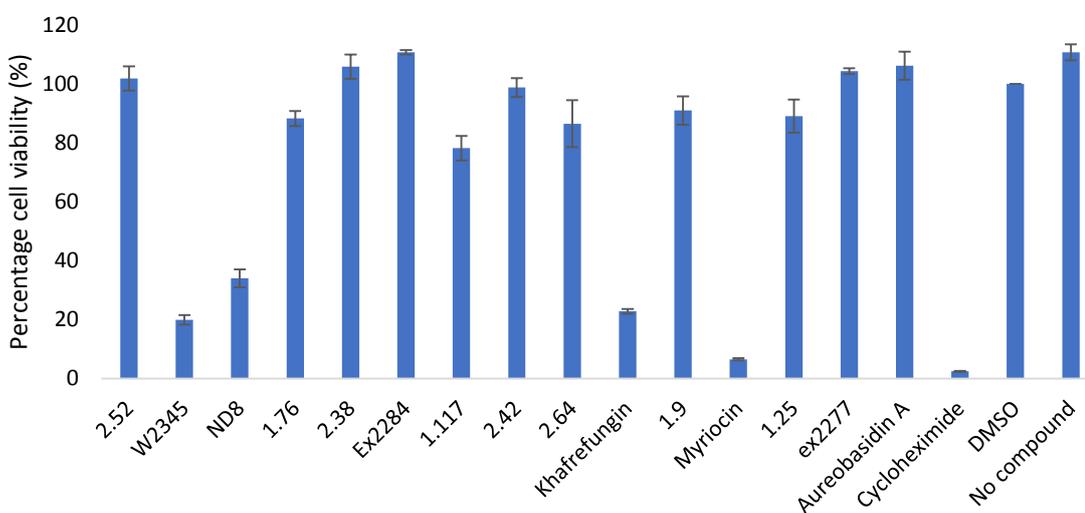
Figure 2-18: Percentage cell viability of *At*IPCS1 complemented yeast cells upon addition of test compounds and incubation for 24 hours; each compound was tested in triplicate at 10 µM.

2.5 Conclusion

Cloning of *Os*IPCS1-3 and *At*IPCS1-3 was achieved and complementation of plant IPCS yeast strains was successful for all but *Os*IPCS3. It was shown that plant IPCS unlike fungal IPCS, were insensitive to Aba. The yeast strains were formatted into an assay which requires quantifying the growth of the yeast strains by measuring OD_{600} after a 24 hours incubation period. The assay conditions were validated for the screening of compounds by quantifying the Z' factor value of an assay with DMSO as negative control and cycloheximide as positive control. Initial screening of a compound library consisting of clemastine analogues, and known inhibitors of sphingolipid biosynthesis, were tested on *At*IPCS1 complemented yeast strains. Myriocin and khafrefungin were identified as hits, and more importantly the assay showed that the *At*IPCS1 complemented yeast strains were insensitive to Aba, which supports the biochemical assay results of plant IPCSs being insensitive to the compound. Other identified hits were the khafrefungin analogues, W2345 and ND8. The activity of W2345 against *At*IPCS1 complemented yeast indicates

differential selectivity for plant IPCS over fungal IPCS, which was a promising indication of the ability of the assay to identify hits which show differential selectivity. Following the validation of assay conditions, the screening of a library of bioactive compounds against the plant IPCSs and AUR1 was carried out; this is presented in the next chapter.

2.6 References

1.   Mina, J. G.; Okada, Y.; Wansadhipathi-Kannangara, N. K.; Pratt, S.; Shams-Eldin, H.; Schwarz, R. T.; Steel, P. G.; Fawcett, T.; Denny, P. W., Y Functional analyses of differentially expressed isoforms of the Arabidopsis inositol phosphorylceramide synthase. *Plant Molecular Biology* **2010,** *73* (4-5), 399-407.
2.   Mina, J.; Okada, Y.; Wansadhipathi-Kannangara, N.; Pratt, S.; Shams-Eldin, H.; Schwarz, R.; Steel, P.; Fawcett, T.; Denny, P., Functional analyses of differentially expressed isoforms of the Arabidopsis inositol phosphorylceramide synthase. *Plant molecular biology* **2010,** *73* (4-5), 399-407.
3.   Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology* **2001,** *305* (3), 567-580.
4.   Yachdav, G.; Kloppmann, E.; Kajan, L.; Hecht, M.; Goldberg, T.; Hamp, T.; Hönigschmid, P.; Schafferhans, A.; Roos, M.; Bernhofer, M., PredictProtein—an open resource for online prediction of protein structural and functional features. *Nucleic acids research* **2014,** *42* (W1), W337-W343.
5.   Levine, T. P.; Wiggins, C. A.; Munro, S., Inositol phosphorylceramide synthase is located in the Golgi apparatus of Saccharomyces cerevisiae. *Molecular Biology of the Cell* **2000,** *11* (7), 2267-2281.
6.   Jones, D. T., Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics* **2007,** *23* (5), 538-544.
7.   Sutterwala, S. S.; Hsu, F. F.; Sevova, E. S.; Schwartz, K. J.; Zhang, K.; Key, P.; Turk, J.; Beverley, S. M.; Bangs, J. D., Developmentally regulated sphingolipid synthesis in African trypanosomes. *Molecular microbiology* **2008,** *70* (2), 281-296.
8.   Lai, Z.; Vinod, K.; Zheng, Z.; Fan, B.; Chen, Z., Roles of Arabidopsis WRKY3 and WRKY4 transcription factors in plant responses to pathogens. *BMC plant biology* **2008,** *8* (1), 68.
9.   Pandey, S. P.; Roccaro, M.; Schön, M.; Logemann, E.; Somssich, I. E., Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *The Plant Journal* **2010,** *64* (6), 912-923.
10.   Lebel, E.; Heifetz, P.; Thorne, L.; Uknes, S.; Ryals, J.; Ward, E., Functional analysis of regulatory sequences controllingPR- 1 gene expression in Arabidopsis. *The Plant Journal* **1998,** *16* (2), 223-233.
11.   Strompen, G.; Grüner, R.; Pfitzner, U. M., An as-1-like motif controls the level of expression of the gene for the pathogenesis-related protein 1a from tobacco. *Plant molecular biology* **1998,** *37* (5), 871-883.
12.   Zhou, J.-M.; Trifa, Y.; Silva, H.; Pontier, D.; Lam, E.; Shah, J.; Klessig, D. F., NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. *Molecular Plant-Microbe Interactions* **2000,** *13* (2), 191-202.

13.	Johnson, C.; Boden, E.; Arias, J., Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. *The Plant Cell* **2003,** *15* (8), 1846-1858.

14.	Cao, H.; Glazebrook, J.; Clarke, J. D.; Volko, S.; Dong, X., The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* **1997,** *88* (1), 57-63.

15.	Heidler, S. A.; Radding, J. A., The AUR1 gene in Saccharomyces cerevisiae encodes dominant resistance to the antifungal agent aureobasidin A (LY295337). *Antimicrobial agents and chemotherapy* **1995,** *39* (12), 2765-2769.

16.	Wang, X. H.; Guo, X. J.; Li, H. Y.; Gou, P., Characteristics of inositol phosphorylceramide synthase and effects of aureobasidin A on growth and pathogenicity of Botrytis cinerea. *Journal of General and Applied Microbiology* **2015,** *61* (4), 108-116.

17.	Hashida-Okado, T.; Yasumoto, R.; Endo, M.; Takesako, K.; Kato, I., Isolation and characterization of the aureobasidin A-resistant gene, aur1R, on Schizosaccharomyces pombe: roles of Aur1p+ in cell morphogenesis. *Current genetics* **1998,** *33* (1), 38-45.

18.	Hashida-Okado, T.; Ogawa, A.; Kato, I.; Takesako, K., Transformation system for prototrophic industrial yeasts using the AUR1 gene as a dominant selection marker. *FEBS letters* **1998,** *425* (1), 117-122.

19.	Kuroda, M.; Hashida-Okado, T.; Yasumoto, R.; Gomi, K.; Kato, I.; Takesako, K., An aureobasidin A resistance gene isolated from Aspergillus is a homolog of yeast AUR1, a gene responsible for inositol phosphorylceramide (IPC) synthase activity. *Molecular and General Genetics MGG* **1999,** *261* (2), 290-296.

20.	Xing, B.; Bachstetter, A. D.; Van Eldik, L. J., Inhibition of neuronal p38α, but not p38β MAPK, provides neuroprotection against three different neurotoxic insults. *Journal of Molecular Neuroscience* **2015,** *55* (2), 509-518.

21.	Balzi, E.; Chen, W.; Ulaszewski, S.; Capieaux, E.; Goffeau, A., The multidrug resistance gene PDR1 from Saccharomyces cerevisiae. *Journal of Biological Chemistry* **1987,** *262* (35), 16871-16879.

22.	Georgiev, V. S., Membrane transporters and antifungal drug resistance. *Current drug targets* **2000,** *1* (3), 261-284.

23.	Schiestl, R. H.; Gietz, R. D., High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Current genetics* **1989,** *16* (5), 339-346.

24.	Hill, J.; Donald, K. I. G.; Griffiths, D. E., DMSO-enhanced whole cell yeast transformation. *Nucleic acids research* **1991,** *19* (20), 5791-5791.

25.	Boeke, J. D.; Trueheart, J.; Natsoulis, G.; Fink, G. R., [10] 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods in enzymology* **1987,** *154*, 164-175.

26.	Nagiec, M. M.; Nagiec, E. E.; Baltisberger, J. A.; Wells, G. B.; Lester, R. L.; Dickson, R. C., Sphingolipid synthesis as a target for antifungal drugs complementation of the inositol phosphorylceramide synthase defect in a mutant strain of Saccharomyces cerevisiae by the AUR1 gene. *Journal of Biological Chemistry* **1997,** *272* (15), 9809-9817.

27.	Denny, P. W.; Shams-Eldin, H.; Price, H. P.; Smith, D. F.; Schwarz, R. T., The protozoan inositol phosphorylceramide synthase a novel drug target that defines a new class of sphingolipid synthase. *Journal of Biological Chemistry* **2006,** *281* (38), 28200-28209.

28.	Tanaka, A. K.; Valero, V. B.; Takahashi, H. K.; Straus, A. H., Inhibition of Leishmania (Leishmania) amazonensis growth and infectivity by aureobasidin A. *Journal of Antimicrobial Chemotherapy* **2007,** *59* (3), 487-492.

29.   Hanada, K., Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2003,** *1632* (1), 16-30.

30.   Mina, J. G.; Mosely, J. A.; Ali, H. Z.; Shams-Eldin, H.; Schwarz, R. T.; Steel, P. G.; Denny, P. W., A plate-based assay system for analyses and screening of the Leishmania major inositol phosphorylceramide synthase. *International Journal of Biochemistry & Cell Biology* **2010,** *42* (9), 1553-1561.

31.   Loening, K. L., ACS Committee on Nomenclature, Annual Report, 1979. *Journal of Chemical Information and Computer Sciences* **1980,** *20* (2), 61-64.

32.   Schiestl, R. H.; Igarashi, S.; Hastings, P., Analysis of the mechanism for reversion of a disrupted gene. *Genetics* **1988,** *119* (2), 237-247.

33.   Chadbourne, F. L.; Raleigh, C.; Ali, H. Z.; Denny, P. W.; Cobb, S. L., Studies on the antileishmanial properties of the antimicrobial peptides temporin A, B and 1Sa. *Journal of Peptide Science* **2011,** *17* (11), 751-755.

34.   Zhang, J.-H.; Chung, T. D.; Oldenburg, K. R., A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of biomolecular screening* **1999,** *4* (2), 67-73.

35.   Boulton, C.; Quain, D., *Brewing yeast and fermentation.* John Wiley & Sons: 2008.

36.   Brown, C. Inhibitors of Leishmania major Inositol Phosphorylceramide Synthase: New Therapies for Leishmaniasis. Durham University, 2016.

CHAPTER 3

SCREENING FOR INHIBITORS OF PLANT IPCS

3.1 Chapter synopsis

In this chapter, the screening strategies utilized to identify herbicidal leads or chemical probes are outlined, and results from the screening of a compound library of 11,440 bioactive compounds are presented. Following the identification of hits in the primary screen, assay conditions for the secondary biochemical screen was carrried out. Variants of phosphatidyl inositol were tested to see which gave maximal activity in terms of plant IPCS turnover. In addition various substrate concentration were tested to ensure the identification of hits was not hindered by saturation of the enzyme with substrate. From the secondary screen 15 compounds were found to show differential selectivity for plant IPCS over AUR1. In addition, compounds were identified that exhibited differential selectivity for *At*IPCS2 over *Os*IPCS2. The top two compounds, showing the most activity against plants IPCS were tested *in vivo* and found to have herbicidal effects that point to the biological pathway targeted, affecting root development and involving PCD.

3.2 Approaches to screening for inhibitors

In agrochemical discovery, two approaches are employed: a phenotypic chemical screen or a target-based chemical screen. In a phenotypic based screen also termed a forward chemical screen, chemicals are tested on plants in a microplate format, and bioactive compounds discovered by identifying an alteration in phenotype upon application of the compound[2]. Following this, compounds are selected and tested against plausible targets (phenotype can give an idea of mode of action) in an assay or a biochemical purification strategy is employed to identify targets. The second approach which is a reverse chemical screen (target-based chemical screen), involves the testing of a compound library against a specific target, with the identified hits then tested *in vivo* to see if the compound remains active which is deduced by a phenotypic alteration in planta[3].

The forward chemical screening approach has been utilized in numerous studies for the identification of bioactive compounds. For example, the screening of a library of compounds against *A. thaliana* seedlings, identified hits that either enhanced or inhibited gravitropic response, ascertained by the presence or absence of root reorientation. From the hits identified, compounds were discovered that also affected the endomembrane system, discovered by observation of aberrant vacuoles displaying a vesiculated morphology[4]. This phenotypic screening strategy to identify inhibitors or enhancers of plant development has also been used to identify auxin analogues that enhance hypocotyl growth[5] and the compound hyperphyllin which enhances leaf formation and shoot meristem enlargement[6]. Chung-lin et al. identified compounds that interfered with *A. thaliana* response to ethylene by screening for the suppression of triple response phenotypic read out in response to ethylene[7]. Hits were subsequently validated as inhibitors of 1-aminocyclopropane-1-carboxylic acid synthase. The discovery of a class of sulfanilamide compounds able to protect wheat crops from the fungal pathogen,

Fusarium graminearum, was also facilitated by an *in vivo* screen for the attenuation of dark lesion-like spots resulting from fungal infection[8].

Following the identification of a compound that has a particular effect of interest on the plant, the compounds are usually then either tested in cellular or enzymatic assays against a series of possible targets in order to facilitate the identification of the compound's mode of action. The identification of small molecule inhibitors can be coupled with omics tools, by utilizing reverse genetics and next generation sequencing to identify changes in the transcriptome or metabolome which may have a corresponding phenotype; this approach to the identification of bioactive compounds and their targets is termed chemical genomics[9].

In this study, a combination of both phenotypic and target-based chemical screening has been utilised for the identification of IPCS inhibitors. A compound library of 11,440 compounds, which have exhibited herbicidal or fungicidal activity in *in vivo* screens were tested against plant IPCS.

3.3 Hypothetical effects of an IPCS inhibitor

To understand the possible impacts of IPCS inhibition, a review of the ramifications of sphingolipid biosynthesis dysfunction is necessary. Various studies have linked the sphingolipid biosynthetic pathway to programmed cell death (PCD) in plants. Initial work carried out by Saucedo- García and colleagues showed that inhibition of phytoceramide synthase in plants treated with fumosin B results in PCD phenotypes at a cellular level, including interiorized and rounded chloroplasts, the disintegration of thylakoids vesicles and DNA fragmentation[1]. This cell death phenotype was found to be attenuated by pre-treatment of plants with myriocin which targets serine palmitoyl transferase (SPT); the inhibition of SPT results in a decrease in 3-ketosphinganine which affects the synthesis

of other compounds downstream in the biosynthetic pathway. This results in a decrease in phytosphingosine, a substrate for phytoceramide synthase, which is inhibited by fumosin B (Figure 3-2). This indicated that the accumulation of the dihydroxylated long chain bases in the form of phytosphingosine is responsible for PCD in plants.
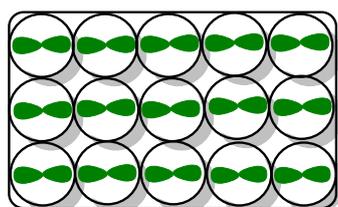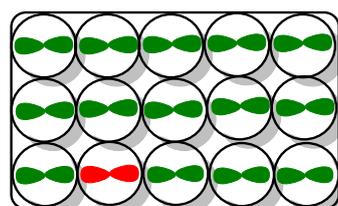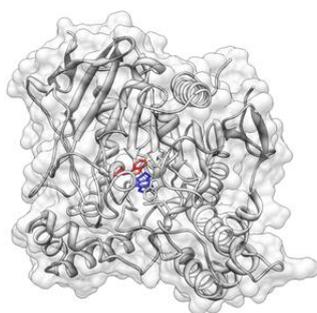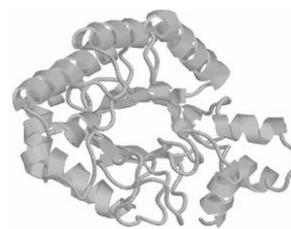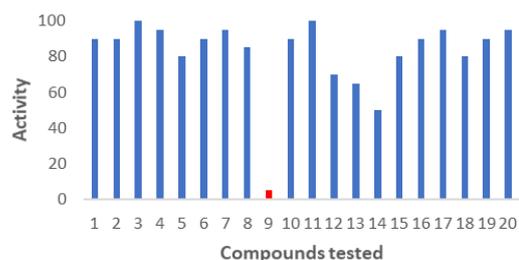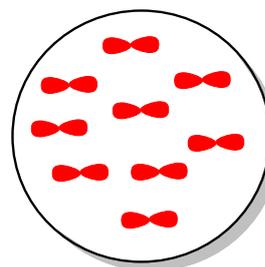
Figure 3-2: Schematic showing the two approaches for hit identification (A) Phenotypic chemical screen which involves *in vivo* testing to identify bioactive compounds which can then be screened against a number of targets to find the compound's target (B) Target-based screening which involves testing a compound library against a specific target to identify compounds that show activity against the target; the identified compound is subsequently tested in vivo to see if it results in any phenotypic alterations in planta.

This finding was corroborated was corroborated by Shi *et al.* who showed that treatment of *A. thaliana* with the sphingoid bases, dihydrosphingosine, phytosphingosine and sphingosine, resulted in cell death, which was preceded by the generation of reactive oxygen intermediates that are linked to PCD in response in abiotic stress[10].

Gene silencing of the second subunit of SPT (LCB2) in *Nicotiana benthamiana was shown to result in an approximate* 50% reduction in LCB2 transcript, *and induced cell death,* with a decrease in levels of phytosphingosine and an increase in the levels of dihydro-sphingosine and sphingosine, which were also associated with elevated SA levels and constitutive expression of the gene pathogenesis-related protein 1 (*PR1*) expression[11], both markers of PCD. The elevation in dihydro-sphingosine levels upon deletion of the first SPT subunit (LCB1) is corroborated by work done by Chen et al.[12] in *Arabidopsis*, and indicates control of the levels of the proapoptotic compound, phytoceramide, by downregulation of LCB C4 hydroxylase resulting in the build-up of dihydroshpingosine, and the subsequence reduction in phytosphingosine and phytoceramide generated; both are bioactive compounds known to trigger PCD. The elevation of sphingosine levels is also as a result of keeping phytoceramide levels down, by breaking down phytoceramide to phytosphingosine, which is achieved by ceramidase.

Disruption of phytoceramide synthesis in *Arabidopsis* by mutation of one of phytoceramide synthase genes, *LOH1*, results in spontaneous cell death, concomitant with a 160-fold increase in *PR1* transcription[13]. Similarly, RNAi suppression of LCB-C4 hydroxylase resulted in the formation of necrotic lesions and constitutive PR1 expression[14]. Mutation of ceramide kinase, ACD5, in *Arabidopsis* resulted in the accumulation of phytoceramide, with constitutive *PR1* expression, the accumulation of free SA and conjugate SA, and necrotic lesions[15]; markers associated with PCD in plants.
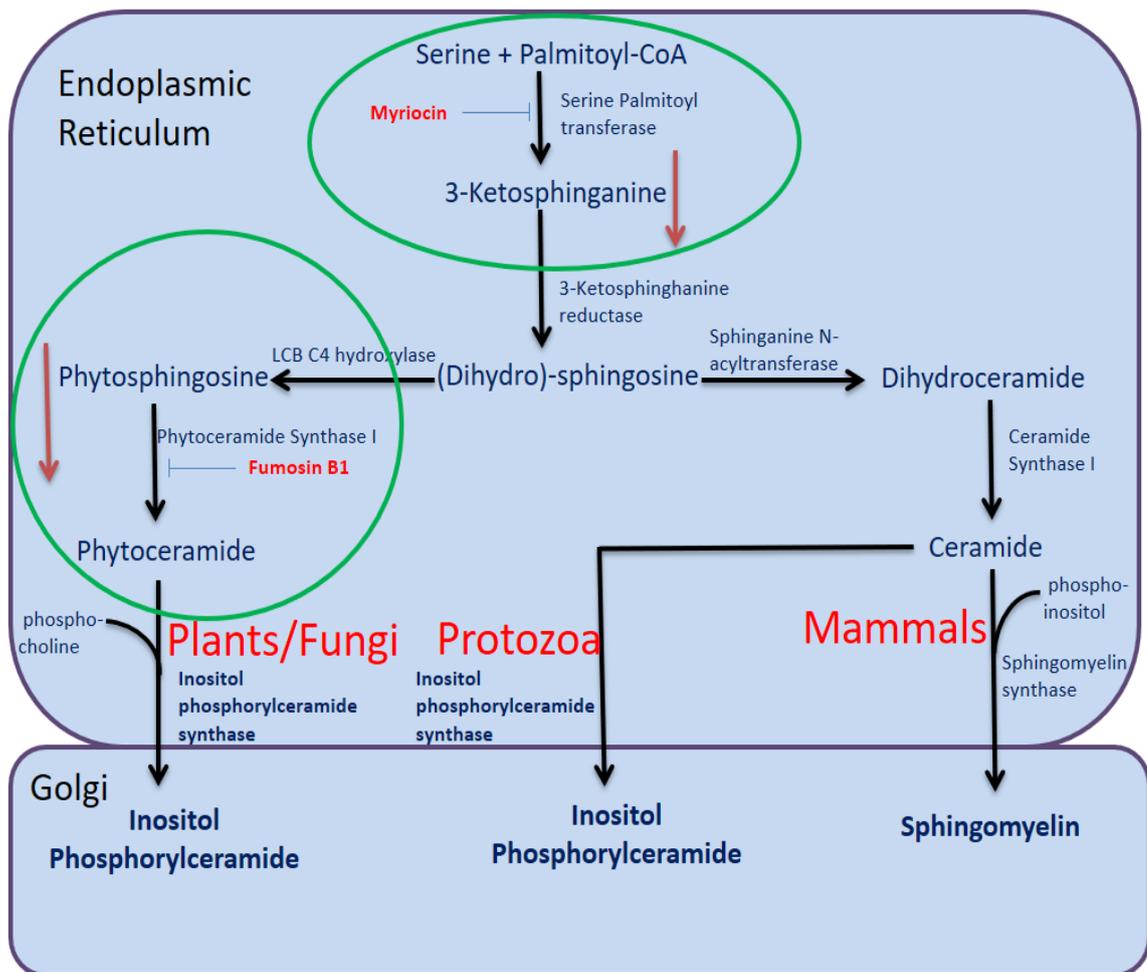
Figure 3-2: Inhibition of phytoceramide synthase I by fumosin B1 results in an increase of phytosphingosine with plants displaying PCD phenotypes[1] which are attenuated by the inhibition of SPT by myriocin, leading to a decrease in 3-ketosphinganine and phytosphingosine.

The findings of the studies outlined above, show that there is a consistent link between the disruption of sphingolipid synthesis, the buildup of intermediates in the biosynthetic pathway and the initiation of PCD. This biosynthetic pathway is clearly tightly regulated and inextricably linked to biological pathways/signaling involved in plant survival.

A homozygous knockout of IPCS2 in *Arabidopsis* was found to lead to elevated levels of SA and phytoceramide, HR-like lesion formation on leaves associated with PCD, and a reduction in plant stature[16]. A working hypothesis for the iinhibition of plant IPCS involves an offset in the balance between the pro-apoptotic compound, phytoceramide (substrate of plant IPCS) and the mitogenic compound, diacylglycerol (by-product), with

inhibition resulting in an accumulation of the former, with a concomitant decrease in the latter[17]. In addition, it would result in the decrease of the product inositol phosporylceramide (IPC), which is attached to a number of oligosaccharides to generate GIPC, a major class of sphingolipids in plants that make up a quarter of the total lipids in the plasma membrane[18].

 GIPCs have been found to be essential for normal growth and reproduction in *Arabidopsis*; homozygous knockout mutants of IPC glucuronosyltransferase 1 (IPUT1) are non-viable. To study the developmental effects of reduced GIPCs in *Arabidopsis*, a pollen specific rescue construct was used to select for homozygous knockout lines of *iput*, and these were found to display phenotypes of severe reduction in growth, increased levels of SA and lesions characteristic of PCD[19]. From these studies, it can be hypothesised that the inhibition of IPCS would result in a decrease in the sphingolipid precursor, IPC resulting in a decrease in GIPC which is essential for plant development. In addition, the build-up of the bioactive compound phytoceramide would initiate PCD, making IPCS a promising herbicidal target.

3.4 Primary screening

To identify putative inhibitors of IPCS, the focused library consisting of 11,440 bioactive compounds was screened against *At*IPCS2, *Os*IPCS2 and AUR1 formatted into the yeast-based assay described in Chapter 2. The IPCS2 isoform was selected for screening due to its well characterised role in plant defense and PCD[16, 20] and the fact that it is the most highly expressed orthologue being log 2-fold greater than both IPCS1 and IPCS3 in *A.thaliana*[21]. Screening was also carried out against the AUR1 complemented yeast strain as a control to filter out compounds that are generically cytotoxic to yeast cells rather than on target.

In the pharmaceutical industry the Lipinski rule of 5 is used as a gold standard for the selection of compounds that have orally-active drug like features[22]. The physico-chemical rules require that the compound has a molecular weight ≤ 500, clogP ≤5, no more than 5 H-bond donors and 10 H-bond acceptors. In the agrochemcical industry there is a need to have compounds that not only trigger cell death at a biochemical level but also have the right physico-chemical properties to penetrate the waxy cuticle and cell membranes of plants. Analyses of commercial herbicides and pesticides carried out by Tice, revealed that most have a molecular weight between 300-400 with an average of 329 for herbicides and 324 for insecticides[23]. In addition, the clogP values of herbicides were in the range 1.2-2.4 with an average of 1.7, which means that most herbicides are hydrophobic in nature which is necessary for the compounds to enter the cells and reach their target. According to Briggs rule of three, for a compound to show activity when sprayed on plants, it must have a molecular weight of 300±100, clogP=3±3, melting point ≤ 300°C and no more than 3 H-donors[24]. Based on these parameters, 50% of the compounds selected for screening had a molecular weight between 300-400 (Figure 3-3A), with 45% possessing a clogP of 3-5 (Figure 3-3B).

In the primary screen against *At*IPCS2, *Os*IPCS2 and AUR1, due to the configuration of the compound plates, the positive and negative control were repeatedly tested in the first row and alternated in terms of placement in these wells. At this stage, the library of compounds was tested in duplicate at 10 µM with the inhibition of yeast growth measured after 24 hours.
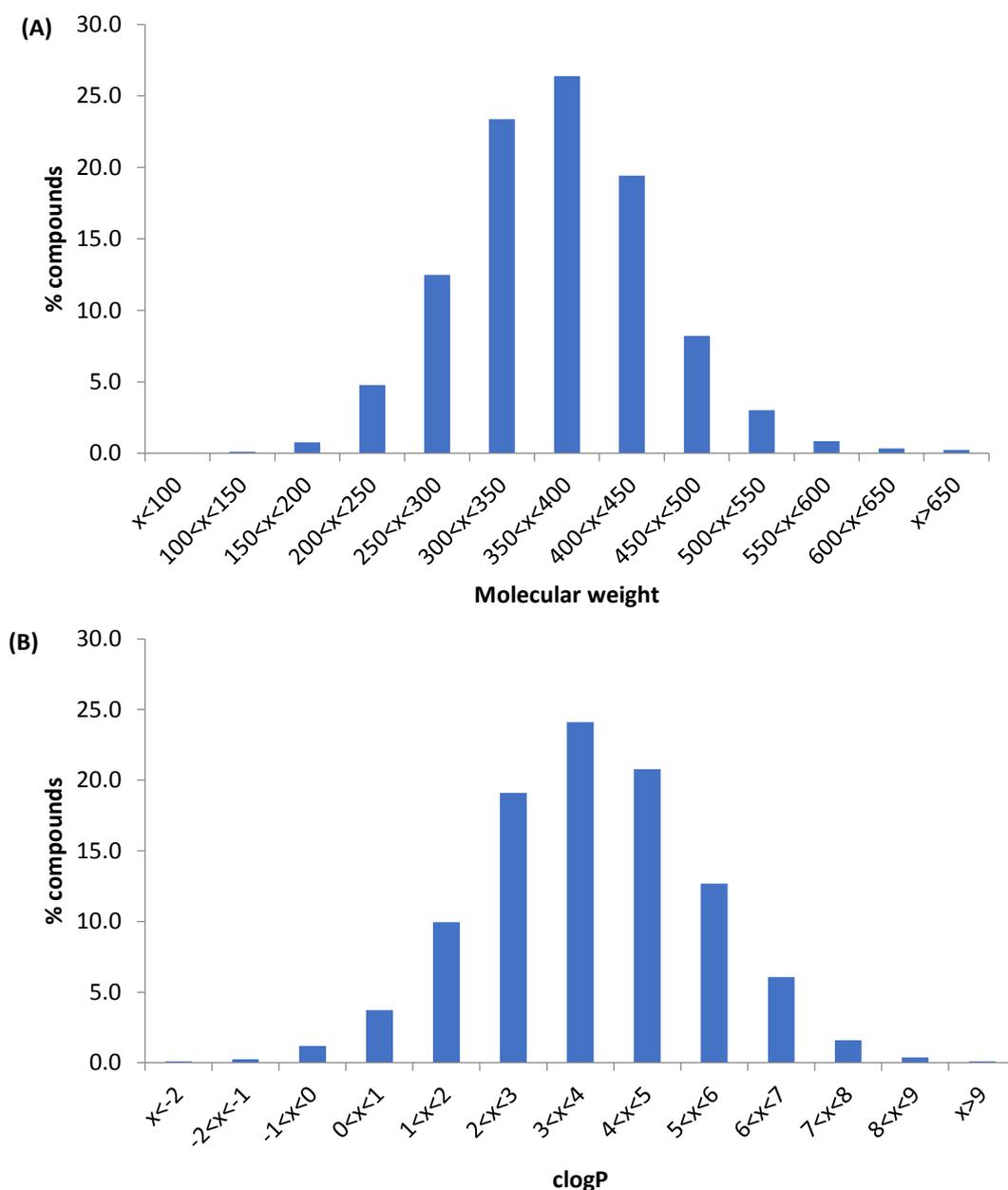
Figure 3-3: (A) Molecular weight and (B) clogP distribution of the library compounds screened.

To ensure reproducibility and reduce well to well variation, the plates were sealed with transparent breathable membranes which eliminated the 'edge efffect'[25] seen in assays incubated over long periods of time. For quality control, each plate in the yeast cell assay was required to have a calculated Z' factor of ≥ 0.5 in order for the assay results to be deemed as valid (Figure 3-4). Compounds exhibiting $\geq 80\%$ inhibition (the threshold) were selected to be taken forward. From the 11,440 compounds screened, 1164 compounds showed $\geq 80\%$ inhibition of AtIPCS2, 986 compounds showed $\geq 80\%$

inhibition of *Os*IPCS2, and 854 compounds had ≥80% activity against AUR1 (Figure 3-5). The false positive hit rate was 2.6%, and these compounds were excluded leaving a final hit rate of 1.5%.
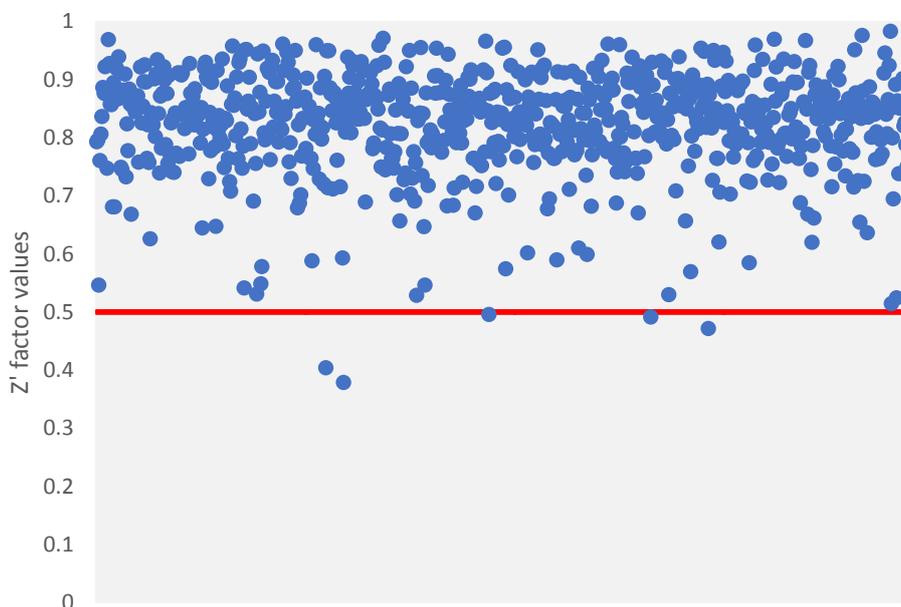
Figure 3-4: Z' factor of plates screened; cut off of 0.5 (red line) for a robust assay. Three plates fell below this value but were found to replicate the data in the n=2 screen and were accepted for further analyses.

To progress to secondary screening the compounds also had to show ≥ 50% differential selectivity for the plant IPCS over the yeast AUR1 orthologue. From the approximately 1000 compounds that showed selectivity, 126 met the criteria above, 40 compounds showed selectivity for AUR1 over the plant IPCS (Figure 3-6) which would be of interest to study for use as antifungals in future work.

Of the 126 plant selective compounds, 4 were compounds were of particular interest as they were known to be pro-apoptotic, and as discussed above, *At*IPCS2 is associated with PCD. One of these compounds showed inhibition of both AtIPCS2 and OsIPCS2 in the yeast-based assay, whilst the other three exhibited selective inhibition of *Os*IPCS2. Using the same assay platform, dose response experiments (50 µM to 68 nM) were carried out in duplicate for the *At*IPCS2 and *Os*IPCS2 hit compounds; 89 of the 105 *At*IPCS2

inhibitors, and 35 of the 39 *Os*IPCS2 hits had IC$_{50}$ values ≤ 10 μM (Figure 3-7 A-C). The

four pro-apoptotic compound hits are highlighted in red in Figure 3-6B and C.
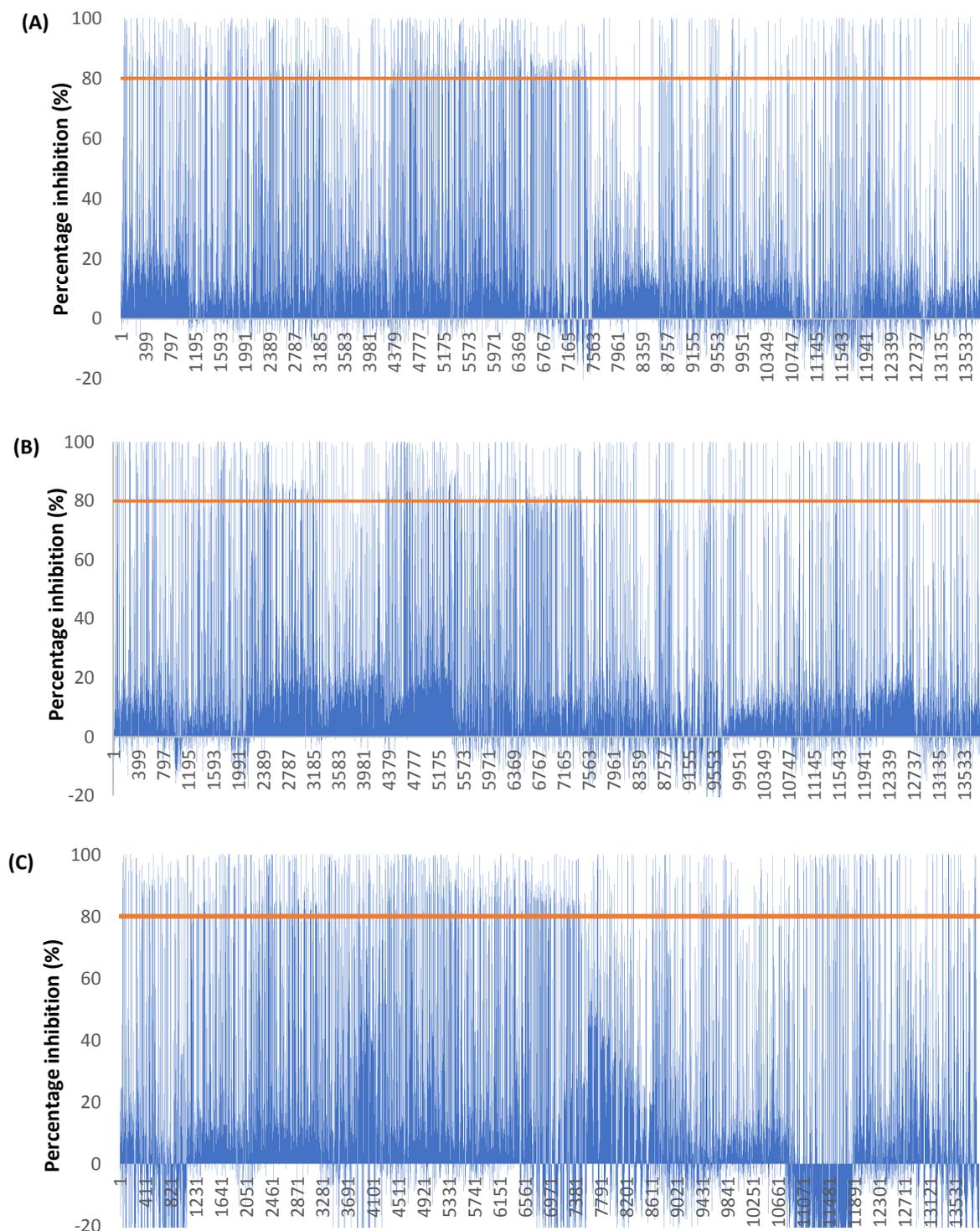
Figure 3-5: Plots showing percentage inhibition for compounds tested against (A) *At*IPCS2 (B) *Os*IPCS2 (C) AUR1 complemented yeast strains.
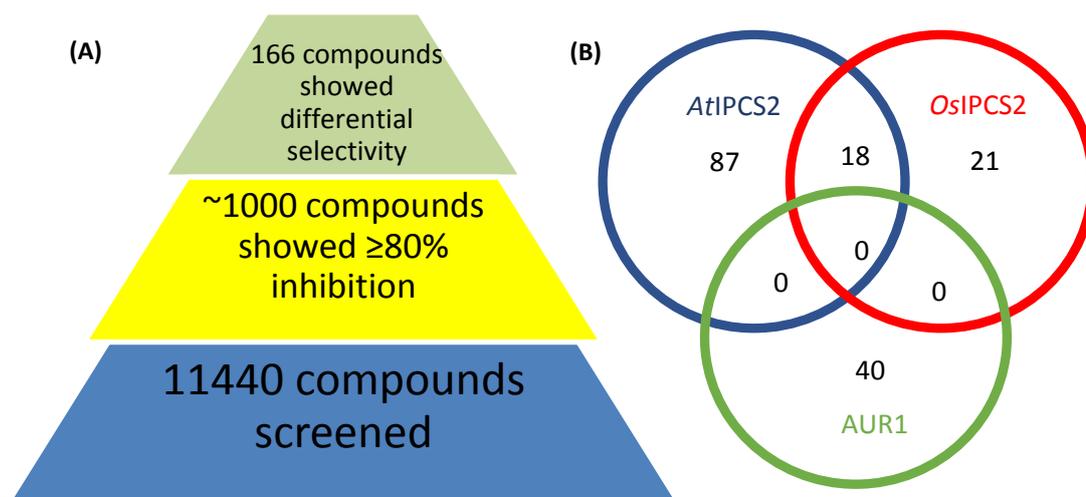
Figure 3-6: (A) A screening pyramid showing the number of compounds selected to progress to each stage of the screening campaign (B) Venn diagram of the 166 hits showing the differential selectivity for *At*ICS2, *Os*IPCS2 and AUR1.

Interestingly one of the pro-apoptotic compounds showed activity against both *At*IPCS2 and *Os*IPCS2 (6/12b; Figure 3-8A and B) whilst the other 3 hit only *Os*IPCS2. However, a closer inspection of the primary screen results shows that these compounds also targeted *At*IPCS2 but fell below the inhibition threshold of ≥80% required for the compounds to be considered hits. Of the 4 pro-apoptotic compounds identified as hits, 3 share a common core and exhibit low µM activity in the assay; compound 6/12b is a dihydropyridazinone, compound 6/8c is an isothiazole carboxamide, and compounds 6/12a and 62/10d are dihydrotriazinones (Figure 3-7). For the biochemical screen, all 126 hits were carried forward in order to prevent the loss of active compounds.

Structures redacted

6/12b – IC$_{50}$ = 3.1µM   6/8c – IC$_{50}$ = 35.8µM   6/12a – IC$_{50}$ = 2.3µM   62/10d – IC$_{50}$ = 4.5µM

6/12b – IC$_{50}$ = 6.8µM

Figure 3-7: Core structures of the four apoptotic inducer compounds with IC$_{50}$ values – highlighted in red are compounds which showed activity against *Os*IPCS2 and in blue those showing activity against *At*IPCS2 in the yeast-based primary assay.
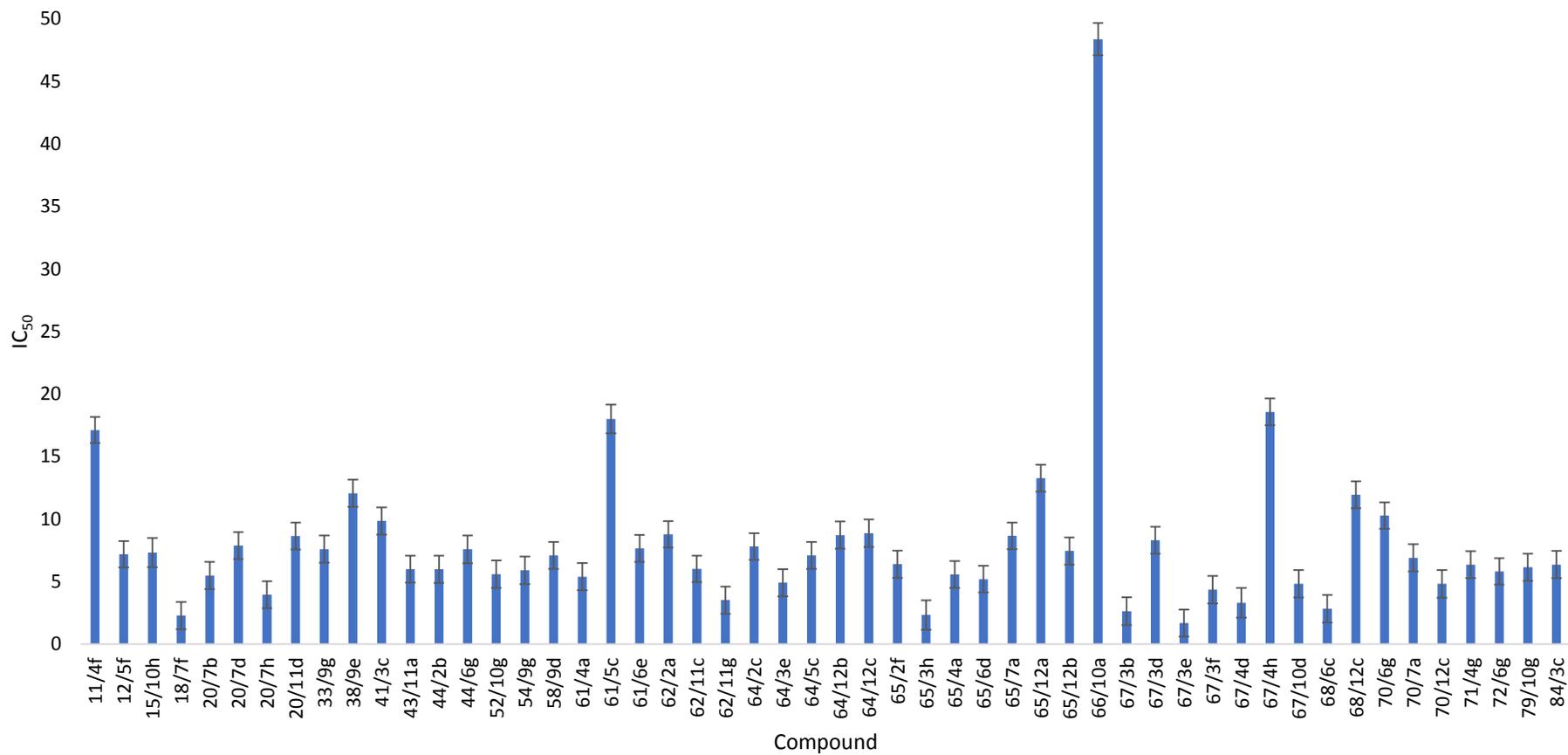
Figure 3-8: (A) IC$_{50}$ values for compounds tested against *At*IPCS2. Each value is representative of the mean of two independent biological replicates.
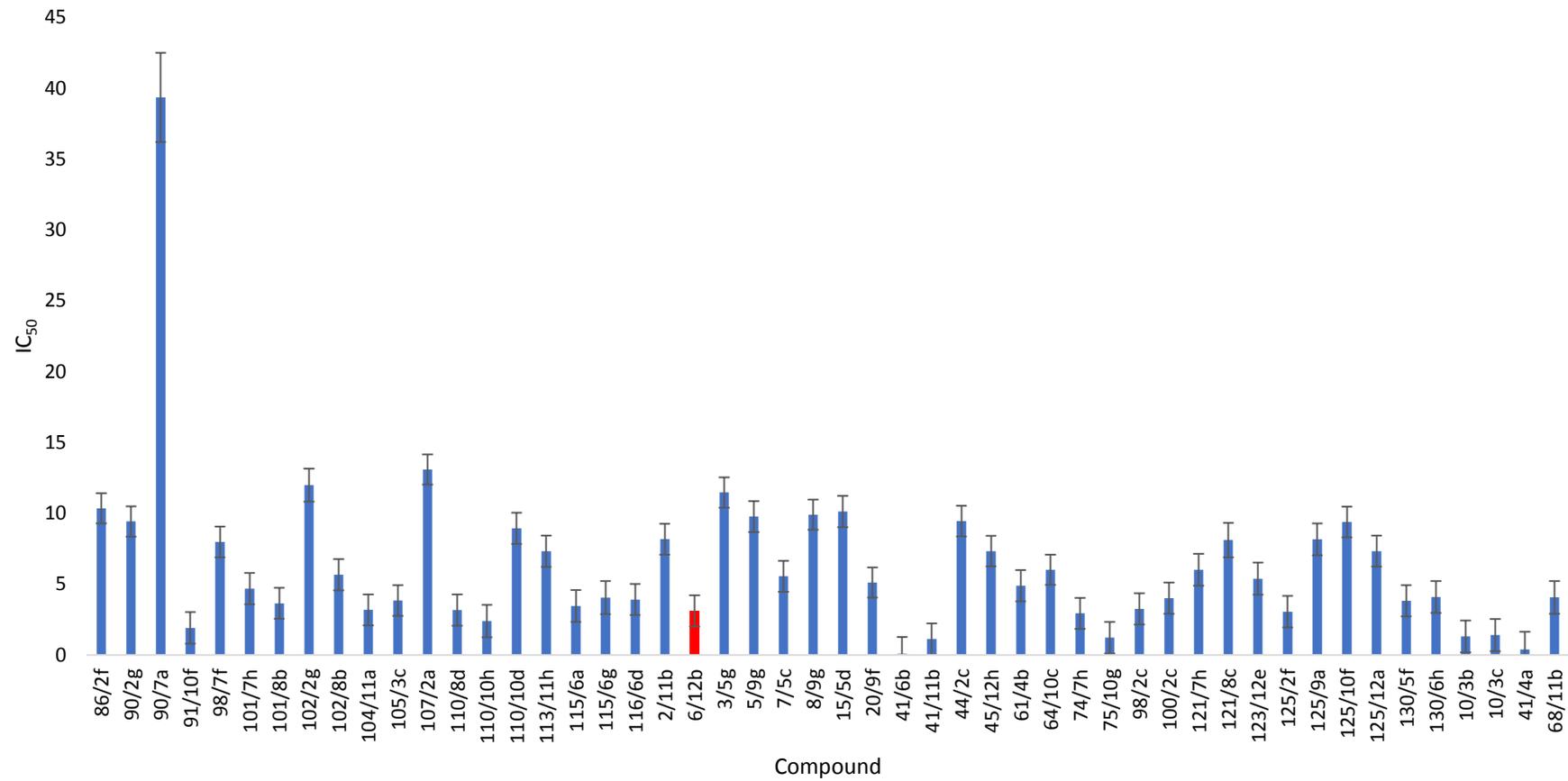
Figure 3-8: (B) IC$_{50}$ values for compounds tested against *At*IPCS2, with known PCD inducer in red. Each value is representative of the mean with standard error calculated for two independent biological replicates.
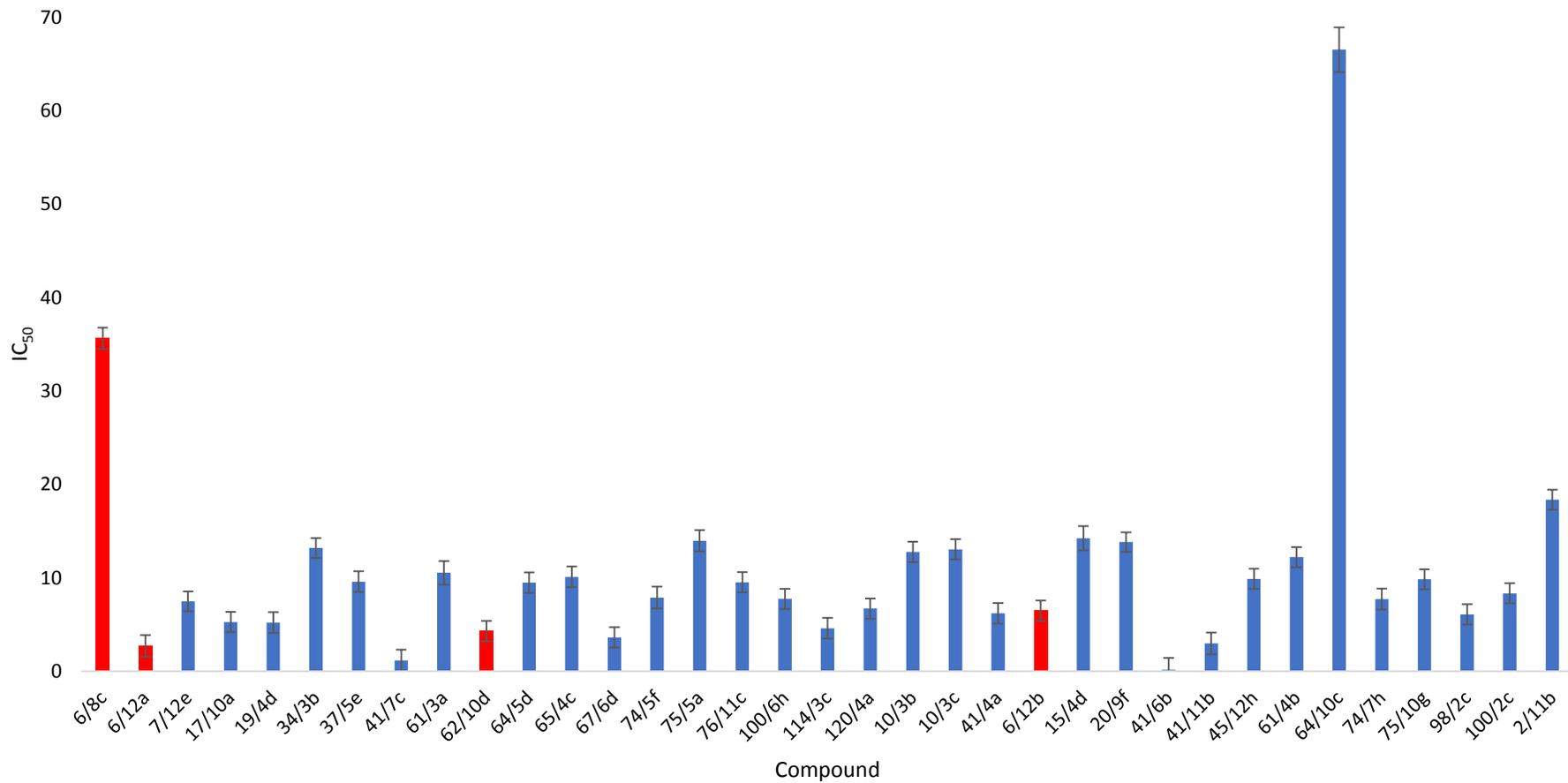
Figure 3-8: (C) IC$_{50}$ values for compounds tested against *Os*IPCS2, with known PCD inducer in red. Each value is representative of the mean with standard error calculated for two independent biological replicates.

3.5 Secondary screening

A 96-well plate assay to screen for inhibitors of IPCS had already been developed within our group[26]. This assay allows for the separation of the fluorescent product IPC from the fluorescent substrate NBD-ceramide/NBD-phytoceramide using ion exchange chromatography. Assay parameters for secondary screening were first established to determine the initial velocity of the reaction so that the optimal concentrations of substrates could be chosen to ensure that the enzyme is not saturated, and the reaction not limited by substrate concentrations. The linear range was used in the enzyme assay to allow accurate quantitation.

Following yeast microsome preparation as described, the activity of *At*IPCS2 and *Os*IPCS2 in the presence and absence of exogenous phosphatidylinositol (PI, soy or bovine) was quantified. The result shows that there is a 330% increase in *At*IPCS2 activity following the addition of exogenous soy PI. Bovine PI results in an increase of activity by about 200%, therefore unsurprisingly the *Arabidopsis* enzyme has a preference for the plant PI. However, *Os*IPCS2 lacks such a marked difference in activity in the presence
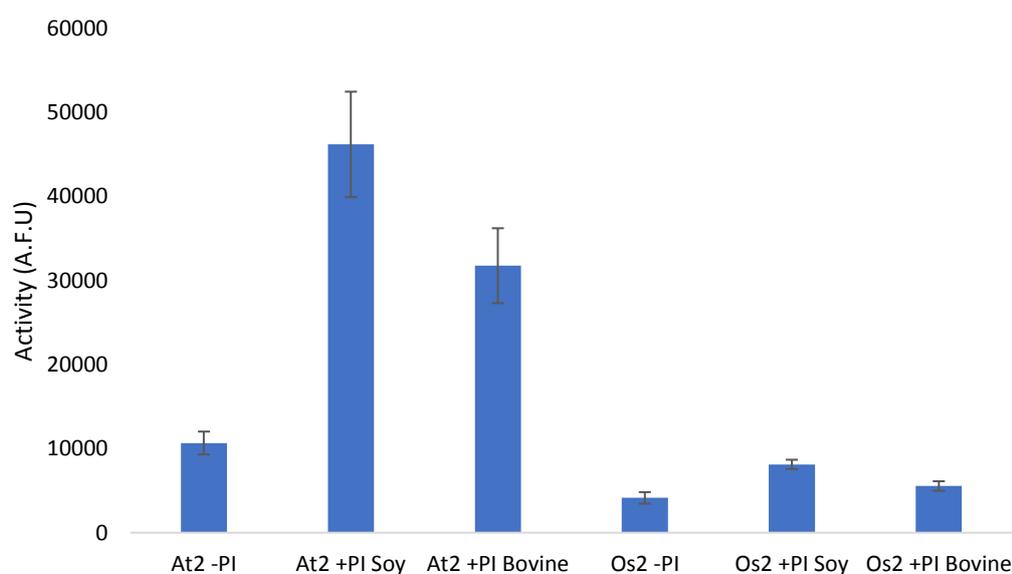
Figure 3-9: *At*IPCS2 and *Os*IPCS2 activity assayed in the presence and absence of exogenous PI variants (1 mM) and NBD ceramide (4 µM). Each value is representative of the mean with standard error calculated for three independent biological replicates.

and absence of exogenous PI from either source (Figure 3-9). The reasons for this are unclear, it is possible that rice preferentially utilises a different PI species.

In addition, to establish if the concentration of NBD-ceramide was limiting due to the much lower activity of *Os*IPCS2 microsomes, NBD-ceramide concentration was increased 3-fold to 12 µM. This lead to an increase in assayed *Os*IPCS2 activity (Figure 3-10), particularly the response to soy PI (2-fold compared to -PI), indicating that NBD-ceramide concentration had been limiting. In plants phytoceramide rather than ceramide is used as a natural substrate. To test if this is a preferred substrate, *At*IPCS2 and *Os*IPCS2 were assayed with 4µM NBD-phytoceramide in the presence and absence of PI (soy and bovine). In the presence of NBD-phytoceramide, *At*IPCS2 activity was similar when assayed with NBD-ceramide, although for reasons unknown, the enzyme now demonstrated a slight preference for bovine PI (Figure 3-11).
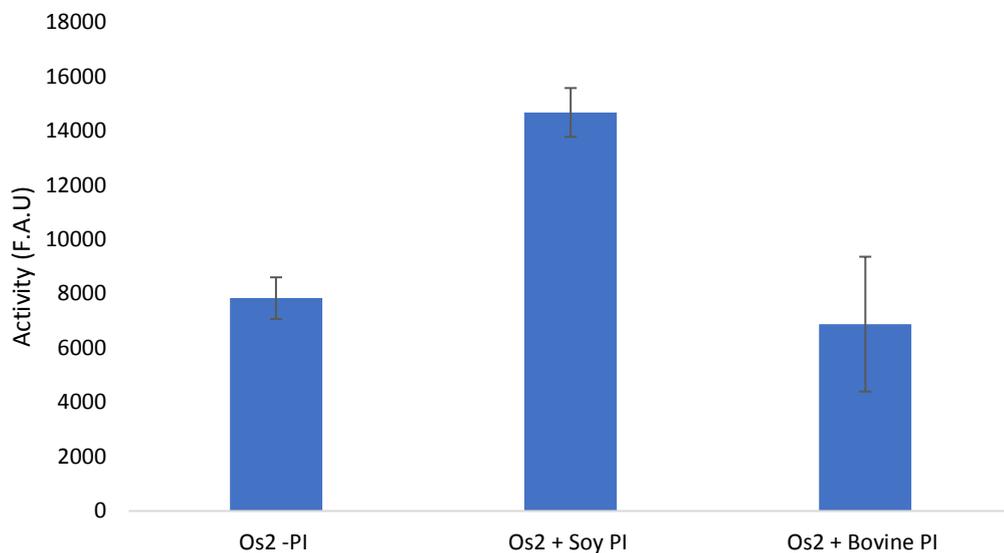
Figure 3-10: *Os*IPCS2 activity increased upon a 3-fold increase in NBD-ceramide to 12 µM and became more responsive to soy PI (1 mM). Each value is representative of the mean and the standard error calculated for three independent biological replicates.
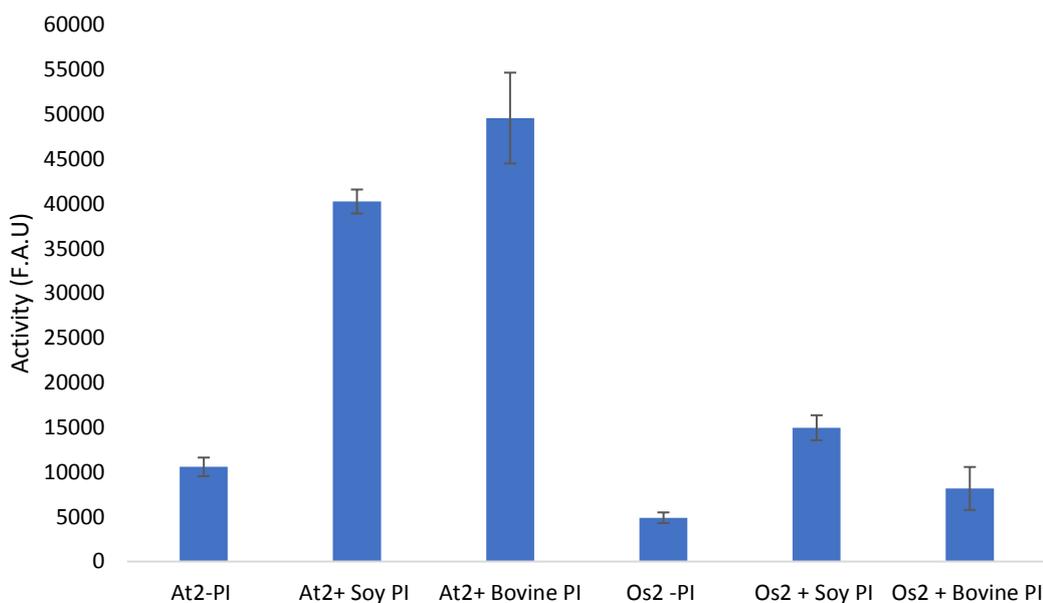
Figure 3-11: *At*IPCS2 and *Os*IPCS2 activity assayed with NBD-phytoceramide (4 µM) in the presence and absence of exogenous PI (1 mM). As with the increased concentration of NBD-ceramide, the activity *Os*IPCS2 is increased when using NBD-phytoceramide (4 µM), and more responsive to soy PI. Each value is representative of the mean and standard error calculated for three independent biological replicates.

In contrast, the activity of *Os*IPCS2 increased markedly (Figure 3-10) showing that phytoceramide is a better substrate for IPCS synthesis in plants. These data established that the optimum substrates were NBD-phytoceramide and soy PI; these were then utilised in all subsequent experiments.

To determine the enzyme concentration required for the assay to achieve a high background to noise ratio, different amounts of *At*IPCS2 (Figure 3-12) were tested in the 96-well plate assay. Following incubation for 60 minutes, the product was quantified, and it was clear that an increase in enzyme concentration resulted in a likewise increase in IPC turnover.

For *Os*IPCS2, increased quantities of microsomes were assayed to compensate for the relatively low activity, and a likewise increase in IPC turnover was observed which was associated with increasing enzyme concentration (Figure 3-13).
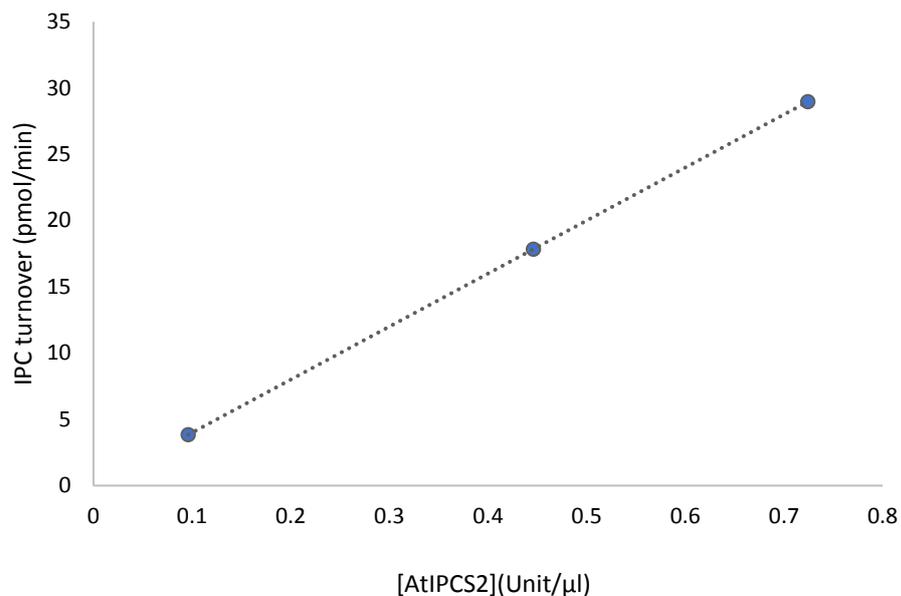
Figure 3-12: *At*IPCS2 concentration against IPC turnover. Each value is representative of the mean and standard error calculated for three independent biological replicates.

On the basis of these results (Figure 3-12 and 3-13), it was decided that 0.44 Unit/µl of *At*IPCS2 and 0.07 Unit/µl of *Os*IPCS2 would be utilised in the secondary screen. In addition to producing a good level of signal, this also meant that the limited amount of microsomes available for screening would be sufficient.
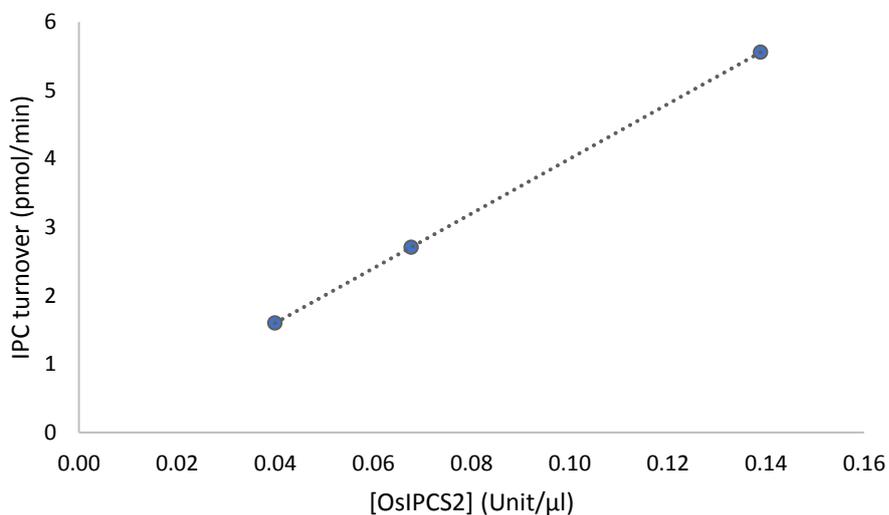
Figure 3-13: *Os*IPCS2 concentration against IPC turnover. Each value is representative of the mean and standard error calculated for three independent biological replicates.

Subsequently, various concentrations of soy PI (with NBD-phytoceramide at 4 µM) were tested to determine the amount necessary to achieve a high signal without jeopardising

the activity of the enzyme. *At*IPCS2 activity peaks at 100 µM PI, and at higher concentrations there was a considerable decrease in activity (Figure 3-14). The observed decrease in *At*IPCS2 activity resembled that seen for the *Leishmania* orthologue, *Lmj*IPCS, which was established as being due to surface dilution kinetics[27]. In comparison, the activity profile for *Os*IPCS2 shows a peak in activity at 200 µM, after which there is a decrease in enzyme turnover (Figure 3-15).
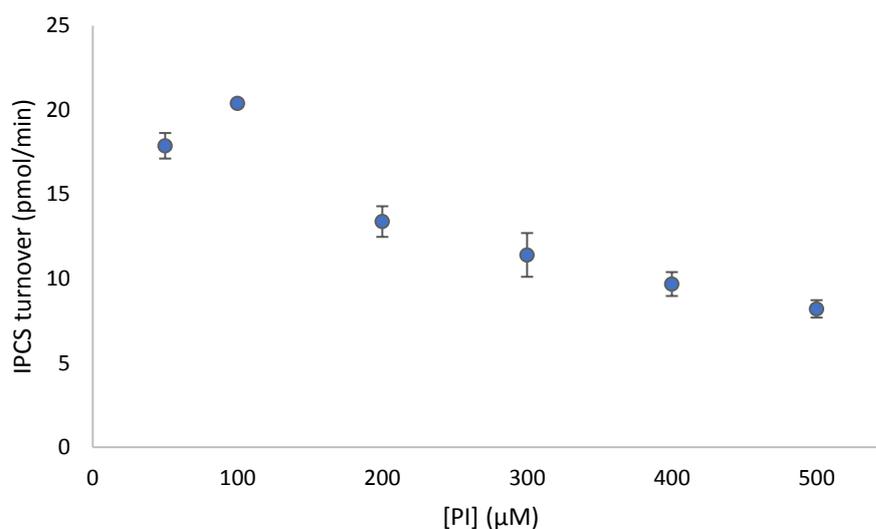
Figure 3-14: Soy PI concentration against *At*IPCS2 IPC turnover. Each value is representative of the mean and standard error calculated for three independent biological replicates.
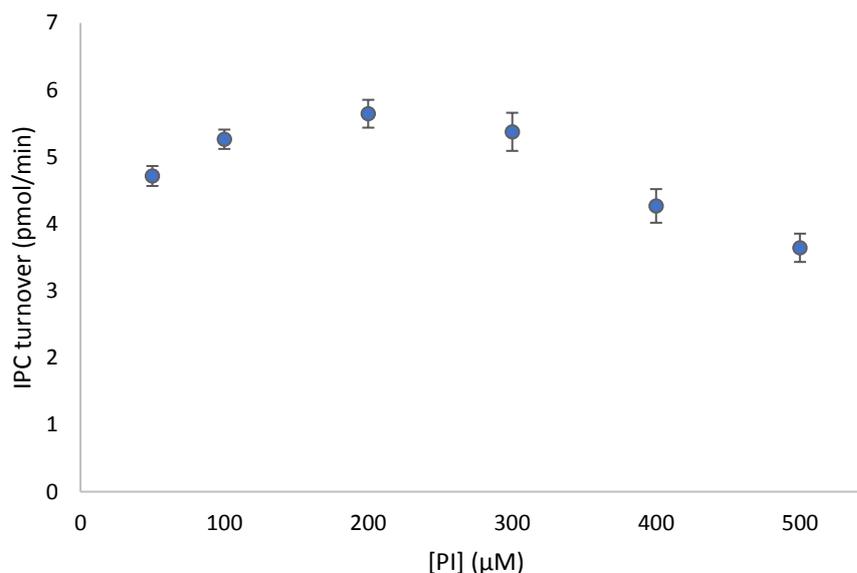
Figure 3-15: Soy PI concentration against *Os*IPCS2 IPC turnover. Each value is representative of the mean and standard error calculated for three independent biological replicates.

Based on these findings it was decided that 100 µM PI should be assayed against varying concentrations of NBD-phytoceramide to identify the optimum amount of NBD-phytoceramide for the assay (Figure 3-16).
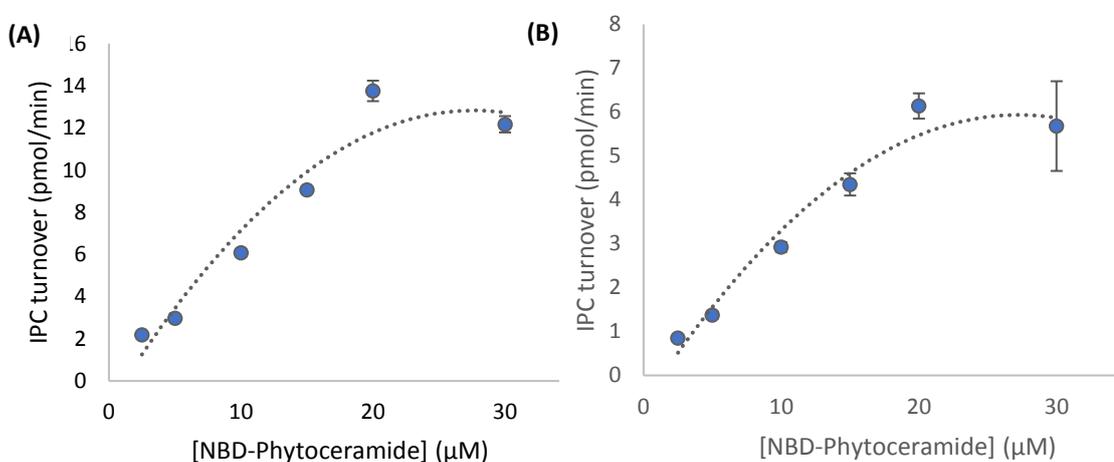
Figure 3-16: NBD-phytoceramide concentration against (A) *At*IPCS2 (B) *Os*IPCS2 IPC turnover. Each value is representative of the mean and standard error calculated for three independent biological replicates.

From this experiment carried out it was determined that 15 µM NBD-phytoceramide was the optimal concentration of NBD-phytoceramide, giving a high signal and being in the linear range. Therefore, the assay conditions were: 0.44 Unit/µl of *At*IPCS2 or 0.07 Unit/µl of *Os*IPCS2 100 µM soy PI and 15 µM NBD-phytoceramide, incubated at 30 °C for 60 minutes.

With AUR1 used as a control in the screening it was necessary to check that the set of conditions defined for assaying *At*IPCS2 and *Os*IPCS2 were suitable for this enzyme. Using the predetermined conditions (100 µM soy PI and 15 µM NBD-phytoceramide) the AUR1 reaction needed to be incubated for a maximum of 40 rather than 60 minutes to achieve a linear range (Figure 3-17).

With the assay parameters validated, screening commenced with 10 µM of each of the 126 compounds tested in duplicate. A threshold of $\geq 30\%$ percentage inhibition was set to identify hits to progress to dose response testing for IC_{50} determination.
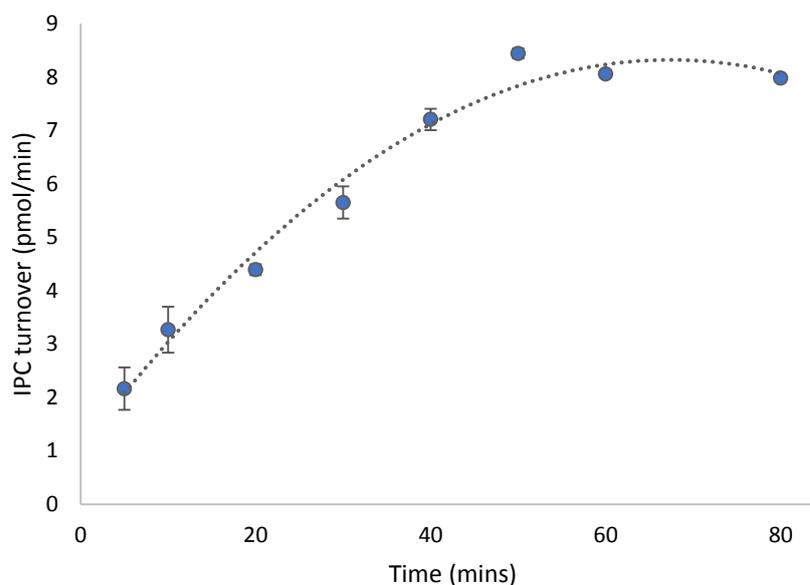
Figure 3-17: Incubation time against AUR1 reaction rate with 100 µM PI and 15 µM NBD-phytoceramide. Each value is representative of the mean and standard error calculated for three independent biological replicates.

Of the 126 compounds screened, 16 compounds were identified as hits; 14 were hits against *At*IPCS2 and 6 against *Os*IPCS2, 5 of these inhibited both enzymes. Dose response analyses was carried out for these compounds in triplicate against *At*ICS2, *Os*IPCS2 and AUR1 to determine the IC_{50} for each hit against the 3 enzymes. Of the 16 compounds tested only 2 showed activity against AUR1 at the highest concentration only (100 µM): compound 2 and 4. There was no activity at lower concentrations. In contrast, all compounds showed an inhibitory effect against the plant IPCSs.

Of the 4 compounds with an $IC_{50} \leq 10$ µM, 2 are phenylamidines (Compound 1 and 2) and the other 2 compounds are triazinone (Compound 3 and 4). These 4 compounds all showed selective inhibition of *At*IPCS2 over *Os*IPCS2 in the yeast-based assay, with IC_{50} values ranging from 3-5 µM for *At*IPCS2 compared to 15-31 µM for *Os*IPCS2. The other three pro-apoptotic compounds (Table 3-1, structures highlighted in red) also share the same triazinone core but had IC_{50} values 15-47 µM for the plant IPCSs and showed differential selectivity for *At*IPCS2 over *Os*IPCS2.

Table 3-1: IC50 determination and chemical structures of the 16 compounds identified as hits in the biochemical assay; highlighted in blue are the compounds that exhibited an IC$_{50}$ value ≤ 10 μM, in red are the pro-apoptotic compounds and in purple is one of the proapoptotic compounds that also had an IC$_{50}$ ≤ 10 μM against AtIPCS2

Compound ID	AtIPCS2 IC$_{50}$ (μM)	AtIPCS2 IC$_{50}$ 95% confidence interval (μM)	R square	OsIPCS2 IC$_{50}$ (μM)	OsIPCS2 IC$_{50}$ 95% confidence interval (μM)	R square	Compound structure
1 (15/10h)	4.7	3.5 to 7.3	0.99	13.9 to ???	13.9 to ???	0.77	Structure redacted
2 (113/11h)	4.0	3.3 to 4.9	0.99	61.9	33.6 to ???	0.94	Structure redacted
3 (6/12b)	8.4	5.2 to 22.0	0.94	43.8	undetermined	0.40	Structure redacted
4 (101/7h)	9.8	8.7 to 11.3	0.99	24.8	21.5 to ???	0.98	Structure redacted

						Structure redacted
5 (6/12a)	15.1	11.1 to 25.2	0.99	~ 38.6	undetermined	0.78	
						Structure redacted
6 (7/4c)	22.5	13.6 to 98.4	0.97	~ 35.3	undetermined	0.91	
						Structure redacted
7 (62/10d)	30.7	15.9 to 310.9	0.98	~ 46.6	undetermined	0.83	
						Structure redacted
8 (101/8b)	53.0	25.1 to ???	0.96	~ 44.2	undetermined	0.5	

9 (121/7h)	44.2	25.9 to 69712	0.96	~ 37.6	undetermined	0.91	**Structure redacted**
10 (120/4a)	56.2	35.5 to 96.8	0.97	~ 54633	undetermined	0.91	**Structure redacted**
11 (15/4d)	59.8	22.9 to ???	0.95	~ 28.9	undetermined	0.11	**Structure redacted**
12 (3/5g)	~0.0002	undetermined	0.14	~ 200.7	undetermined	0.31	**Structure redacted**

13 (41/4a)	53.9	27.1 to 132.1	0.94	undetermined	undetermined	undetermined	**Structure redacted**
14 (74/7h)	16.1	11.2 to 32.6	0.98	69.8	29.61 to ???	0.97	**Structure redacted**
15 (66/10a)	1456	undetermined	0.71	undetermined	undetermined	undetermined	**Structure redacted**
16 (84/7a)	1251	258.1 to ???	0.98	~ 64027	undetermined	0.95	**Structure redacted**

A literature search reveals that these compounds fall in to two compound classes: 1,2,4 triazin-6-ones (compounds 3-6) and phenylamidines (compounds 1 and 2). The 1,2,4 triazin-6-ones are fungicidal compounds which were filed for patent in 2001 by Dow Agrosciences[28]. Variants of this class have exhibited activity against the pathogenic fungi, *L. nodorum*, *M. grisea*, *U. maydis and Z. tritici*[29]. The phenylamidines are active pesticides which were first discovered in the 1960s by Bayer for control of plant fungi[30], with variants of these later patented for use as herbicides[31].

Structural analysis of the hits identified show that an aryl group attached to both nitrogens is necessary for enhanced activity against the plant IPCS; compound 8 which has an alkyne substituent attached to one of the nitrogens as opposed to an aryl group shows the lowest activity of the triazinones identified. The presence of thio or ether linker does not have much of an effect on plant IPCS activity and it seems that the presence of the more electron withdrawing fluorine attached to both aryl rings as opposed to chlorine and perhaps the presence of the fluorine at the meta position doubles the activity of the compound against *At*IPCS2 (deduction made by comparing compound 3 and 5). Addition of a naphthalene substituent attached to sulphur as opposed to an aryl ring also seems to decrease activity due to sterics. Interestingly, compound 7 showed lowest activity of the triazinones possessing an aryl group attached to both nitrogens; the drop in activity is probably due to the change in configuration of all the substituents (structure has been flipped in compound 7) which shows that for good activity it is essential that the configuration of the substituents be maintained for a specific interaction between the electron withdrawing groups and a binding site in *At*IPCS2.

3.6 *In vivo* screening

In vivo testing of the two compounds (compounds 1 and 2) that had IC_{50} values below 5 µM was undertaken against 7 days old *A. thaliana* seedlings. Seedlings grown were first

grown on agar containing no compound and then transferred to agar containing either compound 1 or 2. Seedling transferred to agar containing 10 µM of compound 2 had purple patches associated with anthocyanine biosynthesis in plant response to stress[32], and no lateral roots compared to seedlings grown on agar containing DMSO (Figure 3-18A), whilst seedlings grown on agar containing 10 µM of compound 1, showed no obvious difference compared to wildtype (Figure 3-18B). Seedlings grown on 40 µM of compound 2 exhibited the same phenotypes of purple patches on leaves and no lateral root development. However, seedlings grown on 40 µM of compound 1, had chlorotic phenotype usually associated with PCD[33] and no lateral root development compared to controls (Figure 3-18B). These results indicated that both compounds affected root development.
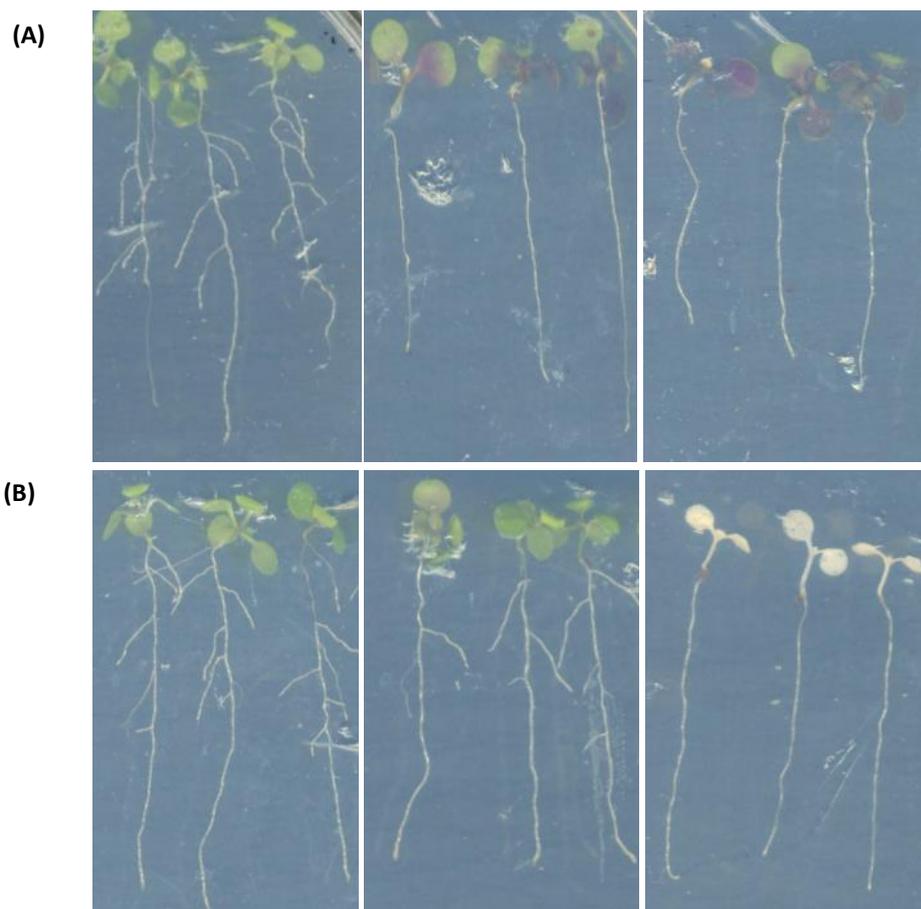
Figure 3-18: From left to right, 10 days old *A. thaliana* seedlings transferred to agar containing (A) 10 µL DMSO, 10 µM and 40 µM compound 2 (B) 40 µL DMSO, 10 µM and 40 µM of compound 1. Image taken 7 days after transfer of seedlings to agar containing the respective compounds.

Root growth was assessed by marking the root length after 4 days growth on the plates, then the changes in root length were measured after 3 days. No change in root length was observed in seedlings grown on 10 µM or 40 µM of compound 2. However, a significant difference was found between the primary root length of seedling grown on 10 µM of compound 1, compared to wildtype, with a reduction in root length by 30% (Figure 3-19) and seedlings grown on 40 µM of compound 1 showing no root development.
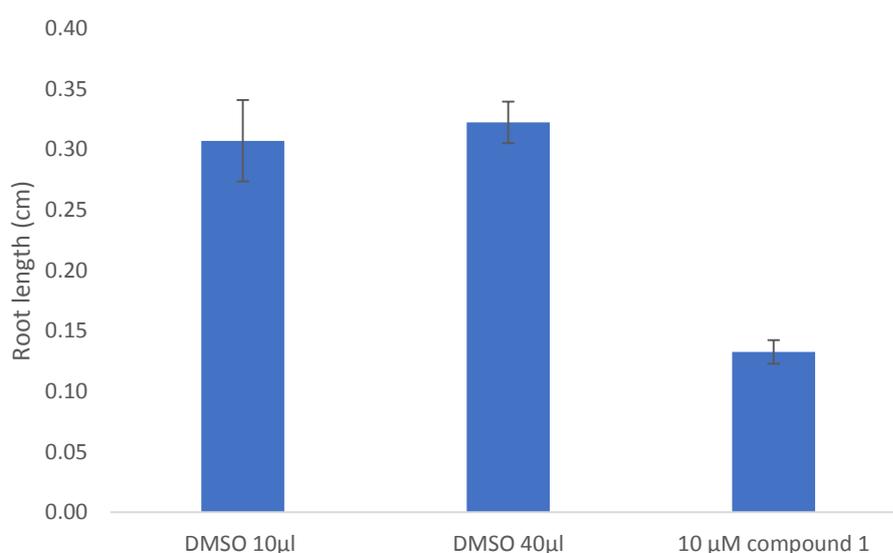
Figure 3-19: Root length measurements of seedlings grown in DMSO and in the presence of 10 µM of compound 1. Each value is representative of the mean of six replicates.

3.7 Conclusion

The screening campaign to identify inhibitors of the plant IPCS led to the discovery of the phenylamidine and triazinone class of compounds as selective inhibitors of *At*IPCS over AUR1, with the phenylamidines validated in an in vivo screen against *A. thaliana* seedlings. The phenylamidines exhibit low µM activity against *At*IPCS2 in vitro causing chlorotic leaves, stunted development of lateral roots and primary root length in treated seedlings. These findings are compelling and raise the question of an interplay between sphingolipid biosynthesis and root development. The differential selectivity of these compounds for *At*IPCS2 over *Os*IPCS2 and AUR1 in the yeast-based assay and the

enzyme assay, also raises the question as to how the compounds are interacting with the target; there is a high degree of sequence identity between *At*IPCS2 and *Os*IPCS2 (68%), both sharing well conserved domains, so it would be interesting to explore how these compounds are able to selectively inhibit one over the other. The differential selectivity of the inhibitors identified for the dicot (*A. thaliana*) over the monocot (*O. sativa*) IPCS can be utilized to create herbicide formulations that selectively kill dicot weeds without harming monocot crops.

3.8 References

1.   Saucedo‑García, M.; Guevara‑García, A.; González‑Solís, A.; Cruz‑García, F.; Vázquez‑Santana, S.; Markham, J. E.; Lozano‑Rosas, M. G.; Dietrich, C. R.; Ramos‑Vega, M.; Cahoon, E. B., MPK6, sphinganine and the LCB2a gene from serine palmitoyltransferase are required in the signaling pathway that mediates cell death induced by long chain bases in Arabidopsis. *New Phytologist* **2011,** *191* (4), 943-957.
2.   Choi, H.; Kim, J.-Y.; Chang, Y. T.; Nam, H. G., Forward chemical genetic screening. In *Arabidopsis Protocols*, Springer: 2014; pp 393-404.
3.   Serrano, M.; Kombrink, E.; Meesters, C., Considerations for designing chemical screening strategies in plant biology. *Frontiers in plant science* **2015,** *6*.
4.   Surpin, M.; Rojas-Pierce, M.; Carter, C.; Hicks, G. R.; Vasquez, J.; Raikhel, N. V., The power of chemical genomics to study the link between endomembrane system components and the gravitropic response. *Proceedings of the National Academy of Sciences of the United States of America* **2005,** *102* (13), 4902-4907.
5.   Savaldi-Goldstein, S.; Baiga, T. J.; Pojer, F.; Dabi, T.; Butterfield, C.; Parry, G.; Santner, A.; Dharmasiri, N.; Tao, Y.; Estelle, M., New auxin analogs with growth-promoting effects in intact plants reveal a chemical strategy to improve hormone delivery. *Proceedings of the National Academy of Sciences* **2008,** *105* (39), 15190-15195.
6.   Poretska, O.; Yang, S.; Pitorre, D.; Rozhon, W.; Zwerger, K.; Uribe, M. C.; May, S. T.; McCourt, P.; Poppenberger, B.; Sieberer, T., The small molecule hyperphyllin enhances leaf formation rate and mimics shoot meristem integrity defects associated with AMP1 deficiency. *Plant physiology* **2016**, pp. 01633.2015.
7.   Lin, L.-C.; Hsu, J.-H.; Wang, L.-C., Identification of novel inhibitors of 1-aminocyclopropane-1-carboxylic acid synthase by chemical screening in Arabidopsis thaliana. *Journal of Biological Chemistry* **2010,** *285* (43), 33445-33456.
8.   Schreiber, K. J.; Nasmith, C. G.; Allard, G.; Singh, J.; Subramaniam, R.; Desveaux, D., Found in translation: high-throughput chemical screening in Arabidopsis thaliana identifies small molecules that reduce Fusarium head blight disease in wheat. *Molecular plant-microbe interactions* **2011,** *24* (6), 640-648.
9.   Hicks, G. R.; Raikhel, N. V., Small molecules present large opportunities in plant biology. *Annual review of plant biology* **2012,** *63*, 261-282.
10.   Brodersen, P.; Petersen, M.; Pike, H. M.; Olszak, B.; Skov, S.; Ødum, N.; Jørgensen, L. B.; Brown, R. E.; Mundy, J., Knockout of Arabidopsis accelerated-cell-

death11 encoding a sphingosine transfer protein causes activation of programmed cell death and defense. *Genes & Development* **2002,** *16* (4), 490-502.

11.   Rivas-San Vicente, M.; Larios-Zarate, G.; Plasencia, J., Disruption of sphingolipid biosynthesis in Nicotiana benthamiana activates salicylic acid-dependent responses and compromises resistance to Alternaria alternata f. sp. lycopersici. *Planta* **2013,** *237* (1), 121-136.

12.   Chen, M.; Han, G.; Dietrich, C. R.; Dunn, T. M.; Cahoon, E. B., The essential nature of sphingolipids in plants as revealed by the functional identification and characterization of the Arabidopsis LCB1 subunit of serine palmitoyltransferase. *The Plant Cell* **2006,** *18* (12), 3576-3593.

13.   Ternes, P.; Feussner, K.; Werner, S.; Lerche, J.; Iven, T.; Heilmann, I.; Riezman, H.; Feussner, I., Disruption of the ceramide synthase LOH1 causes spontaneous cell death in Arabidopsis thaliana. *New Phytologist* **2011,** *192* (4), 841-854.

14.   Chen, M.; Markham, J. E.; Dietrich, C. R.; Jaworski, J. G.; Cahoon, E. B., Sphingolipid long-chain base hydroxylation is important for growth and regulation of sphingolipid content and composition in Arabidopsis. *The Plant Cell* **2008,** *20* (7), 1862-1878.

15.   Bi, F.-C.; Liu, Z.; Wu, J.-X.; Liang, H.; Xi, X.-L.; Fang, C.; Sun, T.-J.; Yin, J.; Dai, G.-Y.; Rong, C., Loss of ceramide kinase in Arabidopsis impairs defenses and promotes ceramide accumulation and mitochondrial H2O2 bursts. *The Plant Cell* **2014,** *26* (8), 3449-3467.

16.   Wang, W.; Yang, X.; Tangchaiburana, S.; Ndeh, R.; Markham, J. E.; Tsegaye, Y.; Dunn, T. M.; Wang, G.-L.; Bellizzi, M.; Parsons, J. F., An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in Arabidopsis. *The Plant Cell* **2008,** *20* (11), 3163-3179.

17.   Cerbón, J.; del Carmen López-Sánchez, R., Diacylglycerol generated during sphingomyelin synthesis is involved in protein kinase C activation and cell proliferation in Madin-Darby canine kidney cells. *Biochemical Journal* **2003,** *373* (3), 917-924.

18.   Markham, J. E.; Li, J.; Cahoon, E. B.; Jaworski, J. G., Separation and identification of major plant sphingolipid classes from leaves. *Journal of Biological Chemistry* **2006,** *281* (32), 22684-22694.

19.   Tartaglio, V.; Rennie, E. A.; Cahoon, R.; Wang, G.; Baidoo, E.; Mortimer, J. C.; Cahoon, E. B.; Scheller, H. V., Glycosylation of inositol phosphorylceramide sphingolipids is required for normal growth and reproduction in Arabidopsis. *The Plant Journal* **2017,** *89* (2), 278-290.

20.   Mangwanda, R.; Myburg, A. A.; Naidoo, S., Transcriptome and hormone profiling reveals Eucalyptus grandis defence responses against Chrysoporthe austroafricana. *BMC genomics* **2015,** *16* (1), 319.

21.   Mina, J.; Okada, Y.; Wansadhipathi-Kannangara, N.; Pratt, S.; Shams-Eldin, H.; Schwarz, R.; Steel, P.; Fawcett, T.; Denny, P., Functional analyses of differentially expressed isoforms of the Arabidopsis inositol phosphorylceramide synthase. *Plant molecular biology* **2010,** *73* (4-5), 399-407.

22.   Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews* **1997,** *23* (1-3), 3-25.

23.   Tice, C. M., Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals? *Pest management science* **2001,** *57* (1), 3-16.

24.   Briggs, G. In *Predicting uptake and movement of agrochemicals from physical properties*, Presentation at SCI Meeting, in: Uptake of Agrochemicals and Pharmaceuticals, Belgrave Square, London, England, 9th December, 1997.

25.   Maddox, C. B.; Rasmussen, L.; White, E. L., Adapting cell-based assays to the high-throughput screening platform: problems encountered and lessons learned. *JALA: Journal of the Association for Laboratory Automation* **2008,** *13* (3), 168-173.

26.   Mina, J., Thesis: Exploring IPC synthase: a potential anti-leischmanial drug target. **2010**.

27.   Mina, J. G.; Mosely, J. A.; Ali, H. Z.; Shams-Eldin, H.; Schwarz, R. T.; Steel, P. G.; Denny, P. W., A plate-based assay system for analyses and screening of the Leishmania major inositol phosphorylceramide synthase. *International Journal of Biochemistry & Cell Biology* **2010,** *42* (9), 1553-1561.

28.   Kelly, M. J.; Evans, K. A.; Gallagher, J. J.; Mulvihill, M. J., Substituted 4, 5-dihydro-1, 2, 4-triazin-6-ones, 1, 2, 4-triazin-6-ones and their use as fungicides and insecticides. Google Patents: 2009.

29.   Owen, W. J.; Sullenberger, M. T.; Loso, M. R.; Meyer, K. G.; Slanec, T. J., Synthesis and antifungal activity of 3‐aryl‐1, 2, 4‐triazin‐6‐one derivatives. *Pest management science* **2015,** *71* (1), 83-90.

30.   Dieter, D.; Hans, A.; Ludwig, E., Method for protecting plants from fungi. Google Patents: 1966.

31.   Kuhn, B.; Jakobi, H.; Mueller, T.; Tiebes, J.; Kehne, H.; Schmutzler, D.; Hills, M.; Rosinger, C.; Kunz, K.; Drewes, M. W., Use of n2-phenylamidines as herbicides. Google Patents: 2008.

32.   Chalker‐Scott, L., Environmental significance of anthocyanins in plant stress responses. *Photochemistry and photobiology* **1999,** *70* (1), 1-9.

33.   Lam, E.; Fukuda, H.; Greenberg, J., *Programmed cell death in higher plants*. Springer Science & Business Media: 2000; Vol. 44.

CHAPTER 4

VALIDATION OF IPCS TRANSGENIC LINES FOR FUNCTIONAL GENOMIC ANALYSES

4.1 Chapter synopsis

In this chapter, the creation and characterisation of *Arabidopsis thaliana* transgenic lines over-expressing *AtIPCS1-3*, and knock-down lines of *AtIPCS1-3* are presented. Transgenic lines of 35s::GFP.*At*IPCS1, 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 were also created for the analysis of protein localisation in *Arabidopsis thaliana* and *Nicotiana benthamiana*. These transgenic lines have been created to probe the changes in the transcriptome of knock-down/over-expression lines in response to a change in the expression of each orthologue. Also described in this chapter is the phenotypic screen of the transgenic lines, which identified an early flowering phenotype indicating a link between IPCS and vernalisation.

4.2 RNA interference, over-expression and localisation of *At*IPCS1-3 using gateway cloning

Creating multiple expression constructs for the target gene of interest (*AtIPCS1-3*) using traditional ligase-mediated cloning is time-consuming and laborious, hence the use of gateway cloning. Gateway cloning exploits the bacteriophage lambda recombination system, bypassing the use of ligase-mediated cloning[1].

For RNA interference (RNAi) knock-down of a target of interest, double stranded RNA from an inverted repeat of the target sequence was expressed, resulting in post-transcriptional gene silencing (PTGS). The classical understanding of RNAi triggered by the presence of an inverted repeat gene in plants involves transcription of the DNA sequence resulting in the formation of double stranded RNA (dsRNA). Following this, cleavage of dsRNA is mediated by the Dicer enzymes producing 21-25 nucleotides long siRNAs, which were first identified in plant lines harbouring a transgene that induced PTGS[2]. RNAi is carried out by the RNA induced silencing complex (RISC) a protein–RNA effector complex, that utilizes the complementarity of the siRNA sequences to their target mRNA, resulting in degradation of the target mRNA mediated by the Argonaut proteins[3]. This leads to silencing or expression knock-down of the target.

Knock-down or silencing of an IPCS gene can be achieved by the introduction of exogenous miRNAs designed to specifically target one of the three IPCS isoforms. The system utilized in this study is attractive as an alternative to the use of miRNAs, in that the siRNA produced from the dsRNA could potentially target the expression of one or more of the IPCS isoforms because of the high sequence identity shared; making it possible to observe the cumulative effect of differential expression of each IPCS isoform on plant development.

The Gateway 'destination' vectors used in this study were designed by Karimi et al.[4]; the pK7WG2 vector was used for over-expression, the pK7GWIWG2 (I) for silencing and the pK7FWG2 for localisation by N-terminal tagging of each IPCS orthologue with green fluorescent protein (GFP).

Each IPCS orthologue (*AtIPCS1-3*) was cloned into the over-expression vector (pK7WG2), the silencing vector (pK7GWIWG2) and pK7FWG2 for GFP N-terminal tagging. The clones were verified as described in Materials and Methods. Subsequently, *Agrobacterium tumefaciens* (strain GV1301) mediated transformation of wild-type *Arabidopsis* (Col0) with the constructs was carried out to generate the T0 generation.

4.2.1 *At*IPCS1-3 localisation

For preliminary localisation of *At*IPCS1-3, the N-terminal GFP-tagged IPCSs constructs were used to transform *Nicotiana benthamiana* using agrobacterium strain GV1301 for transient expression of the protein. The 35s::GFP cytosolic marker was also transiently expressed as a control.
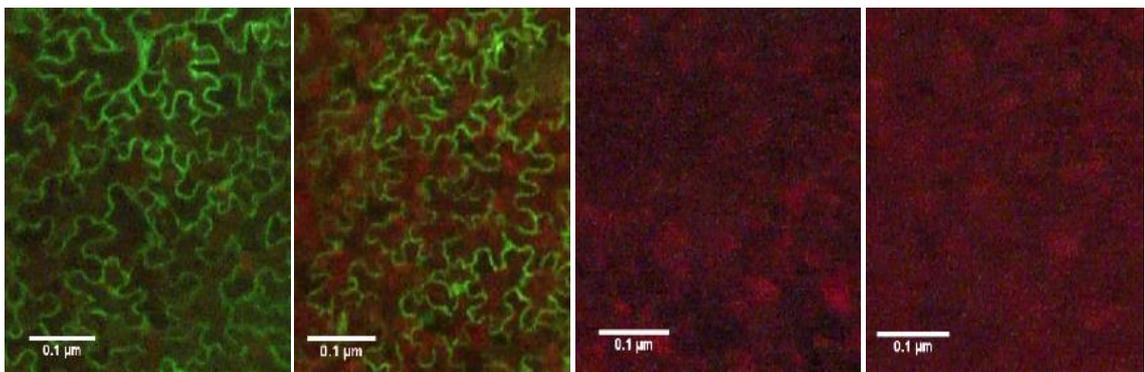
Figure 4-1: From left to right, expression of 35s::GFP cytosolic marker, 35s::GFP.*At*IPCS1, 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 imaged 48 hours after infiltration of *Nicotiana benthamiana* with agrobacterium harbouring respective constructs for each isoform.

The first attempts at the transient expression of 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 failed with no fluorescence observed (Figure 4-1). However, 35s::GFP.*At*IPCS1 exhibited high levels of expression (Figure 4-1). A closer look using confocal microscopy revealed

that the expression pattern was indicative of cytosolic expression (Figure 4-2).

The cytosolic localisation of a multi-transmembrane protein such as *At*IPCS1 was surprising given previous work by Wang et al. which had demonstrated that *At*IPCS2 is localised to the trans Golgi network[5]. The results from 35s::GFP.*At*IPCS1, 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 transient expression suggested that the plant had pre-empted the expression of GFP tagged IPCS, resulting in free cytosolic GFP expression for the 35s::GFP.*At*IPCS1 construct and no expression for 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 constructs. The function of *At*IPCS2 as a negative regulator of PCD is well characterized[5]; *At*IPCS role in defence response could result in the silencing of 35s::GFP.*At*IPCS expression in planta to pre-empt an incompatible reaction or could be due to the instability of the fusion protein.

To ensure that the cytosolic expression of 35s::GFP.*At*IPCS1 was not as a result of over-expression due to multiple t-DNA insertions in a cell which is a consequence of high numbers of bacteria, the optical density of agrobacterium used to infiltrate *N. benthamiana* with the localisation constructs was reduced by 1/3.
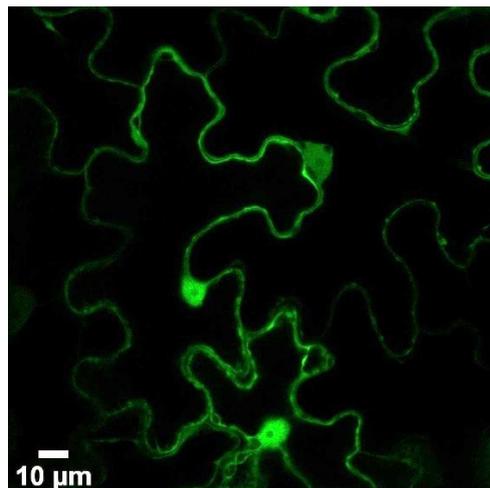
Figure 4-2: Subcellular localisation of 35s::GFP.*At*IPCS1 imaged 48 hours after tobacco infiltration with agrobacterium harbouring 35s::GFP.*At*IPCS1 expression construct.

This was also beneficial for testing 35s::GFP.*At*IPCS1 and 2 to find out if suppression of expression was a result of the detection of high levels of *At*IPCS transcript. 35s::GFP.*At*IPCS1, 2 and 3 were co-infiltrated respectively with RFP tagged sialyl transferase (ST) which is localised to the Golgi[6]. From the results (Figure 4-3) it was clear that the expression and localisation of ST RFP to the Golgi was successful, however, no expression could be seen for 35s::GFP.*At*IPCS1-3. With the unsuccessful expression of 35s::GFP.*At*IPCS1-3 in *N. benthamiana*, 35s::GFP.IPCS expression was checked in *A. thaliana* transgenic lines. From the T2 generation; RNA was isolated from the 10 day old seedlings, and following cDNA synthesis, real-time PCR was used to quantify 35s::GFP.*At*IPCS1, 2 and 3 expression.
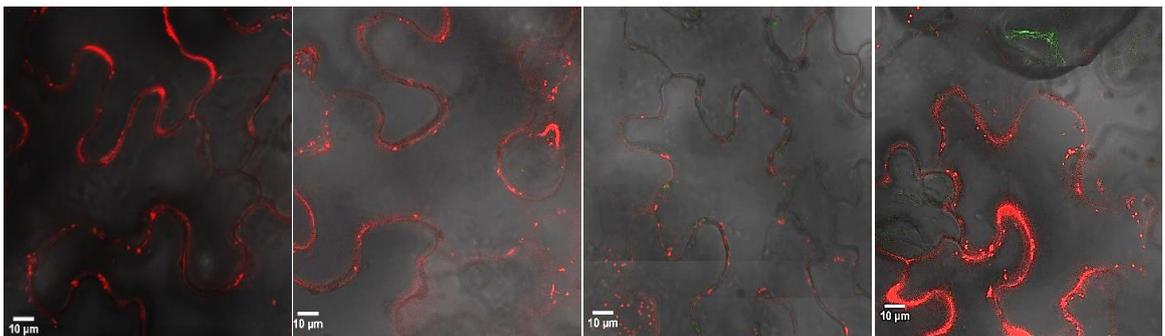
Figure 4-3: left to right, expression of the Golgi marker ST RFP, merged image of GFP tagged *At*IPCS1 and ST RFP, merged image of GFP tagged *At*IPCS2 and ST RFP, merged image of GFP tagged *At*IPCS3 and ST RFP. Plasmid DNA of ST RFP and 35s::GFP.*At*IPCS1, 35s::GFP.*At*IPCS2 and 35s::GFP.*At*IPCS3 respectively was introduced into tobacco cells by infiltration with agrobacterium and imaged with a confocal microscope 72 hours after infiltration.

Analyses were carried out using the comparative cycle threshold (C_t) method which is based on the fold difference in gene expression between the target and an internal standard gene[7]. In this study, the internal standard was *PEX4*, which encodes a ubiquitin-conjugating enzyme-like protein involved in transporting proteins targeted to the peroxisome[8].

Transgenic lines harbouring 35s::GFP.*At*IPCS2 and 3 were lost due to a viral infection, however, 35s::GFP.*At*IPCS1 transgenic plants were retrieved, and seeds from this line were grown and checked for expression of IPCS1 using real-time PCR and gene specific primers. As expected these transgenic lines had higher expression of *At*IPCS1 compared to Col0 (Figure 4-4).

However, under the confocal microscope, only low levels of fluorescence were observed, surrounded by dark regions that were indicative of dead cells (data not shown). These and the data above indicated that 35s::GFP.*At*IPCS constructs were not suitable for localisation analyses.
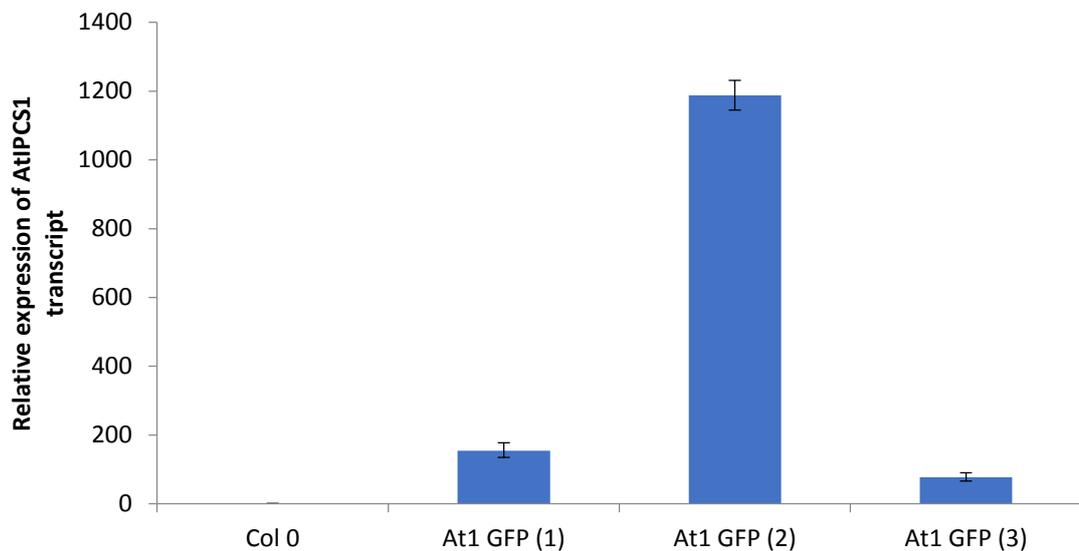
Figure 4-4: Relative transcript levels of *At*IPCS1 in *Arabidopsis* transgenic lines compared to wild-type: ten-week -old seedlings were grown on MS agar. Nomenclature of transgenic line is as follows: the first number identifies the IPCS isoform and the second number in brackets refers to the clonal line. Total RNA prepared from these seedlings and cDNA synthesised from 2 μg of total RNA, then used for transcript analysis using SYBR green RT-PCR. Results are an average with standard error calculated for three technical replicates in each clonal line.

4.2.2 *AtIPCS1-3* RNAi transgenic lines

In a similar manner, expression levels in RNAi *AtIPCS1, 2 and 3 Arabidopsis* knock-down lines were also quantified by real time PCR. To circumvent the detection of

accumulating *IPCS* mRNA in the cytoplasm destined for degradation, primers were designed to detect the 3' untranslated region (UTR) so that only the detection of mRNA destined for translation would be quantified.

Quantification of each *IPCS* isoform was probed with three specific primers pairs for each isoform to establish how the knock-down of one of the isoform would affect the expression pattern of the other (Figure 4-5). *AtIPCS1* transcript levels were 57 and 77% relative to Col0 for At1 RNAi 5 and 6 transgenic lines respectively. Interestingly *AtIPCS1* levels were also reduced in At2 RNAi 3, At2 RNAi 4 and At3 RNAi 2 (45, 85 and 46% respectively); these also exhibited the expected reduction of *AtIPCS2* transcript levels (61, 86 and 56% respectively). The transgenic lines At2 RNAi 1 and At2 RNAi 7 had *AtIPCS2* transcript levels of 53% and 61% relative to Col0, and At3 RNAi 2 exhibiting *AtIPCS2* levels of 56%.

Quantification of *AtIPCS3* transcripts in the transgenic lines, revealed that At2 RNAi 3 and At3 RNAi had significantly less *AtIPCS3* transcript compared to Col0 (19 and 45% respectively). Of all the transgenic lines tested, At2 RNAi 3 showed reduction in the expression of all three IPCS isoforms. This was of particular interest, for the further analyses of the global effects of sphingolipid dysfunction.

To understand the modulation of the three IPCS isoform in relation to one another, scatter graphs showing the transcript levels detected for one isoform plotted against the transcript levels detected for another isoform in the same transgenic lines were drawn with 85% reduction in *IPCS* transcript level set as the cut off value (Figure 4-6A-C). No correlation was found between *AtIPCS1* and *AtIPCS2* expression levels, however lines that had shown a decrease in *At*IPCS2 transcripts levels had an observable increase in *AtIPCS3* expression (Figure 4-6C – points highlighted in yellow).
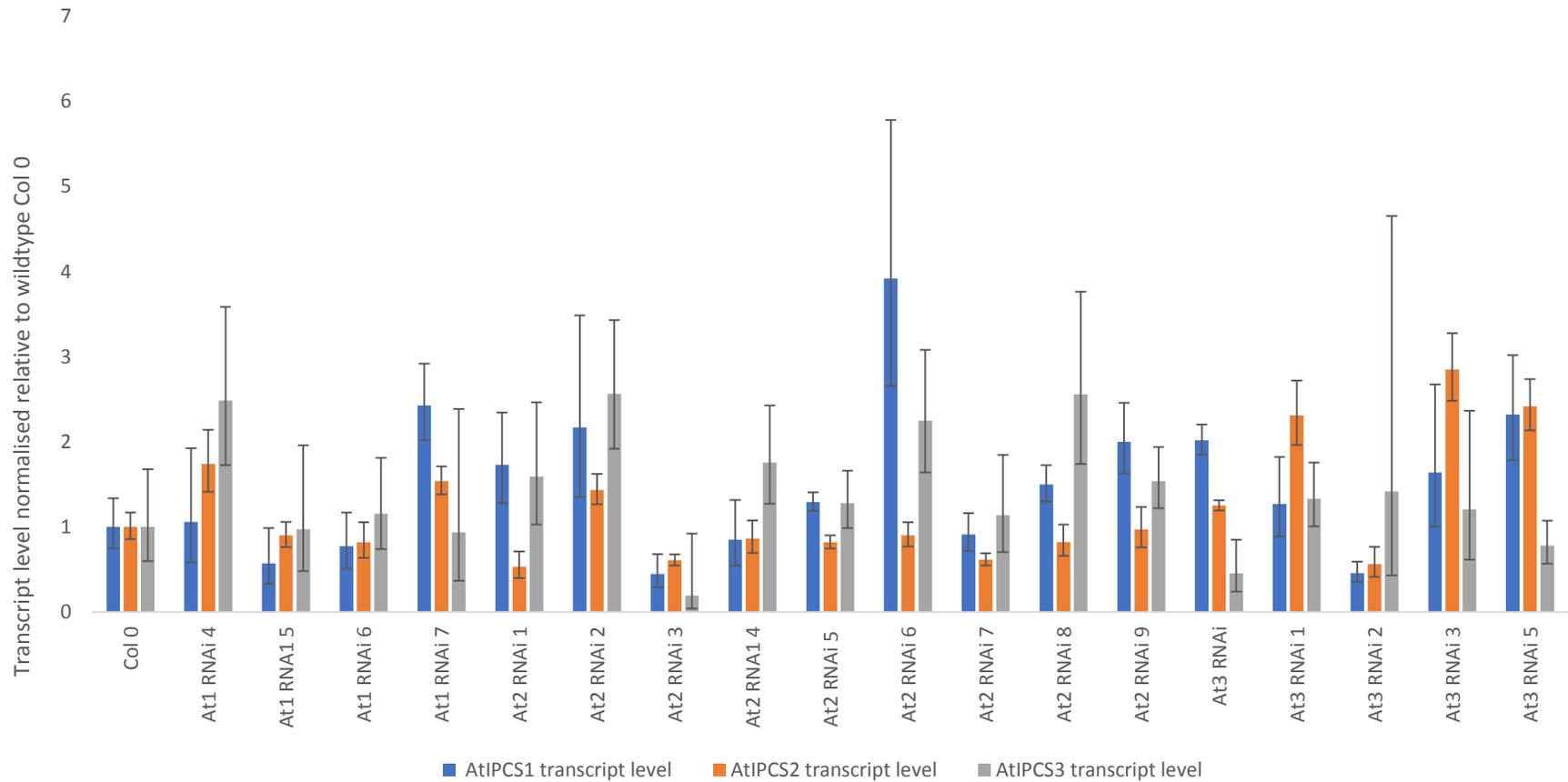
Figure 4-5: Relative transcript levels of *AtIPCS1-3* in *Arabidopsis* transgenic lines compared to wildtype using specific primers designed for each isoform and designed to detect the 3' UTR region of each isoform. Nomenclature of transgenic line is as follows: the first number identifies the IPCS isoform and the second number refers to the clonal line Total RNA was extracted from ten-day-old seedlings grown on MS agar and cDNA synthesised from 2 µg of total RNA. Transcript levels were analysed using SYBR green RT-PCR and are the average with standard error calculated for three technical replicates in each clonal line.
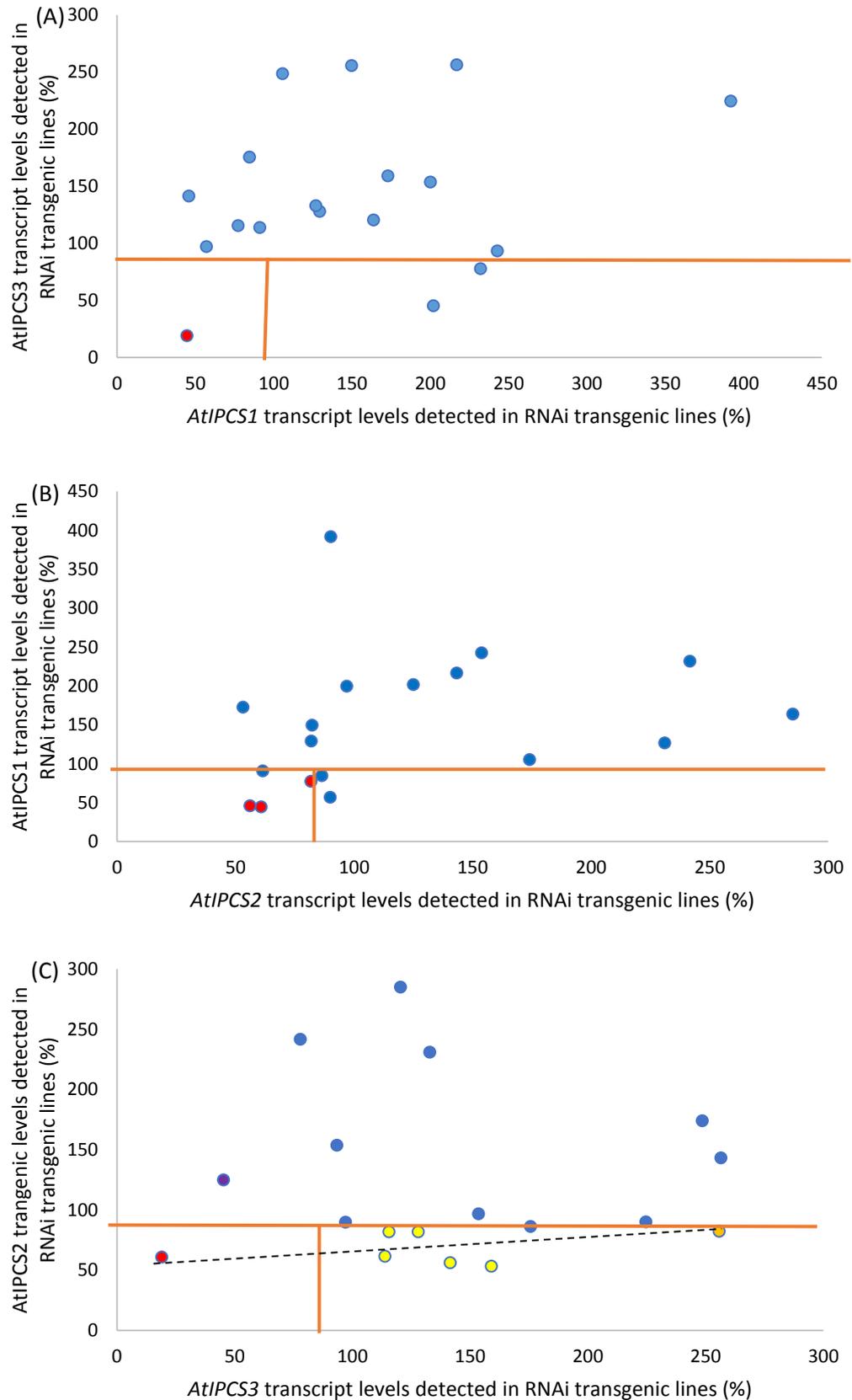
Figure 4-6: Percentage transcript levels of *AtIPCS1-3* in the transgenic lines relative to Col0 (A) *AtIPCS1* vs *AtIPCS3* levels (B) *AtIPCS2* vs *AtIPCS1* levels (C) *AtIPCS3* vs *AtIPCS2*. Highlighted in red are transgenic lines that showed significant reduction in transcript level for both IPCS isoforms and highlighted in yellow are transgenic lines which showed a relationship between the differential expression of *AtIPCS2* and *AtIPCS3*.

Based on the initial results, three transgenic lines were selected for each isoform of *At*IPCS which showed a reduction in transcript levels: At1RNAi 5, At1 RNAi 6, At1 RNAi 7, At2 RNAi 1, At2 RNAi 3, At2 RNAi 7, At3 RNAi, At3 RNAi 2 and At3 RNAi 5. To confirm genetic stability of the reduced transcript levels of *IPCS*, the transgenic lines were grown to obtain the T3 generation and the transcript levels of *At*IPCS1, 2 and 3 quantified (Figure 4-7).

Quantification of *AtIPCS1* showed that the transgenic lines which had previously shown a decrease in transcripts of 57% and 77% for At1 RNAi 5 and At1 RNAi6 (Figure 4-9) had increased to 106% and 93% respectively. This indicated that RNAi of *AtIPCS1* transcripts was not genetically stable. In contrast, the reduction in *AtIPCS2* transcripts proved to be genetically stable with similar levels observed for the selected transgenic RNAi and At3 RNAi 5 had 26% and 16% of *AtIPCS3* wild-type transcript, compared to
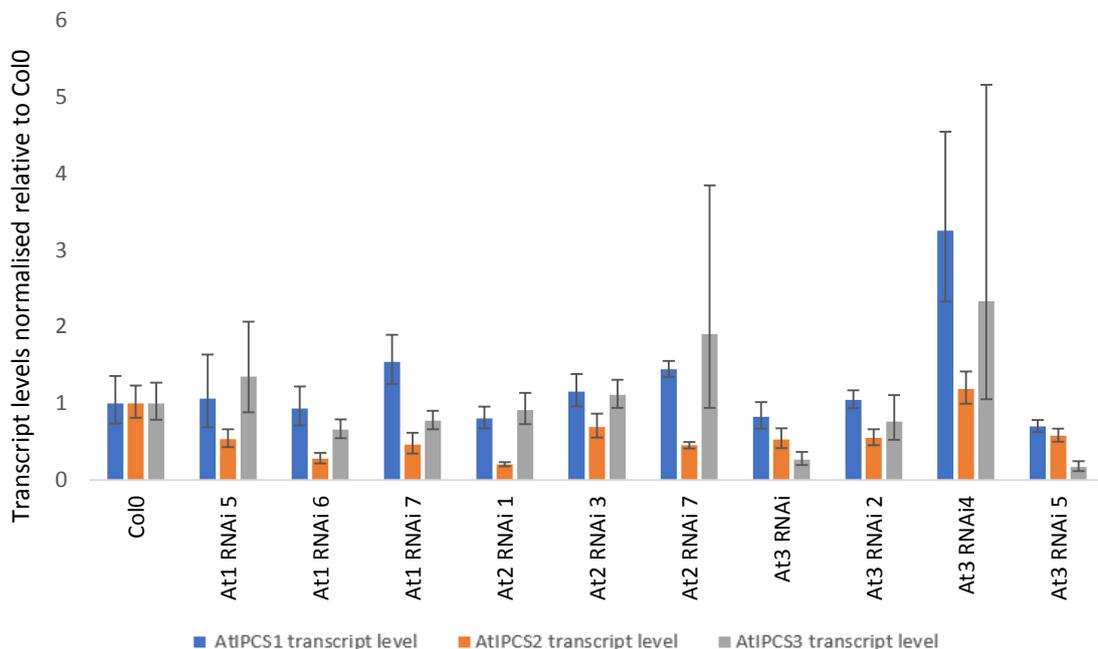
Figure 4-7: Relative transcript levels of *AtIPCS1, 2* and *3* in the selected *Arabidopsis* transgenic lines from the T3 generation compared to wildtype. Nomenclature of transgenic line is as follows: the first number identifies the IPCS isoform and the second number refers to the clonal line. Total RNA was extracted from ten-day-old seedlings grown on MS agar and cDNA synthesised from 2 µg of total RNA. Transcript levels were analysed using SYBR green RT-PCR and are the average with standard error calculated for three technical replicates in each clonal line.

lines: At2 RNAi 1 (20%), At2 RNAi 3 (69%) and At2 RNAi 7 (45%). Similarly, At3 the earlier qPCR result which yielded 45% and 78%. Interestingly, all lines, except At3 RNAi 4, showed a reduction in *AtIPCS2* mRNA levels, 20-69% of wild-type.

It was observed that there was large variation between the biological duplicate results, in some part is owed to pipetting errors, but also to the fact that the analysed lines comprise of pools of different clonal lines consisting of hemizygous, homozygous and wildtype expression of the different isoforms of *AtIPCS*. To isolate a stable homozygous line, T3 transgenic clonal lines were collected separately and a progeny test carried out. The progeny test is based on the knowledge that the transformed T0 generation are heterozygous for the introduced transgene and are kanamycin resistance[9].

Following self-pollination of the T1 generation, a mixture of homozygous (kanamycin resistant), hemizygous (kanamycin resistant) and wildtype (kanamycin sensitive) T2 generation is formed with a segregation ratio of 3:1 for transgenic lines harbouring the transgene. The same pattern of antibiotic selection also occurs in the T3 generation but with a segregation ratio of 5:3. If the parent of the T3 generation is homozygous, all seeds grown on kanamycin plates will survive as all of the progeny of a homozygous parent would be homozygous and therefore kanamycin sensitive. In contrast parent lines that are hemizygous will have some progeny that are kanamycin sensitive and not viable.

Of the 60 independent transgenic lines tested, 6 transgenic lines tested positive as homozygous lines with all seeds grown on kanamycin medium surviving. transgenic lines had reduced expression for at least one isoform of *AtIPCS* (Figure 4-8). However, RNAi effectively reduced transcript levels when targeted against *AtIPCS2* and *3*; *At*IPCS1 RNAi did not lead to a decrease in *AtIPCS1* mRNA (Figure 4-8) which suggests that it has an essential role linked to the early developmental stages in plants.
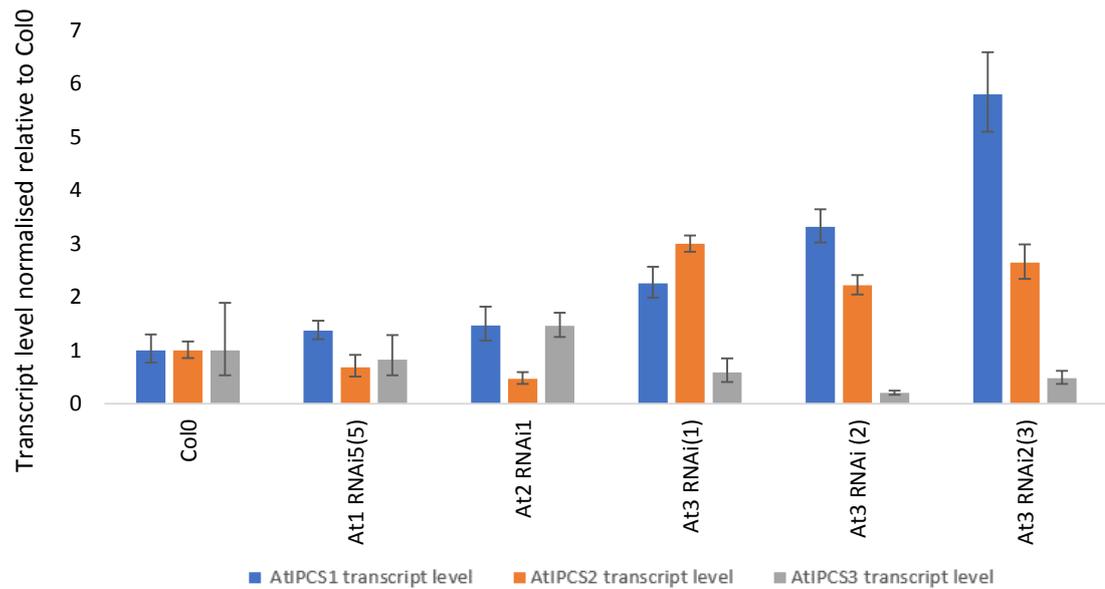
Figure 4-8: Relative transcript level of *AtIPCS1, 2* and *3* n T3 homozygous *Arabidopsis* transgenic lines compared to wildtype using specific primers designed to detect the 3' UTR in each isoform. Nomenclature of transgenic line is as follows: the first number identifies the IPCS isoform and the second refers to the clonal line Total RNA was extracted from ten-day-old seedlings grown on MS agar and cDNA synthesised from 2 µg of total RNA. Transcript levels were analysed using SYBR green RT-PCR and are the average with standard error calculated for three technical replicates in each clonal line.

4.2.3 *AtIPCS1-3* overexpression transgenic lines

Quantitation of transgenic lines over expressing *At*IPCS1-3 was carried out in the same manner. It was found that over-expression of *At*IPCS1 and *At*IPCS3 was highly efficient with very high levels of transcript detected, hundreds of folds higher than Col0. In contrast, over-expressing *At*IPCS2 had transcript levels only 5-7-fold higher than Col0 (Figure 4-9A-C).

The lower relative over-expression of *AtIPCS2* in the transgenic lines, was despite the fact that this isoform is the most highly expressed in *Arabidopsis*. These data suggested that the level of *AtIPCS2* expression is tightly regulated, which could indicate that this is because *AtIPCS2* functions as a negative regulator of PCD[5], so it is vital for plant survival that the expression levels are tightly regulated. Therefore, transgenic lines with higher levels of *AtIPCS2* transcripts may not be viable, resulting in a selection pressure for lines
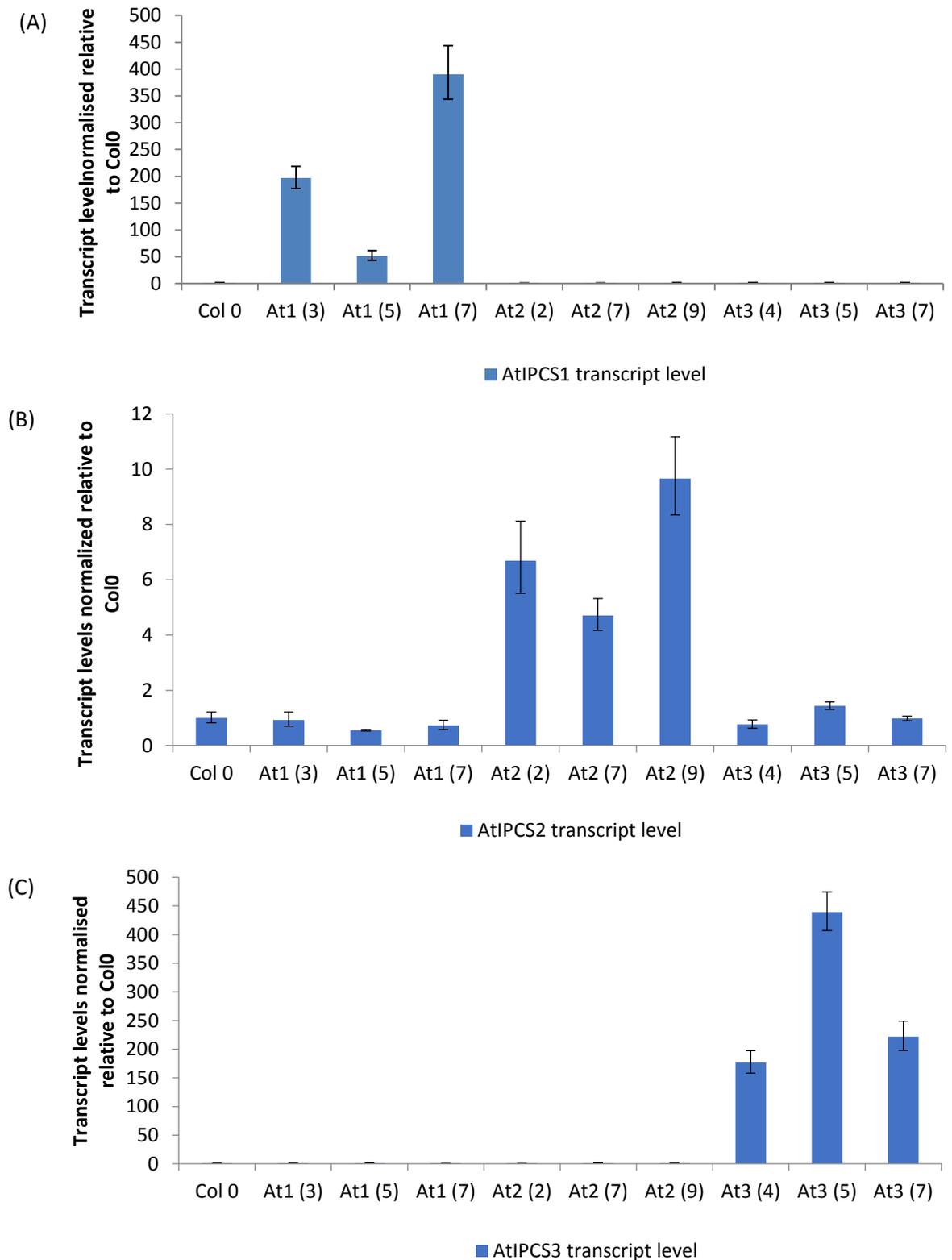
Figure 4-9: Quantification of *AtIPCS1, 2* and 3 transcripts in respective overexpression transgenic lines *Arabidopsis* compared to wildtype using specific primers designed to detect each *IPCS* isoform (A) *AtIPCS1* (B) *AtIPCS2* (C) *AtIPCS3*. Nomenclature of transgenic line is as follows: the first number identifies the IPCS isoform and the second number in brackets refers to the clonal line. Total RNA was extracted from ten-day-old seedlings grown on MS agar and cDNA synthesised from 2 µg of total RNA. Transcript levels were analysed using SYBR green RT-PCR and are the average with standard error calculated for three technical replicates in each clonal line.

with lower over-expression levels of *AtIPCS2*. These data also showed that the over-expression is target specific, with the transcript levels of the other isoforms found to be relatively unaffected (Figure 4-9A-C).

4.3 Phenotypic analysis of transgenic lines

Phenotypic analyses of the transgenic lines revealed no obvious phenotype at 10 days, however one transgenic line At3 RNAi 4 had seedlings displaying chlorotic leaves (Figure 4-10). When these were transferred to soil there was no growth.
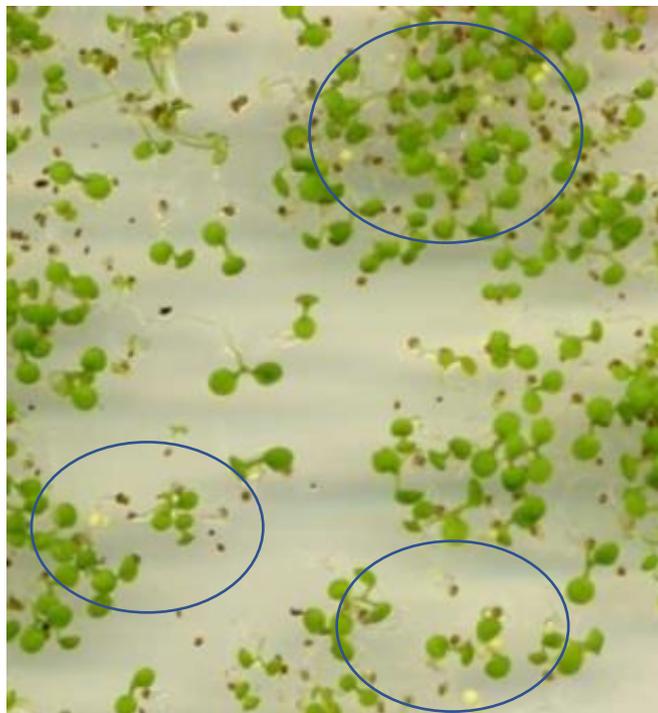
Figure 4-10: Phenotypic analysis of 13 days old at At3 RNAi 4 transgenic lines grown on MS agar; some seedlings have chlorotic leaves which are highlighted in blue circles.

Work carried out by Boyes et al. outlined the principle growth stages of Col0 from seed imbibition to senescence including the average period in days for each stage for plate-based and soil-based phenotypic analysis[10]. Phenotypic analysis of the transgenic lines compared to Col0 showed that RNAi transgenic lines had earlier inflorescence emergence
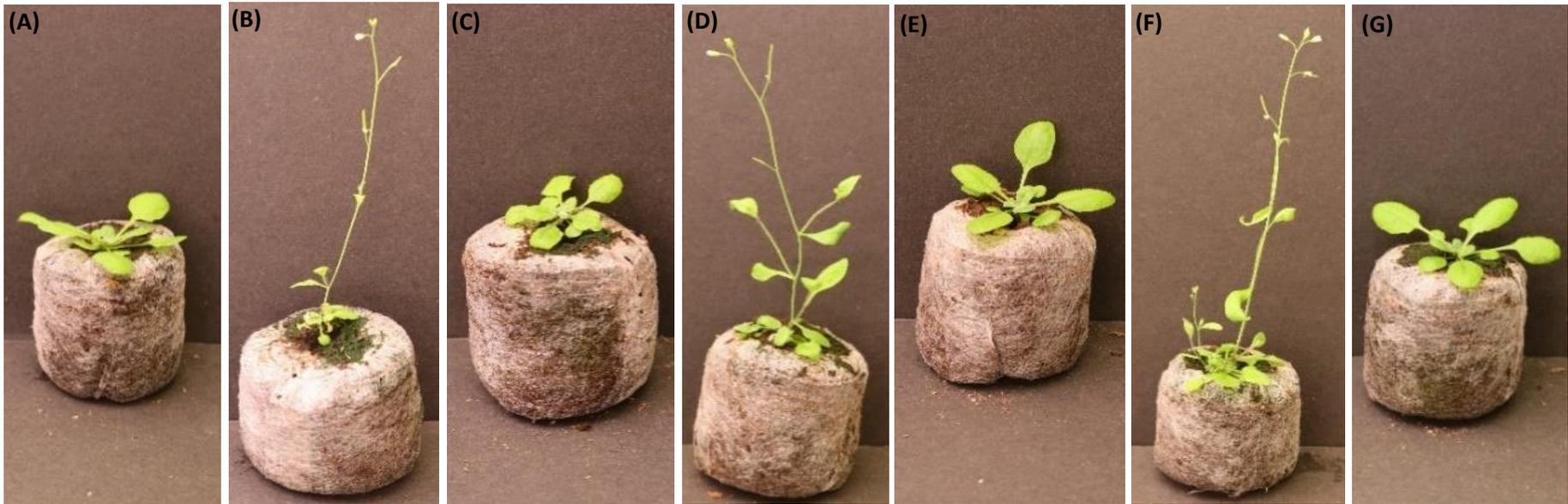
Figure 4-11: Phenotypic analyses of transgenic lines at 31 days (A) Col0 (B) At1 RNAi 7 (4) (C) At1 over exp 7 (D) At2 RNAi 1(3) (E) At2 over exp 9 (F) At3 RNAi (2) (G) At3 over exp 5.

Figure 4-12: Phenotypic analyses of transgenic lines at 44 days (A and B) Col0 which had flowered with varying degrees in height (C) At1 RNAi 7 (4) (D) At1 over exp 7 (E) At2 RNAi 1(3) (F) At2 over exp 9 (G) At3 RNAi (2) (F) At3 over exp 5.

and flowering compared to over expressing transgenic lines and Col0. RNAi transgenic lines flowering was first captured at 31 days (Figure 4-11) whilst Col0 and the over-expressors at the same time point still showed no signs of flowering. IPCS overexpressing lines flowered (40 days) slightly earlier than Col0 (44 days) and were found to growing quicker (Figure 4-12).

This phenotype of early flowering has been shown to be a phenotype of plants grown in cold condition for extended periods of time and is triggered by a decrease in Flowering Locus C (FLC) transcript levels[11]. This process termed vernalisation, is not an inheritable trait and requires low temperatures exposure in successive generations to promote flowering[12]. In addition there is an FLC independent vernalisation pathway that promotes early flowering[13], as does the autonomous pathway, which also controls flowering time in plants and requires the repression of FLC expression for early flowering[14]. The early flowering observed in RNAi transgenic lines are inheritable, implicating IPCS has a role to play in accelerating plant flowering time in *Arabidopsis*.
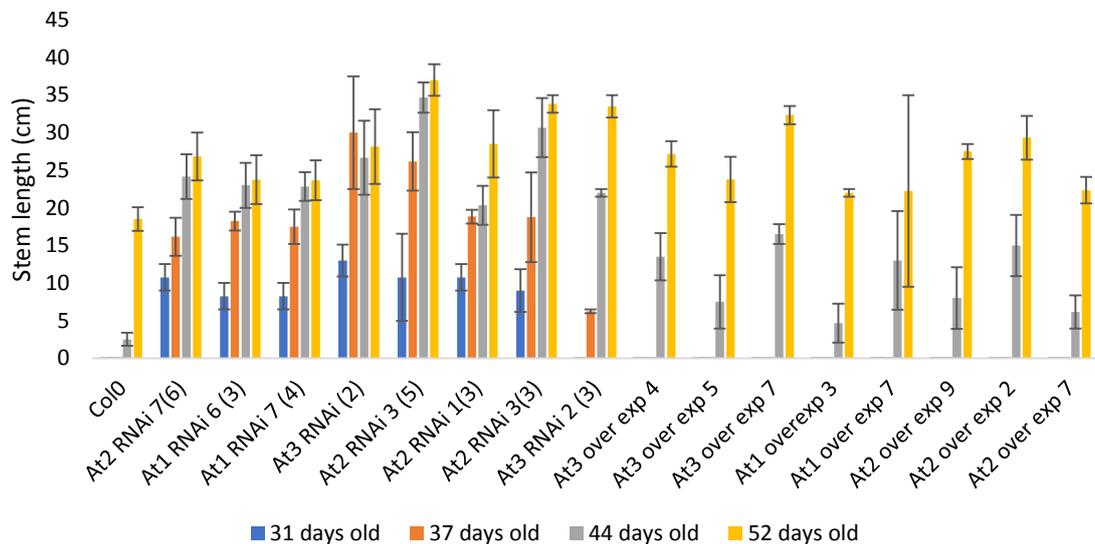
Figure 4-13: Tracking the growth of transgenic lines over 52 days; RNAi lines stem emergence (31 days) was earlier than overexpression transgenic lines and Col0 (44 days).

Another phenotype observed in the RNAi lines were much smaller rosette leaves and weak stems which needed support to stand upright, compared to Col0 and over-

expressing lines which had bigger rosette leaves and strong stems that were upright (Figure 4-12).

In addition, some RNAi transgenic lines showed early onset of senescence at 44 days in comparison to over expressing transgenic lines and Col0 (Figure 4-12). A comparison of plant growth showed that the RNAi lines exhibited initial acceleration in flowering relative to the overexpressing lines and Col0 (Figure 4-13). However, after late inflorescence emergence the overexpressing lines grew at a quicker rate and surpassed the RNAi lines in height.

4.4 Transcript quantification of *At*IPCS homozygous lines from European seed bank

In an alternative approach, homozygous tDNA mutants of *At*IPCS1-3 were sought from the European seed bank, but only mutants for *At*IPCS2 (N696590) and *At*IPCS3 (N674600) were available. Transcript levels of *At*IPCS1 and *At*IPCS2 were quantified in the homozygous t-DNA mutant lines (Figure 4-14).
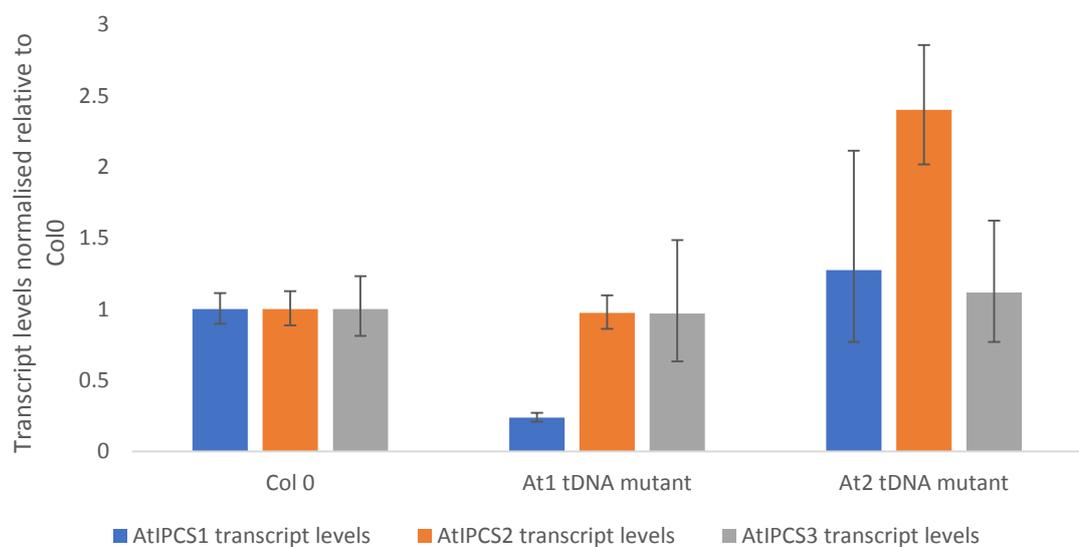
Figure 4-14: Transcript level of *AtIPCS1*, *2* and *3* in homozygous t-DNA mutants *At*IPCS1 and *At*IPCS2.

The *At*IPCS1 t-DNA mutant showed a significant reduction in *At*IPCS1 transcripts, however, the *At*IPCS2 t-DNA mutant exhibited high levels of *At*IPCS2 mRNA. In *At*IPCS2 t-DNA mutants, the insertion was found to be located upstream of the *At*IPCS2 promoter, so disruption of expression would be far from certain. Indeed, studies of multiple t-DNA mutants have shown that insertion in the intron or exon regions are more effective than insertions occurring in the promoter or regions upstream of the promoter[15]. In comparison, the t-DNA insert for *At*IPCS1 was located in the intron sequence. Although this would have been spliced out, the significant reduction in *At*IPCS1 transcript levels suggested that transcripts were produced with decreased efficiency.

4.5 Conclusion

*At*IPCS1-3 RNAi and over expressing transgenic lines were created and verified by qPCR. Homozygous lines were verified by progeny testing for RNAi lines. Phenotypic analyses of the transgenic lines yielded some interesting results, implicating IPCS in plant development particularly with regards to the robustness of the stem. In addition, IPCS was found to be implicated in vernalisation as indicated by the early flowering phenotype in RNAi *At*IPCS1-3 knock down transgenic lines, which is genetically stable and hereditary. The overexpressing lines also exhibited an early flowering phenotype compared to Col0 but was much later compared to the RNAi transgenic lines, indicating that reduction in *At*IPCS transcripts has a more pronounced effect on the flowering time of plants. In generally the RNAi lines were unhealthy, with earlier onset of senescence and withering brown leaves when compared to over-expressers and Col0. These observations show that IPCS is necessary in plants development. For a deeper understanding of which biological pathways are affected by a change in IPCS expression, transcriptomic analyses were carried out, which will be discussed in the next chapter.

4.6 References

1. Earley, K. W.; Haag, J. R.; Pontes, O.; Opper, K.; Juehne, T.; Song, K.; Pikaard, C. S., Gateway- compatible vectors for plant functional genomics and proteomics. *The Plant Journal* **2006,** *45* (4), 616-629.

2. Hamilton, A. J.; Baulcombe, D. C., A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **1999,** *286* (5441), 950-952.

3. Hammond, S. M.; Bernstein, E.; Beach, D.; Hannon, G. J., An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *nature* **2000,** *404* (6775), 293.

4. Karimi, M.; Inzé, D.; Depicker, A., GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in plant science* **2002,** *7* (5), 193-195.

5. Wang, W.; Yang, X.; Tangchaiburana, S.; Ndeh, R.; Markham, J. E.; Tsegaye, Y.; Dunn, T. M.; Wang, G.-L.; Bellizzi, M.; Parsons, J. F., An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in Arabidopsis. *The Plant Cell* **2008,** *20* (11), 3163-3179.

6. Wee, E. G.-T.; Sherrier, D. J.; Prime, T. A.; Dupree, P., Targeting of active sialyltransferase to the plant Golgi apparatus. *The Plant Cell Online* **1998,** *10* (10), 1759-1768.

7. Livak, K. J.; Schmittgen, T. D., Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods* **2001,** *25* (4), 402-408.

8. Zolman, B. K.; Monroe-Augustus, M.; Silva, I. D.; Bartel, B., Identification and functional characterization of Arabidopsis PEROXIN4 and the interacting protein PEROXIN22. *The Plant Cell* **2005,** *17* (12), 3422-3435.

9. Passricha, N.; Saifi, S.; Khatodia, S.; Tuteja, N., Assessing zygosity in progeny of transgenic plants: current methods and perspectives. *Journal of Biological Methods* **2016,** *3* (3), e46.

10. Boyes, D. C.; Zayed, A. M.; Ascenzi, R.; McCaskill, A. J.; Hoffman, N. E.; Davis, K. R.; Görlach, J., Growth stage–based phenotypic analysis of Arabidopsis a model for high throughput functional genomics in plants. *The Plant Cell* **2001,** *13* (7), 1499-1510.

11. Simpson, G. G.; Dean, C., Arabidopsis, the Rosetta stone of flowering time? *Science* **2002,** *296* (5566), 285-289.

12. Sheldon, C. C.; Rouse, D. T.; Finnegan, E. J.; Peacock, W. J.; Dennis, E. S., The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences* **2000,** *97* (7), 3753-3758.

13. Michaels, S. D.; Amasino, R. M., Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. *The Plant Cell* **2001,** *13* (4), 935-941.

14. Simpson, G. G., The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. *Current opinion in plant biology* **2004,** *7* (5), 570-574.

15. Wang, Y. H., How effective is T-DNA insertional mutagenesis in Arabidopsis? *Journal of Biochemical Technology* **2008,** *1* (1), 11-20.

CHAPTER 5

RNASEQ ANALYSES OF TRANSENIC LINES

5.1 Chapter synopsis

In this Chapter, RNASeq was carried out on transgenic lines overexpressing each isoform of IPCS in *Arabidopsis thaliana*. Analyses of the changes in the transcriptome revealed the role of *At*IPCS in plant defence response to biotic and abiotic stress. In addition, RNASeq analyses were also carried out on an *AtIPCS1* t-DNA insertion mutant to compare changes in the transcriptome with the over-expressor lines, thereby uncovering the effects of the dose dependent expression of *AtIPCS* and the global function of IPCS in plants.

5.2 Analyses of the transcriptome

Following the verification of the transgenic lines (Chapter 4) RNASeq was carried out on 10-day old over-expressor lines and 10 day old *AtIPCS1* t-DNA mutant line at principal growth stage 1.02; the RNAi lines were not utilized for RNASeq due to the limited time available. Principle component analysis (PCA) was carried out to quantify the variance between data sets (levels of gene transcription in each plant cell) in different samples (technical and biological replicates).

A PCA plot was created that captured the variance in the transcription patterns in each sample, accounting for 73% of the variance identified. From this plot, it was clear to see that all technical replicates of each biological sample, apart from one *AtIPCS2* over-expressor, At2-over-exp-2, clustered closely together, demonstrating that the transcriptomic data from the technical replicates, are similar, as expected (Figure 5-1). The variability in the variance between technical replicates for At2 over-exp-9 cannot be explained. However, a sample-to-sample heat map showing the similarities and dissimilarities between samples, reveals that all technical replicates of a given biological

sample, including At2-over-exp-9 cluster together (Figure 5-2). This showed that the technical replicates of At2-over-exp-9 are actually quite similar, despite the differences in variance revealed in the PCA plot.
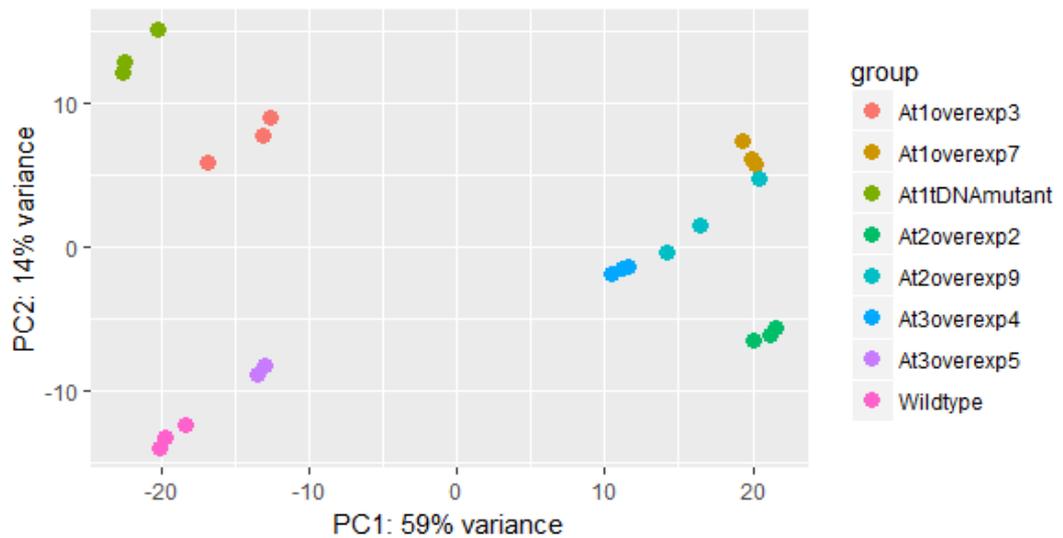
Figure 5-1: PCA plot of variance between transcriptome of technical and biological replicates for *At*IPCS1-3 over-expresser and *At*IPCS1 t-DNA mutant transgenic lines.
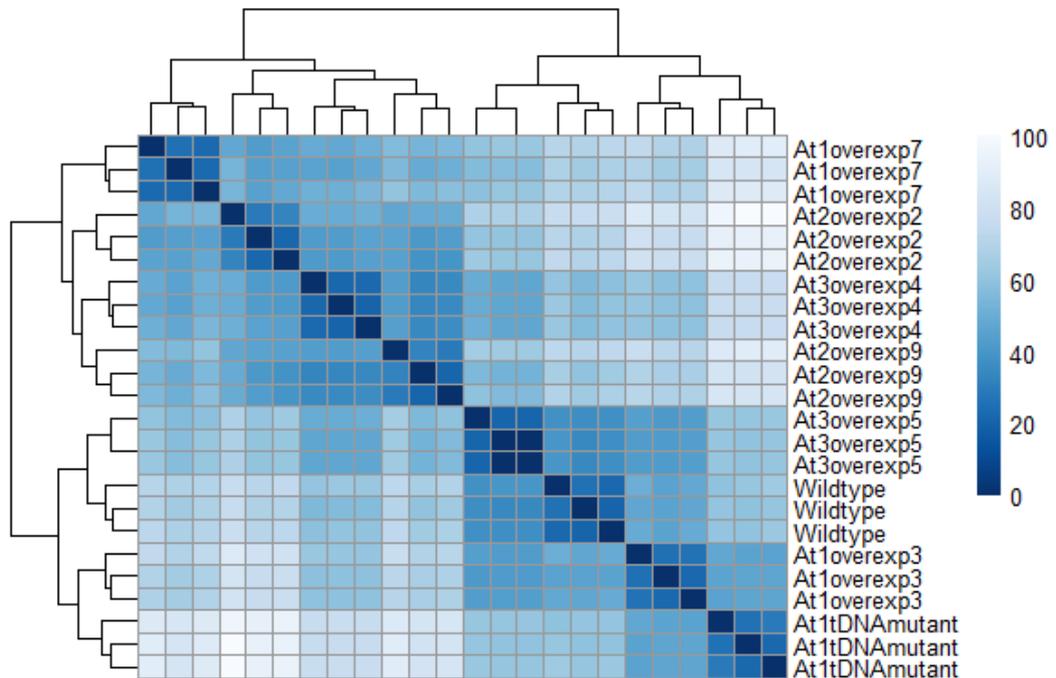
Figure 5-2: Heat map showing sample to sample distances for technical and biological replicates.

Functional characterisation of *Os*IPCS1-3 has been carried out, and has revealed that the

expression of each IPCS isoform is modulated in response to abiotic stress (drought, cold and salt) to different degrees, with each isoform showing an organ specific expression pattern[1]. *OsIPCS1* and *OsIPCS2* were found to be upregulated in response to cold and drought in the roots and stem, but specifically down-regulated in the leaves in response to cold. *OsIPCS3* expression was found to be upregulated in all tissues, in response to drought, cold and salt stress.

To date, there has been limited functional characterisation of the three IPCS isoforms in *A. thaliana*. Work done by Wang et al, has characterised *At*IPCS2 as a negative regulator of PCD[2] in plants, but little is known about *At*IPCS1 and *At*IPCS3. These two homologues are likely redundant in function, and exhibit tissue-specific differential expression, with *AtIPCS2* found to be highly expressed in all tissues, whilst *AtIPCS1* expression is $\log_2$ fold ≥ 2 lower in expression in all tissues and *AtIPCS3* is highly expressed in flowers, and stems comparable to *AtIPCS2* expression levels[3]. *At*IPCS3 shows a $\log_2$ fold ≥ 3 lower expression levels than *AtIPCS2* in roots, rosette leaves, cauline leaves, and siliques and is lower in expression in comparison to *AtIPCS1* in all tissues apart from in siliques, flowers and stems[3]. The differential expression of the three isoforms, could be an indication of as yet unidentified, tissue specific regulatory roles and modulation of plant response to stimuli, based on the specific expression pattern of each isoform.

RNASeq of two biological replicates for each IPCS isoform over-expressor lines was carried out, with Col0 as control. Genes that were identified as upregulated with a $\log_2$ fold difference in expression in both biological replicates, compared to Col0, were carried forward for further analyses. It was found that 275 genes were upregulated in the *AtIPCS2* over-expressors, 70 in *AtIPCS1* and 19 genes in the *AtIPCS3* over-expressors; of these 15 were found to be upregulated in response to the overexpression of all three isoforms (Figure 5-3A). Regarding the down regulated genes, again *At*IPCS2 over expressors

transgenic lines had the highest number showing a log$_2$ fold change ≥ 2 in expression, with 135 compared to 54 and 59 genes for *AtIPCS1* and *AtIPCS3* respectively (Figure 5-3B). Of these, 26 genes were downregulated in all three IPCS isoforms over-expressors.

The acquisition of high-throughput data from genomic and proteomic experiments, results in labour-intensive analyses to identify the involvement of a set of genes in a specific pathway. In order to achieve this, pathway analysis (also called enrichment analysis) is undertaken to identify the underlying biological function of these set of genes; this is referred to as the gene ontology[4]. The set of genes enriched under a specific GO are then compared to a set of genes associated with a specific biological pathway, and p-values generated by comparing the observed frequency of a GO term with the frequency expected by chance; GO annotation with a p-value ≤ 0.05 are considered to be enriched[5]. Singular enrichment analysis was carried out to categorise the genes identified, into the relevant biological pathways within which they function.
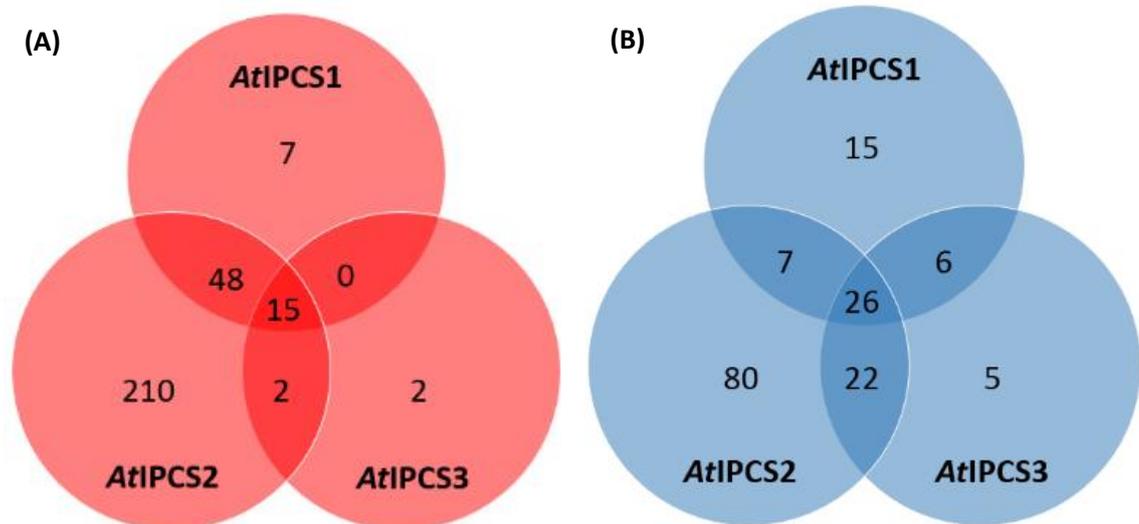
Figure 5-3: Venn diagrams showing the number of genes that were (A) upregulated (B) down regulated in the overexpression lines when compared to Col0.

5.3 Genes downregulated in response to the overexpression of *AtIPCS1, 2* and *3*

Analyses of genes that were downregulated and upregulated in response to the overexpression of *AtIPCS1, 2* and *3* was carried out using agriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php). Genes downregulated in both *AtIPCS1* over expressors (At1-over-exp-3 and At1-over-exp-7) revealed a significant enrichment ($p = 1.87e-12$) for genes annotated with the GO term GO:0006950 response to stress, with 43.4% (23/53) of the genes associated with this term compared to 6.14% in the *Arabidopsis* transcriptome (Table 5-1). Other GO terms with siginificant enrighment include response to biotic stimulus (13/53) and defence response (13/53), with response to salicylic acid (7/53) and jasmonic acid stimulus (5/530) also showing enrichment. Of the 23 genes enriched under resposne to stress, 13 were associated with defence response and 9 were found to be downregulated by a $\log_2$ fold ≥ 2 in at least one biological replicate (highlighed in bold in Table 5-2).

Of particular interest are genes that showed a dose dependent decrease associated with higher transcript levels of *AtIPCS1* in the At1 over-exp-7 transgenic line; these genes included *plant defensin 1.2B*, the salicylic acid inducible pathogenesis-related genes, *PR1* and *PR2*, and an *ethylene/jasmonate responsive element plant defensin*, a well characterized component of the defence network against pathogens[6] (Figure 5-2). *PR1* and *PR2*, and the *ethylene/jasmonate responsive element plant defensin*, showed a 2-fold decrease in transcript levels in At1 over-exp-7 transgenic lines, compared to At1 over-exp-3 which over expresses *AtIPCS1* at a lower level.

This implied that there is a link between the levels of *AtIPCS1* transcript and the expression of these genes. The roles of *PR* genes are well established in the plant response to biotic and abiotic stress; *PR2* and *PR5* expression level are elevated in response to plant

Table 5-1: GO enrichment of genes down regulated in response to the constitutive overexpression of *AtIPCS1*

GO term	Description	Number in input list	Number in Background	p-value
GO:0050896	Response to stimulus	32	4057	3.50E-18
GO:0006950	Response to stress	23	2320	1.50E-14
GO:0051707	Response to other organism	13	599	2.10E-12
GO:0009607	Response to biotic stimulus	13	638	4.60E-12
GO:0006952	Defence response	13	766	4.30E-11
GO:0051704	Multi-organism process	13	776	5.00E-11
GO:0009814	Defence response, incompatible interaction	7	143	1.80E-09
GO:0009751	Response to salicylic acid stimulus	7	200	1.70E-08
GO:0045087	Innate immune response	7	347	6.30E-07
GO:0009628	Response to abiotic stimulus	12	1471	7.60E-07
GO:0006955	Immune response	7	367	9.10E-07
GO:0002376	Immune system process	7	368	9.30E-07
GO:0009753	Response to jasmonic acid stimulus	5	215	1.50E-05
GO:0010033	Response to organic substance	10	1342	1.60E-05
GO:0042221	Response to chemical stimulus	12	2085	2.60E-05
GO:0009605	Response to external stimulus	5	429	0.00035
GO:0009266	Response to temperature stimulus	5	485	0.00062
GO:0009416	Response to light stimulus	5	596	0.0015
GO:0009314	Response to radiation	5	613	0.0017
GO:0009719	Response to endogenous stimulus	6	1068	0.0038
GO:0005618	Cell wall	6	403	2.30E-05
GO:0030312	External encapsulating structure	6	407	2.40E-05

Table 5-2: Genes down regulated by the constitutive overexpression of *AtIPCS1*, highlighted in red are genes responsive to JA signalling, in blue are genes responsive to SA signalling, in purple are genes responsive to SA and JA and in bold are genes with log$_2$ fold change $\geq$ 2

Gene ID	Gene annotation	At1-over-exp-3		At1-over-exp-7	
		Log$_2$ fold change	p-value	Log$_2$ fold change	p-value
AT1G31580	ECS1 (CARBOXYPEPTIDASE)	-1.6	1.60E-73	-1.1	3.71E-37
AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3)	**-1.8**	**6.29E-17**	**-4.0**	**1.33E-53**
AT2G26020	PLANT DEFENSIN 1.2B	**-1.9**	**8.15E-18**	**-5.0**	**1.44E-66**
AT5G52310	LTI78 (LOW-TEMPERATURE-INDUCED 78)	-1.1	6.25E-52	-1.5	5.40E-102
AT2G14610	PR1 (PATHOGENESIS-RELATED GENE 1)	**-3.4**	**1.83E-172**	**-7.3**	**4.08E-238**
AT1G75040	PATHOGENESIS-RELATED GENE 5	-1.4	3.27E-22	-1.7	2.13E-33
AT1G06160	OCTADECANOID-RESPONSIVE ARABIDOPSIS AP2/ERF 59	-1.0	1.14E-10	-1.4	9.92E-17
AT1G06040	STO (SALT TOLERANCE)	-1.1	2.13E-28	-1.5	1.78E-49
AT2G14560	LURP1 (LATE UPREGULATED IN RESPONSE TO HYALOPERON OSPORA PARASITICA)	**-2.7**	**5.62E-112**	**-3.4**	**1.13E-162**
AT5G10140	FLC (FLOWERING LOCUS C)	-1.2	2.18E-09	-1.5	4.16E-14
AT1G19670	CLH1 (CORONATINE-INDUCED PROTEIN 1)	-1.7	4.25E-75	-1.8	9.20E-91
AT4G14400	ACD6 (ACCELERATED CELL DEATH 6)	-1.7	3.11E-73	-1.5	5.03E-60
AT5G57380	VIN3 (VERNALIZATION INSENSITIVE 3)	-1.1	0.002783851	-1.1	0.001246826
AT5G37260	RVE2 (REVEILLE 2)	-1.4	2.65E-46	-1.2	1.18E-38
AT3G57260	PATHOGENESIS-RELATED PROTEIN 2	**-1.9**	**1.17E-30**	**-5.3**	**4.64E-102**
AT3G56400	WRKY70	-1.0	3.44E-33	-1.8	3.96E-89
AT3G45860	RECEPTOR LIKE PROTEIN-KINASE	**-1.5**	**6.64E-09**	**-2.9**	**3.18E-24**
AT5G44420	ETHYLENE AND JASMONATE RESPONSIVE PLANT DEFENSIN	**-2.7**	**1.49E-54**	**-6.1**	**2.23E-122**
AT4G12490	AZI3, LPID TRANSFER PROTEIN (LTP)	**-1.3**	**7.93E-05**	**-4.6**	**4.17E-43**
AT4G24350	PHOSPHORYLASE FAMILY PROTEIN	-1.2	8.48E-38	-1.2	1.08E-41
AT3G25760	AOC1 (ALLENE OXIDE CYCLASE 1)	-1.2	6.06E-40	-1.7	4.85E-72
AT2G40750	WRKY54	-1.3	4.41E-14	-1.2	1.08E-13
AT1G09080	ATP binding BIP3	**-1.1**	**7.65E-06**	**-2.5**	**6.79E-21**

infection with the pathogen, *Erwinia amylovora*[7]. PR1 expression is elevated in response to SA[8], ethylene and methyl jasmonate[9], and all three PR genes are upregulated in response to drought[10]. The down regulation of these PR genes (*PR1*, *PR2* and *PR5*) in response to the constitutive overexpression of *AtIPCS1* implies that *At*IPCS1 is part of the plant defence signaling network.

Interestingly, between the two *AtIPCS1 over-expressors,* a 3-fold difference in transcript levels of tyrosine amino transferase 3 (*TAT3*) was observed; this protein has been shown to increase in expression in response to wounding, and was hypothesized to have increased activity following pathogen attack in order to meet the high demand for prenylquinones that neutralize reactive oxygen species (ROS)[11]. Another gene that showed a three-fold decrease in transcript levels was the systematic acquired resistance (SAR) specific protein, *azelaic acid induced 1* (*AZI1*) which interacts with DIR1 (Defective in Induced Resistance), forming a complex that contributes to the biosynthesis and accumulation of glycerol-3-phosphate (G3P), a molecule necessary for the spreading of SAR signaling to distal tissues in plants[12, 13].

The results from the transcriptomic analyses of genes downregulated in response to the overexpression of *AtIPCS1*, suggested that it is a negative regulator of the plant defence response to pathogens. This affects both the SA and JA signaling pathways, which counters the well-established JA/ET and SA antagonistic relationship found in most plant defence response against pathogens[14, 15]. Studies have shown that SA signaling is usually instigated in response to biotrophic and hemibiotropic pathogens, whilst JA signaling is activated by necrotrophic pathogens, with both signaling pathways acting as antagonists[16]. However, work by Yoashioka *et al.* showed that *crp22* (constitutive expresser of PR genes) mutants demonstrate a synergistic interplay between JA and SA signaling, with higher levels of SA, and constitutive expression of the *PR* genes and the

ethylene responsive defensin gene, *PDF1.2*[17]. The downregulation of genes involved in SA and JA plant defence signaling in response to the constitutive overexpression of *At*IPCS1 indicates that IPCS in plant defence against pathogens (necrotrophs, bio and hemibiotrophs) by utilizing both signaling networks.

The WRKY transcriptional factors also have well characterized roles in plant defence[18]; *WRKY 54* and *70* were shown to be downregulated in response to *AtIPCS1* overexpression (Table 5-2), and at low levels of expression are coupled to a decrease in SA biosynthesis after pathogen response[19]. In a separate study *wrky54/wrky70* double mutants have been shown to result in enhanced tolerance to osmotic stress[20] and to function as negative regulators of leaf senescence[21], thereby show casing their roles in plant response to biotic and abiotic stress. This affirms the role of *At*IPCS1 as a modulator of plant response to stress.

A decrease in gene expression of the Flowering Locus C (*FLC*) and vernalisation insensitive 3 (*VIN3*) genes, was also observed on *AtIPCS1* overexpression. This could explain, in part, the phenotypic differences (early flowering) observed between the over-expressors and Col0. The over-expressors exhibited a slightly earlier flowering (4 days) relative to Col0. When plants are exposed to cold, VIN3 expression is induced and forms a heterodimer with vernalization 5 (VRN5) which is necessary for the repression of FLC, and results in early flowering[22]; it therefore follows that a decrease in VIN3 expression, should result in a decrease of the repression of *FLC* (up regulation of *FLC*) leading to either normal or delayed flowering. This is not what was observed from the transcriptomic analysis. Instead, a decrease in the transcript for *VIN3* and *FLC* was seen, which would suggest that there are other components which down regulate FLC, independent of *VIN3*. The down regulation of *FLC* combined with the slightly early flowering phenotype obser-

Table 5-3: GO enrichment of genes down regulated in response to the constitutive overexpression of *AtlPCS2*

GO term	Description	Number in input list	Number in background	p-value
GO:0050896	Response to stimulus	72	4057	2.80E-34
GO:0006950	Response to stress	46	2320	2.20E-22
GO:0042221	Response to chemical stimulus	39	2085	5.60E-18
GO:0010033	Response to organic substance	32	1342	1.20E-17
GO:0051707	Response to other organisms	21	599	6.90E-15
GO:0009607	Response to biotic stimulus	21	638	2.30E-14
GO:0051704	Multi-organism process	22	776	9.60E-14
GO:0006952	Defence response	21	766	6.90E-13
GO:0009628	Response to abiotic stimulus	27	1471	2.60E-12
GO:0009719	Response to endogenous stimulus	21	1068	2.80E-10
GO:0009753	Response to jasmonic acid stimulus	11	215	5.70E-10
GO:0009416	Response to light stimulus	15	596	5.10E-09
GO:0009314	Response to radiation	15	613	7.40E-09
GO:0048511	Rhythmic process	7	75	1.80E-08
GO:0007623	Circadian rhythm	7	75	1.80E-08
GO:0009751	Response to salicylic acid stimulus	9	200	6.40E-08
GO:0009605	Response to external stimulus	12	429	6.20E-08
GO:0009723	Response to ethylene stimulus	9	199	6.10E-08
GO:0045087	Innate immune response	11	347	6.70E-08
GO:0010200	Response to chitin	8	151	1.10E-07
GO:0002376	Immune system process	11	368	1.20E-07
GO:0006955	Immune response	11	367	1.20E-07
GO:0009743	Response to carbohydrate stimulus	9	240	2.80E-07
GO:0051716	Cellular response to stimulus	15	840	4.00E-07
GO:0009725	Response to hormone stimulus	16	982	5.30E-07
GO:0009611	Response to wounding	8	197	7.40E-07
GO:0009617	Response to bacterium	8	247	3.80E-06
GO:0010035	Response to inorganic substance	8	279	9.00E-06

Table 5-4: Genes down regulated by the constitutive overexpression of *AtIPCS2*, highlighted in red are genes responsive to JA signalling, in blue are genes responsive to SA signalling, in purple are genes responsive to SA and JA and in bold are genes with log$_2$ fold change ≥ 2

Gene ID	Gene annotation	At2-over-exp-2		At2-over-exp-9	
		Log$_2$ fold change	p-value	Log$_2$ fold-change	p-value
AT5G47220	ERF2 (ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 2)	-1.7	4.58E-25	-1.1	7.33E-15
AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3)	**-2.6**	**3.03E-28**	**-2.3**	**4.60E-26**
AT1G16420	MC8 (METACASPASE 8)	-1.6	3.84E-06	-1.2	0.000470364
AT4G12480	PEARLI 1; LIPID BINDING A PUTATIVE LIPID TRANSFER PROTEIN	**-1.9**	**1.94E-31**	**-5.1**	**7.83E-172**
AT1G78290	SERINE/THREONINE PROTEIN KINASE	-1.8	8.25E-47	-2.0	1.39E-54
AT4G37990	ELI3-2 (ELICITOR-ACTIVATED GENE 3-2)	**-2.3**	**7.52E-12**	**-2.3**	**2.18E-12**
AT3G45290	MLO3 (MILDEW RESISTANCE LOCUS O 3)	-1.1	5.16E-09	-1.3	7.16E-12
AT2G26020	PDF1.2B (PLANT DEFENSIN 1.2B)	**-4.9**	**1.47E-60**	**-4.9**	**3.31E-62**
AT2G32680	RLP23 (RECEPTOR LIKE PROTEIN 23)	**-2.2**	**1.35E-13**	**-2.3**	**3.16E-15**
AT2G26560	PLA2A (PHOSPHOLIPASE A 2A)	**-1.4**	**1.57E-26**	**-2.6**	**1.57E-78**
AT2G14610	PR1 (PATHOGENESIS-RELATED GENE 1)	**-7.3**	**1.05E-206**	**-7.9**	**4.60E-217**
AT2G46830	CCA1 (CIRCADIAN CLOCK ASSOCIATED 1)	**-2.0**	**3.12E-12**	**-1.1**	**1.04E-05**
AT1G06160	ORA59 (OCTADECANOID-RESPONSIVE ARABIDOPSIS AP2/ERF 59)	**-2.5**	**9.89E-37**	**-1.8**	**1.67E-24**
AT3G49620	DIN11 (DARK INDUCIBLE 11)	**-3.2**	**7.04E-106**	**-4.1**	**5.37E-147**
AT1G06040	STO (SALT TOLERANCE)	-1.3	2.23E-38	-1.1	8.10E-27
AT1G48000	MYB112 (MYB DOMAIN PROTEIN 112)	**-2.0**	**1.97E-10**	**-1.2**	**5.54E-05**
AT2G14560	LURP1 (LATE UPREGULATED IN RESPONSE TO HYALOPERONOSPORA PARASITICA)	**-4.1**	**5.67E-173**	**-3.3**	**1.62E-149**
AT1G22770	GI (GIGANTEA)	**-2.4**	**6.55E-94**	**-2.3**	**8.70E-91**
AT5G10140	FLC (FLOWERING LOCUS C)	-1.4	5.23E-11	-1.3	4.55E-10
AT3G1550	ANAC055 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 55)	-1.3	2.83E-20	-1.5	4.16E-28
AT1G76930	EXT4 (EXTENSIN 4)	**-1.2**	**1.21E-28**	**-3.1**	**3.77E-175**
AT5G44420	AN ETHYLENE- AND JASMONATE-RESPONSIVE PLANT DEFENSIN	**-6.0**	**3.77E-111**	**-6.1**	**3.17E-118**
AT2G23680	STRESS-RESPONSIVE PROTEIN	-1.1	3.58E-14	-1.3	1.09E-18
AT1G61120	TPS04 (TERPENE SYNTHASE 04)	**-2.6**	**2.43E-15**	**-1.7**	**5.55E-08**
AT1G75040	PR5 (PATHOGENESIS-RELATED GENE 5)	-1.5	1.00E-25	-1.8	5.50E-34
AT1G52890	ANAC019 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 19)	-1.2	2.90E-05	-1.5	5.04E-07
AT3G57260	PR2 (PATHOGENESIS-RELATED GENE 2)	**-4.8**	**2.99E-86**	**-4.4**	**4.08E-87**
AT1G18330	EPR1 (EARLY-PHYTOCHROME-RESPONSIVE1)	-1.6	3.85E-29	-1.9	4.62E-40
AT1G78410	VQ MOTIF-CONTAINING PROTEIN	-1.8	9.42E-08	-1.6	1.24E-06
AT5G19880	PEROXIDASE	-1.4	0.000104883	-1.0	0.004183815

Table 5-4 continued: Genes down regulated by the constitutive overexpression of *AtIPCS2*, highlighted in red are genes responsive to JA signalling, in blue are genes responsive to SA signalling, in purple are genes responsive to SA and JA and in bold are genes with log$_2$ fold-change $\geq$ 2

		At2-over-exp-2		At2-over-exp 9	
Gene ID	Gene annotation	Log$_2$ fold-change	p-value	Log$_2$ fold-change	p-value
AT2G18050	HIS1-3 (HISTONE H1-3)	**-2.5**	**2.90E-96**	**-1.7**	**3.90E-55**
AT3G04720	PR4 (PATHOGENESIS-RELATED 4)	-1.6	7.55E-69	-1.5	1.20E-55
AT2G43510	TI1; SERINE-TYPE ENDOPEPTIDASE INHIBITOR	**-2.5**	**6.34E-132**	**-3.9**	**2.98E-227**
AT3G45860	RECEPTOR-LIKE PROTEIN KINASE	**-2.9**	**3.03E-22**	**-3.0**	**8.20E-25**
AT3G51660	MACROPHAGE MIGRATION INHIBITORY FACTOR FAMILY PROTEIN	-1.2	1.40E-16	-1.1	6.63E-15
AT1G71030	MYBL2 (ARABIDOPSIS MYB-LIKE 2)	-1.2	1.04E-29	-1.7	4.26E-55
AT1G57630	DISEASE RESISTANCE PROTEIN (TIR CLASS)	**-1.5**	**8.34E-09**	**-2.5**	**4.73E-18**
AT5G59780	MYB59 (MYB DOMAIN PROTEIN 59)	-1.2	6.92E-29	-1.7	9.77E-56
AT4G12490	LIPID TRANSFER PROTEIN (LTP)	**-2.8**	**8.38E-18**	**-4.7**	**1.76E-45**
AT4G23210	PROTEIN KINASE FAMILY PROTEIN ENCODES A CYSTEINE-RICH RECEPTOR-LIKE KINASE (CRK13)	**-1.7**	**8.20E-24**	**-3.0**	**1.31E-52**
AT1G09080	ATP BINDING BIP3	-1.6	4.53E-10	-1.6	6.04E-11
AT4G02520	GSTF2 (GLUTATHIONE S-TRANSFERASE PHI 2)	**-1.2**	**2.82E-06**	**-2.9**	**4.44E-28**
AT2G25000	WRKY60	-1.2	7.64E-11	-1.4	9.79E-15
AT2G02990	RNS1 (RIBONUCLEASE 1)	-1.1	1.29E-08	-1.7	1.82E-18

ved in the over-expressers is in line with the current understanding of early flowering in plants which requires reduced *FLC* transcripts[23].

Analyses of genes downregulated in *AtIPCS2* overexpressing lines, demonstrated a significant enrichment for GO terms, GO:0050896 response to stimulus (72/134), GO:0006950 response to stress (46/134). That response 34.3% of the genes associated with response to stress, which make up 6.1% of the *Arabidopsis* transcriptome (Table 5-3). Other GO terms showing significant enrichment include, GO: 0042221 (response to chemical stimulus), GO: 0010033 (response to organic substances), and GO: 0051707 (response to other organisms). Of the 46 genes enriched under the GO term response to stress, 17 genes were associated with defence response. Interestingly, genes that showed a dose dependent decrease associated with higher transcript levels of *AtIPCS1* (*plant defensin 1.2B*, *PR1*, *PR2*, and *ethylene/jasmonate responsive element plant defensin*) did not show the same dose dependent decrease in relation to higher *AtIPCS2* transcripts (Table 5-4) despite the 1/3 difference in *AtIPCS2* transcripts between At2 over-exp-2 and At2 over-exp-9 (higher levels of *AtIPCS2* transcripts in At2 over- exp-9). However, it is immediately clear that the plant defence system is more sensitive to expression levels of *AtIPCS2* than *AtIPCS1*; despite relatively low levels of *AtIPCS2* transcript increase (6-10 fold) compared to *AtIPCS1* transcripts (200-390 fold). The relatively low increase in *AtIPCS2* transcript level compared to *AtIPCS1* in the overexpression lines reflects *AtIPCS2* being the most expressed IPCS in plants compared to *AtIPCS1* and *AtIPCS3* in Col0. Genes identified to be down regulated in response to *AtIPCS2* overexpression, are comparable to the fold-change in expression of genes also observed to be down regulated in the transgenic line with the highest level of *AtIPCS1* transcripts (At1 over exp 7).

Analyses of the genes down regulated in *AtIPCS2* over-expressers associated with defence response are higher in numbers, with some identified that were not seen in the

AtIPCS1 over-expressors analyses. These include, *ERF2* which induces SAR via the JA signalling pathway resulting in resistance to *F. oxysporum*[24]; *MC8*,which promotes programmed cell death in response to stress (hydrogen peroxide and UV light), and *MLO3* which confers resistance to powdery mildew in barley[25], and is probably in part, responsible for the reduced susceptibility to powdery mildew of *erh1*/Col-gl mutants[2]. Also downregulated was *pEARL1*, which is a *cis*-jasmone-inducible gene found to be upregulated in response to the enhancement of plant defence against herbivorous insects[26]; *RLP23* which is down regulated in plant SAR response[27], *PLA2A* which shows elevated expression in response to *Botrytis cinerea* and *Pseudomonas syringae* with increased susceptibility to these pathogens observed in *Arabidopsis*[28], *EXT4* which is upregulated in response to hydrogen peroxide and the production of superoxide anion up on treatment with digitonin[29]. *TPS04* which is responsible for the synthesis of the trepenoid, (E,E)-4,8,12-trimethyl-1,3,7,11-tridecatetraene, and has been found to be produced as a by-product of activated JA signaling[30]. *ANAC019* which is a regulator of JA signalling defence response[31] and lines overexpressing this gene were found to have increase drought tolerance[32], *GSTF2* which exhibits elevated transcript levels when *Arabidopsis* is inoculated with *P. syringae* and is inducible by SA and ethylene[33], and *CRK13* which also plays a role in plant resistance to *P. syringae* by the accumulation of SA[34]. Also, among this list of down regulated genes identified in *AtIPCS2* over-expressers but not *AtIPCS1* over-expressers are *peroxidase*, *TIR* class of disease resistance proteins and *PR4*; which all play a role in plant defence response. From these analyses, it is clear that *AtIPCS2* expression affects a wider network of genes involved in plant defence responses compared to *AtIPCS1* which might be a result of its higher expression pattern in all tissues of *Arabidopsis*[3].

A dose dependent decrease was found in the transcription of genes identified in response to the increase in *AtIPCS2* transcript levels in the transgenic lines. Interestingly 6 genes,

CCA1, *ORA59*, *LURP1*, *MYB112*, *TPS04* and *HIS1-3*, show a $\log_2$ fold decrease in expression with an inverse relationship to *AtIPCS2* expression. The rational as to why this has occurred is unclear. A higher number of genes are also identified for *AtIPCS2* over-expressors implicated in plant response to cold stress (*Gigantea*)[35], drought stress (*serine/threonine protein kinase* and *HIS1-3*)[36, 37], salt and osmotic stress (*WRKY60*)[38].

It was found that genes involved in the regulation of circadian control of light responsive genes were down regulated. These included *CCA1* which encodes a transcription factor involved in the induction of the light-harvesting chlorophyll *a/b* gene associated with photosystem II[39], and *EPR1* which is inducible by irradiation with red light and regulated by phytochrome A and B[40]. Overexpression of both genes is associated with delayed flowering, and in the case of *EPR1*, also results in cotyledon opening[40] whilst *CCA1* overexpression results in plants with longer hypocotyls and the disruption of the circadian clock which affects several genes[39]. These genes are necessary for plant development, and in response to *AtIPCS2* overexpression are down regulated and probably contributed to the slightly earlier flowering (4 days) phenotype of these lines compared to Col0. In addition, the cold response gene *GI* was down regulated. The gene product regulates circadian rhythm and photoperiodic flowering by upregulating the expression of CONSTANS and FLOWERING LOCUS T (FT) which results in early flowering[41].

Also, down regulated was *FLOWERING LOCUS C* (*FLC*) which is a MADS box protein and the regulator of flowering in the autonomous pathway[42] and in response to vernalisation[43]; down regulation of *FLC* promotes early flowering in both cases. Floral transition occurs not only through vernalisation and the autonomous pathway via the repression of *FLC*, but also through the photoperiodic pathway, GA pathway and the FLC independent vernalisation pathway[44] through a host of different intermediary genes and transcriptional regulators (Figure 5-4). A decrease in *FLC* transcript results in an early fl-

owering phenotype whilst plants which have delayed flowering show an increase in *FLC* transcript[45, 46].
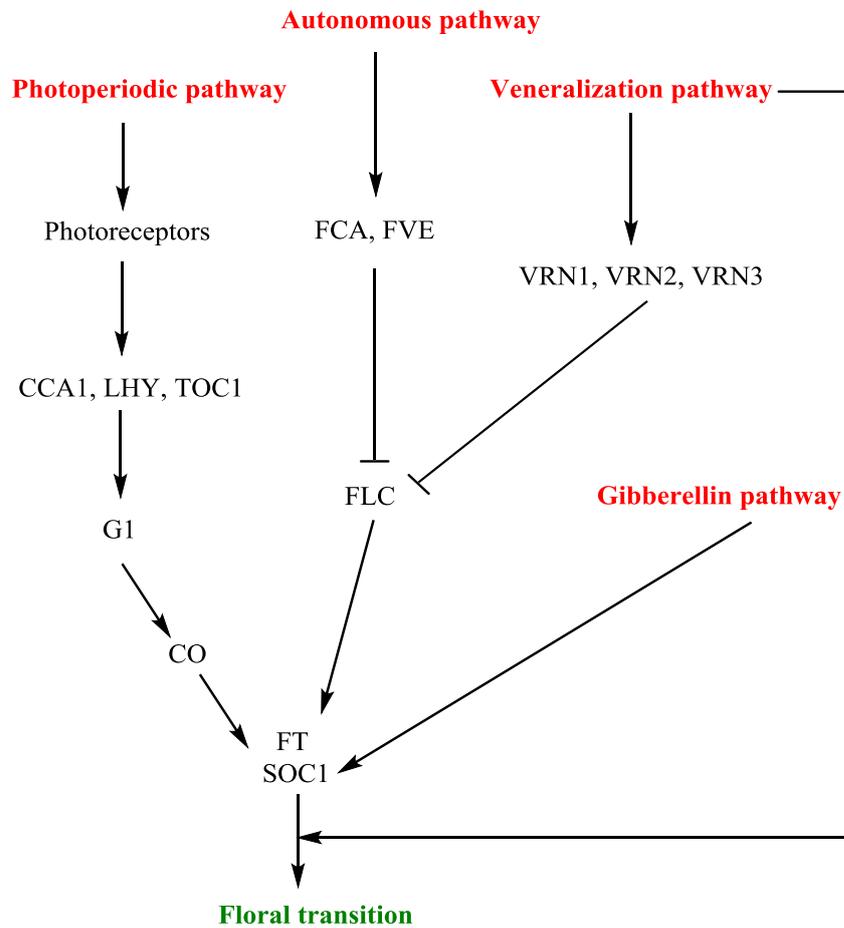
Figure 5-4: Schematic of the different pathways that trigger floral transition. The vernalisation and autonomous pathway rely on the repression of *FLC* which in turn results in the expression of *FT* and *SOC1* induce the expression of genes responsible for the development of floral organs. Repression of FLC by the expression of any of the following upstream genes: FCA, FVE, VRN1, VRN2 and VRN3 results in floral transition. The photoperiodic pathway and the gibberellin pathway although do not require the repression of *FLC* to induce floral transition, have genes upstream of the respective pathways that can induce the expression of *FT* and *SOC1* resulting in the transition of a plants from a vegetative phase to a reproductive phase. The *FLC* independent vernalisation pathway also does not require *FLC* to induce floral transition and is believed to occur through the induction of MADS box transcription factors which are closely related to *FLC*.

In this study the over-expressor lines have a phenotype of early flowering and shows a $\log_2$ fold ≥ 1 reduction in *FLC* which suggests that *At*IPCS may have a role in plant response to vernalisation. As stated in chapter 4, the early flowering phenotype observed

Table 5-5: GO enrichment of genes down regulated in response to the constitutive overexpression of *AtIPCS3*

GO term	Description	Number in input list	Number in Background	p-value
GO:0050896	Response to stimulus	33	4057	2.70E-17
GO:0006950	Response to stress	24	2320	2.40E-14
GO:0009628	Response to abiotic stimulus	15	1471	6.00E-09
GO:0006952	Defence response	11	766	2.90E-08
GO:0051707	Response to other organism	10	599	3.40E-08
GO:0009607	Response to biotic stimulus	10	638	6.00E-08
GO:0051704	Multi-organism process	10	776	3.60E-07
GO:0009416	Response to light stimulus	8	596	4.40E-06
GO:0009314	Response to radiation	8	613	5.40E-06
GO:0009639	Response to red or far red light	5	210	2.20E-05
GO:0042221	Response to chemical stimulus	12	2085	8.00E-05
GO:0010035	Response to inorganic substance	5	279	8.30E-05
GO:0009791	Post-embryonic development	7	705	0.00012
GO:0009266	Response to temperature stimulus	6	485	0.00012
GO:0009409	Response to cold	5	328	0.00017
GO:0010033	Response to organic substance	8	1342	0.0011
GO:0030528	Transcription regulator activity	12	2417	0.00032
GO:0003700	Transcription factor activity	10	2173	0.0019

Table 5-6: Genes down regulated by the constitutive overexpression of *AtIPCS3*, highlighted in red are genes responsive to JA signalling, in blue are genes responsive to SA signalling, in purple are genes responsive to SA and JA and in bold are genes with log$_2$ fold-change ≥ 2

Gene ID	Gene annotation	At3 over-exp-4		At3 over-exp-5	
		Log$_2$ fold-change	p-value	Log$_2$ fold-change	p-value
AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3)	**-1.8**	**6.10E-18**	**-2.3**	**1.65E-24**
AT1G16420	MC8 (METACASPASE 8)	-1.0	0.002428624	-1.2	0.000589589
AT1G78290	SERINE/THREONINE PROTEIN KINASE	-1.0	2.73E-20	-1.3	1.51E-29
AT2G26020	PDF1.2B (PLANT DEFENSIN 1.2B)	**-2.6**	**3.65E-30**	**-3.4**	**3.34E-40**
AT5G65080	MAF5 (MADS AFFECTING FLOWERING 5)	-1.5	2.11E-07	-1.7	5.09E-08
AT2G14610	PR1 (PATHOGENESIS-RELATED GENE 1)	**-4.6**	**1.09E-252**	**-4.9**	**2.85E-235**
AT1G06040	STO (SALT TOLERANCE)	-1.0	6.04E-26	-1.5	2.41E-48
AT4G12480	PEARLI 1; LIPID BINDING A PUTATIVE LIPID TRANSFER PROTEIN	**-1.2**	**8.99E-13**	**-2.0**	**3.20E-32**
AT1G48000	MYB112 (MYB DOMAIN PROTEIN 112)	-1.2	6.45E-05	-1.7	3.23E-08
AT2G14560	LURP1 (LATE UPREGULATED IN RESPONSE TO HYALOPERONOSPORA PARASITICA)	**-3.2**	**9.88E-155**	**-2.2**	**5.56E-82**
AT1G61120	TPS04 (TERPENE SYNTHASE 04)	-1.6	4.61E-07	-1.8	6.72E-09
AT5G10140	FLC (FLOWERING LOCUS C)	-1.4	8.80E-12	-1.1	3.25E-08
AT1G19670	CLH1 (CORONATINE-INDUCED PROTEIN 1)	-1.1	1.14E-34	-1.2	2.36E-40
AT1G22770	GI (GIGANTEA)	**-1.3**	**4.07E-33**	**-2.4**	**1.27E-94**
AT3G57260	PR2 (PATHOGENESIS-RELATED GENE 2)	**-2.0**	**2.97E-36**	**-2.2**	**8.40E-38**
AT1G78410	VQ MOTIF-CONTAINING PROTEIN	-1.1	0.00072986	-1.2	0.000163912
AT2G43510	TI1; SERINE-TYPE ENDOPEPTIDASE INHIBITOR	-1.1	2.90E-37	-1.8	9.93E-77
AT3G45860	RECEPTOR-LIKE PROTEIN KINASE	-1.5	4.12E-09	-1.7	4.81E-10
AT5G44420	PDF1.2A (PLANT DEFENSIN 1.2A)	**-2.7**	**1.34E-55**	**-3.6**	**7.78E-78**
AT4G37990	ELI3-2 (ELICITOR-ACTIVATED GENE 3-2)	-1.8	3.22E-08	-1.6	6.77E-07
AT4G12490	LIPID TRANSFER PROTEIN (LTP)	**-1.9**	**1.32E-08**	**-2.6**	**1.46E-15**
AT2G40750	WRKY54	-1.1	5.31E-11	-1.2	1.02E-11
AT4G02520	GSTF2 (GLUTATHIONE S-TRANSFERASE PHI 2)	-1.4	3.33E-08	-1.6	3.75E-10
AT1G18330	EPR1 (EARLY-PHYTOCHROME-RESPONSIVE1)	-1.3	1.98E-23	-1.6	6.68E-30

is inherited, which is contrary to the current understanding of early flowering in plants. This is based on reduced *FLC* expression, with the levels of *FLC* returning to normal levels in subsequent generations and is reduced in plants undergoing vernalisation by exposure to cold temperatures[43].

Down regulation of *GI* would result in delayed flowering compared to Col0 which was not observed. The down regulation of these genes points to a link between IPCS expression and circadian control of light responsive genes which may be a result of MYB transcription factor binding sites, which were identified upstream of *At*IPCS1, 2 and 3 promoter in chapter 2; this class of transcription factors are known to control plant development and plant response to abiotic and biotic stress[47].

Genes down regulated in response to the overexpression of *AtIPCS3* were found to be have a significant enrichment of GO term GO: 0050896 (response to stimulus) with 56% (33/59) of genes associated with this term compared to 10.7% in the *Arabidopsis* transcriptome. Other GO terms showing significant enrichment include GO:0006950 (response to stress), GO:0009628 (response to abiotic stimulus) and GO:0006952 (defence response) (Table 5-5). A dose dependent down regulation of genes identified in response to the overexpression of *At*IPCS3 was observed for *TAT3*, *PDF 1.2B*, *pEARL1*, *LURP1*, *G1*, *PDF1.2A* and *LPT*. Most of the genes identified as modulated in *At*IPCS3 over-expressors, had also been identified as responsive to *At*IPCS1 and *At*IPCS2 overexpression. It is clear, that there is a clustering of specific GO terms for genes down-regulated in *AtIPCS1*, *2* and *3* transgenic lines; these are mainly connected with the response to biotic and abiotic stress.

LURP1, which is upregulated in response to plant infection with oomycete *Hyaloperonospora parasitica* and required for basal defence against this fungal

pathogen[48], shows an inverse relationship in terms of expression, in response to *AtIPCS2* and *AtIPCS3* transcript levels (Table 5-4 and Table 5-6). This is different from the positive relationship found between *LURP1* down regulation and *AtIPCS1* expression (Table 5-2), the reasons for this differential relationship is unclear. Analyses of genes down regulated in response to the expression of all three *AtIPCS* isoforms was also carried out (Figure 5-7) and revealed 26 genes, of which 14 came under the GO term: 0050896 *(*response to stimulus), 7 genes under GO term: 0006950 *(*defence response) and 6 genes under GO term: 0009607 (response to biotic stress) with all genes identified under the GO term response to stress already identified under the GO term response to stimulus. Likewise, genes identified under GO term defence response were found to be also enriched under response to biotic stimulus.

Most of the genes enriched under response to stress for over-expressors of the three IPCS isoforms were expected as they had been identified in the analysis of genes enriched under GO term response to stimulus (Table 5-8). However, some genes were identified that were not grouped under response to stimulus but had been identified in the analyses of all three IPCS isoform over-expressors. These includes the glutamate receptor, *GLR2.7*, which encodes a functional channel involved in the allocation of Ca^{2+} ions to different calcium sinks for adaptation to K^+ and Na^+ stress and for normal development in plants[49], and *BDA1* which functions downstream of the SNC2 (Suppressor of NPR1, Constitutive2) receptor and has been shown to be a regulator of plant immunity. Loss of function of *BDA1* results in constitutive activated cell death and enhanced susceptibility to *P.syringae*[50]. *HYH*, a positive regulator of blue light specific induction of gene expression and plant development (shown to inhibit hypocotyl growth)[51], and *AIG1*, a GTP binding protein identified in *Arabidopsis* and found to confer resistance to plants infected with *Pseudomonas syringae* pv. *Maculicola* strain carrying the AvrRpt2 effector[52].

Table 5-7: GO enrichment of genes down regulated in response to the constitutive overexpression of *AtIPCS1*, *2* and *3*

GO term	Description	Number in input list	Number in Background	p-value
GO:0050896	Response to stimulus	14	4057	7.60E-08
GO:0006952	Defence response	7	766	6.80E-07
GO:0006950	Response to stress	10	2320	1.70E-06
GO:0051707	Response to other organism	6	599	2.90E-06
GO:0009607	Response to biotic stimulus	6	638	4.10E-06
GO:0051704	Multi-organism process	6	776	1.20E-05
GO:0009628	Response to abiotic stimulus	5	1471	0.003
GO:0042221	Response to chemical stimulus	5	2085	0.013

Table 5-8: Genes down regulated by the constitutive overexpression of *AtIPCS1*, *2* and *3* respectively

Genes down regulated in all overexpressing IPCS transgenic lines and enriched under GO term :0050896 (Response to stimulus)	
Gene ID	**Gene annotation**
AT2G29120	GLR2.7 (GLUTAMATE RECEPTOR)
AT2G14610	PR1 (PATHOGENESIS-RELATED GENE 1)
AT5G54610	BDA1, ANKYRIN PROTEIN
AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3);
AT3G45860	RECEPTOR-LIKE PROTEIN KINASE
AT5G44420	ETHYLENE- AND JASMONATE-RESPONSIVE PLANT DEFENSIN
AT3G57260	PR2 (PATHOGENESIS-RELATED GENE 2)
AT2G26020	PDF1.2B (PLANT DEFENSIN 1.2B)
AT1G06040	STO (SALT TOLERANCE)
AT2G14560	LURP1 (LATE UPREGULATED IN RESPONSE TO HYALOPERONOSPORA PARASITICA)
AT5G10140	FLC (FLOWERING LOCUS C)
AT3G17609	HYH (HY5-HOMOLOG)
AT4G12490	LIPID TRANSFER PROTEIN (LTP)
AT1G33960	AIG1 (AVRRPT2-INDUCED GENE 1)
Genes down regulated in all overexpressing IPCS transgenic lines and enriched under GO term: 0006950 (Defence response)	
AT5G52310	LTI78 (LOW-TEMPERATURE-INDUCED 78)
AT5G57380	VIN3 (VERNALIZATION INSENSITIVE 3)
AT5G37260	RVE2 (REVEILLE 2)
AT4G24350	PHOSPHORYLASE FAMILY PROTEIN
AT3G56400	WRKY70
AT3G25760	AOC1 (ALLENE OXIDE CYCLASE 1)
AT4G14400	ACD6 (ACCELERATED CELL DEATH 6)

The transcription factor *WKRY 70*, which has been shown to be capable of integrating the SA and JA signalling pathways in plant defence against pathogens[53], was also identified as downregulated in response to the overexpression of all IPCS isoforms , as was the SA inducible gene *ACD6* which is a positive regulator of plant defence against pathogens [54]. From these analyses, the role of *At*IPCS1-3 in plant defence against pathogens has been established. The expression of *AtIPCS1, 2* and *3* affects both the SA and JA signalling defence response, and so it is possible to conclude that the overexpression of *AtIPCS* isoforms, would make plants more susceptible to a wide range of pathogens.

5.5 Genes upregulated in response to the overexpression of *AtIPCS1*, *2* and *3*

Genes that were upregulated in response to *AtIPCS1* overexpression were also analysed to determine other pathways affected. Significant enrichment (p = 9.10E-27) for genes annotated with the GO term GO:0015979 (photosynthesis), with 26.1% (18/69) of the genes associated with this term compared to 0.43 % in the *Arabidopsis* transcriptome (Table 5-9). Other GO terms with siginificant enrighment include lipid localization (8/69) and post-embryonic development (12/69). Of the 18 genes enriched under photosynthesis, 7 were found to be downregulated by a $\log_2$ fold-change ≥ 2 (highlighted in bold in Table 5-10). Genes upregulated and enriched under GO: 0010876 (lipid localization) and GO:0009791 (post-embryonic development) all show a dose dependent increase in expression relative to an increase in *AtIPCS1* transcripts in the transgenic lines.

NdhH a subunit of NDH (NADH dehydrogenase-like complex) which is responsible for factilitating the movement of electrons from the stroma to plastonquinone was found to upregulated. It has been shown that plant mutants which are unable to fold the NADhH

Table 5-9: GO enrichment of genes up-regulated in response to the constitutive overexpression of _AtIPCS1_

GO term	Description	Number in input list	Number in Background	p-value
GO:0015979	Photosynthesis	18	162	9.10E-27
GO:0019684	Photosynthesis, light reaction	12	103	2.60E-18
GO:0006091	Generation of precursor metabolites and energy	15	285	8.40E-18
GO:0010876	Lipid localization	8	24	8.20E-16
GO:0009791	Post-embryonic development	12	705	6.40E-09
GO:0044237	Cellular metabolic process	37	8722	4.10E-08
GO:0008152	Metabolic process	41	10614	5.70E-08
GO:0009987	Cellular process	42	11684	2.90E-07
GO:0009737	Response to abscisic acid stimulus	8	378	5.30E-07
GO:0009845	Seed germination	5	83	6.10E-07
GO:0048316	Seed development	9	530	6.00E-07
GO:0010154	Fruit development	9	557	9.00E-07
GO:0033036	Macromolecule localization	8	462	2.30E-06
GO:0048608	Reproductive structure development	10	978	1.20E-05
GO:0003006	Reproductive developmental process	10	978	1.20E-05
GO:0006869	Lipid transport	5	163	1.50E-05
GO:0007275	Multicellular organismal development	14	2020	1.50E-05
GO:0032501	Multicellular organismal process	14	2094	2.30E-05
GO:0009409	Response to cold	6	328	3.40E-05
GO:0055114	Oxidation reduction	5	203	4.00E-05
GO:0022414	Reproductive process	10	1161	5.00E-05
GO:0000003	Reproduction	10	1186	6.00E-05
GO:0032502	Developmental process	14	2304	6.40E-05
GO:0006412	Translation	11	1445	6.20E-05
GO:0009725	Response to hormone stimulus	9	982	7.80E-05
GO:0009719	Response to endogenous stimulus	9	1068	0.00015
GO:0009266	Response to temperature stimulus	6	485	0.00028
GO:0050896	Response to stimulus	18	4057	0.00028
GO:0044085	Cellular component biogenesis	6	571	0.00065
GO:0051179	Localization	11	1922	0.00071
GO:0010033	Response to organic substance	9	1342	0.00076
GO:0048856	Anatomical structure development	10	1726	0.0012
GO:0042221	Response to chemical stimulus	10	2085	0.0045
GO:0044249	Cellular biosynthetic process	17	4925	0.0065
GO:0045735	Nutrient reservoir activity	10	63	1.20E-16
GO:0005198	Structural molecule activity	11	659	3.60E-08
GO:0003735	Structural constituent of ribosome	10	494	2.70E-08
GO:0016651	Oxidoreductase activity, acting on NADH or NADPH	5	103	1.70E-06
GO:0046906	Tetrapyrrole binding	5	136	6.20E-06
GO:0008289	Lipid binding	5	323	0.00034
GO:0008324	Cation transmembrane transporter activity	5	503	0.0024

Table 5-10: Genes upregulated by the constitutive overexpression of _AtIPCS1_ under GO term: 0015979 (photosynthesis)

Gene ID	Gene annotation	_At1-over-exp 3_		_At1-over-exp 7_	
		Log$_2$ fold change	p-value	Log$_2$ fold-change	p-value
ATCG00730	A CHLOROPLAST GENE ENCODING SUBUNIT IV OF THE CYTOCHROME B6/F COMPLEX PETD	1.3	2.93E-05	1.6	1.02E-06
ATCG01110	ENCODES THE 49KDA PLASTID NAD(P)H DEHYDROGENASE SUBUNIT H PROTEIN (NDhH)	**2.8**	**3.73E-19**	**4.0**	**4.72E-37**
ATCG01100	NADH DEHYDROGENASE ND1 NDHA	**1.6**	**2.92E-06**	**2.6**	**2.62E-15**
ATCG00520	ENCODES A PROTEIN REQUIRED FOR PHOTOSYSTEM I ASSEMBLY AND STABILITY	**1.3**	**0.000136246**	**2.0**	**6.06E-09**
ATCG01090	ENCODES SUBUNIT OF THE CHLOROPLAST NAD(P)H DEHYDROGENASE COMPLEX NDHI	**1.5**	**1.14E-05**	**2.5**	**1.60E-13**
ATCG00270	PSII D2 PROTEIN PSBD	1.8	5.79E-08	1.9	4.30E-09
ATCG00680	ENCODES FOR CP47, SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	1.0	0.000857587	1.2	0.00014644
ATCG00710	ENCODES A 8 KD PHOSPHOPROTEIN THAT IS A COMPONENT OF THE PHOTOSYSTEM II	1.1	0.000866598	1.5	1.80E-06
ATCG00360	ENCODES A PROTEIN REQUIRED FOR PHOTOSYSTEM I ASSEMBLY AND STABILITY	**1.5**	**6.99E-06**	**2.3**	**5.20E-12**
ATCG00300	ENCODES PSBZ, WHICH IS A SUBUNIT OF PHOTOSYSTEM II.	1.1	0.002252376	1.6	6.17E-06
ATCG00440	ENCODES NADH DEHYDROGENASE D3 SUBUNIT OF THE CHLOROPLAST NAD(P)H DEHYDROGENASE COMPLEX NDHC	**1.3**	**0.000167954**	**2.0**	**3.59E-09**
ATCG00720	ENCODES THE CYTOCHROME B(6) SUBUNIT OF THE CYTOCHROME B6F COMPLEX	1.1	0.000543643	1.4	5.73E-06
ATCG00340	ENCODES THE D1 SUBUNIT OF PHOTOSYSTEM I AND II REACTION CENTERS	1.2	0.000120455	1.6	1.29E-06
ATCG00580	PSII CYTOCHROME B559	1.1	0.000365459	1.3	8.33E-05
AT5G01600	ENCODES A FERRETIN PROTEIN THAT IS TARGETED TO THE CHLOROPLAST	1.1	1.19E-63	1.5	4.61E-116
ATCG01060	ENCODES THE PSAC SUBUNIT OF PHOTOSYSTEM I.	**1.3**	**5.19E-05**	**2.2**	**3.40E-12**
ATCG00540	ENCODES CYTOCHROME F APOPROTEIN	1.1	0.000430677	1.9	3.70E-09
ATCG00280	ENCODING A CP43 SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	1.5	2.31E-06	1.6	1.08E-06

subunit due to a defective chloroplast chaperonin subunit (Cpn60β4), cannot produce NADHh for the assembly of the NDH complex assembly with no phenotye reported as associated with the defect[55]. From the list of genes up regulated, all are subunits of photosynthesis systmen I and II, which should contribute to an increase in the the plant's ability to efficiently convert light energy to chemical energy for plant development.

For genes enriched under GO term lipid localization (Table 5-11), all apart for the seed storage proteins, were found to be expressed specifically in immature seeds but it is also thought that these genes may be expressed in vegetative tissues during plant development[56]. Suppression of *olesin 1* resulted in large oil bodies (4–10 μm^2) compared to Col0 (1 μm^2) in embryo cells and a delay in germination[57, 58]. Double mutants, *ole1 ole3* and *ole1 ole2,* were found to have even larger oil bodies (>10 μm^2) and had lower levels of olesin resulting in delays in germination[57]. The mutant phenotype of large oil bodies in *ole1* single mutant, was reversed by complementation with a functional OLESIN gene, and the delay in germination found to be reversible by cold treatment or by growing the seeds in the presence of sucrose[58]. In addition, upon undergoing freezing treatment, the delay in germination for the single mutants (ole1-4) was found to increase with an inverse relationship to the amount of olesin produced in the oil bodies. This indicated that olesin levels affect seed freezing tolerance[57]. Storage proteins (2S albumin 1-4) were also upregulated in response to *AtIPCS1* overexpression, and are known to serve as nitrogen and sulphur reserves for plant germination[59, 60].

In addition there is a clear dose dependent relationship between the increase in *AtIPCS1* expression and the increasing upregulation of the genes which are involved in the nutrient reservoir and germination onset in seeds pathways. From these findings, it appears that *At*IPCS1 is linked to the seed development pathway, with overexpression of this isofrom

resulting in an increase in expression of seed specific proteins which could, theoretically, lead to higher rates of sucessful germination.

Further analyses of upregulated genes show 12 are enriched under GO: 0009791 (post embryonic development) (Table 5-12), with *OLESIN1-3* already identified as involved in seed germination. The other 9 genes upregulated include, *LEA* a protein known to accumulate during the final maturation stage of seeds termed desiccation, and in vegetative tissues in response to stress such as cold and drought[61]; *CRU3*, a seed storage protein that is phosphorylated in response to plant abscisic acid (ABA) treatment and found to play a role in protecting proteins needed for germination from oxidation[62], *S1* the calcium binding oil body surface protein that facilitates the degradation of eicosenoic acid found in storage triacylglycerol in oil bodies and affects seedling growth post germination[63], and *EPR1* which is expressed specifically in the endosperm of germinating seeds but with an as yet uncharacterised role in plants, and with no phenotype detected in insertional mutants[64]. In *Arabidopsis*, *EM6* which is expressed in all tissues in the embryo, has been shown to be necessary for normal seed development, with a T-DNA insertional mutant exhibiting a phenotype of premature seed dehydration at the distal end 13 days after flowering[65]. PER1, a peroxidase proven to have antioxidant activity by protection of *E.coli* from cumen hydroperoxide, with overexpression shown to result in a delay in germination in seeds under-going osmotic stress[66], indicating that the upregulation of the protein serves to inhibit germination in response to stress.

FER1, a ferritin protein which accumulates in response to iron, has been shown to be involved in pathways that affect the aging process in *Arabidopsis*; a knock out mutant, exhibits a phenotype of early senescence in leaves, a reduction in chlorophyll concentration and reduced photochemical efficiency in PSII [67]. Conversely upregulation

of this gene should result in high photochemical efficiency of PSII and delayed aging; this was not observed in any of the *AtIPCS* over-expressors. The upregulated gene *M10*, is specifically expressed in seeds and repressed 2 hours after imbibition; it accumulates in response to cold stress but is repressed in response to salt/drought stress and treatment with abscisic acid (ABA)[68].

Next, genes up regulated in response to *AtIPCS2* overexpression were analysed, and a significant enrichment found for GO term: 0015979 (photosynthesis) with 13.3% (30/226) of the genes associated with this term compared to 6.14% to 0.004% in *Arabidopsis* transcriptome (Table 5-13). Other GO terms significantly enriched include lipid localization (17/226) and post embryonic development (12/226). For GO term photosynthesis 30 genes are identified from 162 involved in this pathway in *Arabidopsis*, and for the GO term lipid localization 17 genes are identified as upregulated from 24 genes known to be involved in that process, which strongly indicated the there is a link between proteins involved in lipid localization and IPCS expression. Closer inspection of genes up regulated in response to *AtIPCS2* overexpression reveals that most of those involved in photosynthesis (Table 5-14) that were upregulated in response to *AtIPCS1* overexpression were also upregulated in response to *AtIPCS2* expression. However, there were some of these genes that were only upregulated in *AtIPCS2* over- expressing lines, these include: the F subunit of *NDH* (*NDHF*) which has been shown to affect the expression of H subunit (*NDHH*) in mutants exhibiting low transcript levels of *NDHF*, resulting in a decrease in NDH activity quantified by a measure of chlorophyll florescence and with no obvious phenotype[69], PSB28 in *Synechocystis* 6803 has been shown to be involved in the biogenesis of chlorophyll binding proteins (CP47 and PsaA/PsaB) by facilitating the conversion of Mg protoporphyrin monomethyl ester to chlorophyll[70]. Insertional mutants of *PSB28* have shown an accumulation of Mg protop-

Table 5-11: Genes upregulated by the constitutive overexpression of *At*IPCS1 under GO term:0010876 (lipid localization)

Gene ID	Gene annotation	At1 over-exp-3		At1 over-exp-7	
		Log$_2$ fold change	p-value	Log$_2$ fold-change	p-value
AT5G54740	2S SEED STORAGE PROTEIN 5	5.1	3.42E-97	8.2	3.86E-259
AT4G25140	OLEO1 (OLEOSIN 1)	3.5	4.38E-45	7.1	2.82E-192
AT4G27140	2S SEED STORAGE PROTEIN 1	4.2	5.35E-65	8.0	1.03E-238
AT4G27150	2S SEED STORAGE PROTEIN 2	3.8	2.16E-40	7.3	2.16E-147
AT4G27160	2S SEED STORAGE PROTEIN 3	5.0	1.84E-81	8.5	2.51E-239
AT5G40420	OLEO2 (OLEOSIN 2)	1.4	1.45E-07	4.6	3.63E-70
AT4G27170	2S SEED STORAGE PROTEIN 4	2.7	1.03E-19	5.7	5.84E-88
AT3G27660	OLEO4 (OLEOSIN 4)	2.1	1.35E-15	5.2	7.26E-100

Table 5-12: Genes upregulated by the constitutive overexpression of _At_IPCS1 under GO term: 0009791 (post embryonic development)

		At1 over-exp-3		_At1 over-exp-7_	
Gene ID	Gene annotation	Log$_2$ fold change	p-value	Log$_2$ fold change	p-value
AT3G22500	LATE EMBRYOGENESIS ABUNDANT (LEA) PROTEIN ECP31	1.9	1.70E-09	4.7	8.46E-63
AT3G27660	OLEO4 (OLEOSIN 4)	2.1	1.35E-15	5.2	7.26E-100
AT4G28520	CRU3 (CRUCIFERIN 3)	5.8	9.91E-131	9.6	0
AT4G25140	OLEO1 (OLEOSIN 1)	3.5	4.38E-45	7.1	2.82E-192
AT4G26740	S1 (SEED GENE 1)	1.7	1.44E-08	5.1	8.42E-81
AT2G27380	EPR1	2.9	3.08E-31	6.0	6.18E-135
AT2G40170	EM6 (LATE EMBRYOGENESIS ABUNDANT 6)	1.3	1.18E-06	3.8	1.09E-51
AT1G48130	1-CYSTEINE PEROXIREDOXIN (AtPER1)	1.4	1.87E-11	4.1	1.33E-98
AT5G01600	FER1; FERRITIN PROTEIN	1.1	1.19E-63	1.5	4.61E-116
AT5G40420	OLEO2 (OLEOSIN 2)	1.4	1.45E-07	4.6	3.63E-70
AT1G03880	CRU2 (CRUCIFERIN 2)	2.4	1.77E-17	6.9	2.47E-149
AT2G41280	M10 ENCODES A HYDROPHILIC PROTEIN SIMILAR TO LATE EMBRYOGENESIS ACTIVATED (LEA)	1.3	7.58E-05	3.8	1.77E-35

orphyrin IX monomethyl ester, lower levels of CP47 and PsaA/PsaB combined with slower photoautotrophic growth; there is currently no functional study of PSB28 that has been performed in *Arabidopsis*[70].

The calvin cycle enzyme, *SBPASE* activity in tobacco plants overexpressing the *Arabidopsis* gene was found to have a positive effect on rates of photosynthesis. This was associated with an up to 50% increase in the accumulation of the carbohydrates, sucrose and starch, and an increase in total leaf surface area and shoot biomass of up to 30%[71]. The two *PQL* isoforms 1 and 2 were also upregulated; work done by Yabuta et al. show that an insertional mutant did not display the usual increase in chlorophyll fluorescence upon exposure to actinic light which is associated with the activity of the NDH complex[72]. The minor light harvesting complex, *LHCA6*, was also upregulated. A *lhca6* mutant showed a reduction in NDH activity in immature and mature leaves (~60% compared to wildtype *Arabidopsis*) with no obvious phenotype[73]. Again, the level of upregulation of genes identified in response to *AtIPCS2* overexpression was found to be about the same or in some cases exceeded the levels for the same genes in *AtIPCS1* over-expressors. This is despite the fact that the fold change in *AtIPCS1* expression was much than *AtIPCS2*.

Under the GO term lipid localization 9 /17 genes encode for lipid transfer proteins (LPT) (Table 5-15); LPTs like the yeast transfer lipid Sec14, facilitate the movement of phospholipids between membranes and has been shown to complement sec14 mutants[74]. Some LPTs have been shown to have a role in plant defence, for example, insertional mutant *dir-1-1* (defective in induced resistance) exhibited increase susceptibility to *P. syringae*. DIR1 is hypothesised to function as a chaperon of lipid signals during SAR[75]. The LTP, At5g05960, has been shown to be upregulated in the t-DNA mutant, *nhx1* (vacuolar Na^+/H^+ antiporter) in response to salt stress[76]. The LPT orthologues TaLPT 1

Table 5-13: GO enrichment of genes upregulated in response to the constitutive overexpression of _AtIPCS2_

GO term	Description	Number in input list	Number in Background	p-value
GO:0015979	Photosynthesis	30	162	1.70E-31
GO:0010876	Lipid localization	17	24	1.80E-26
GO:0006091	Generation of precursor metabolites and energy	26	285	2.50E-20
GO:0008152	Metabolic process	143	10614	2.00E-18
GO:0019684	Photosynthesis, light reaction	17	103	1.40E-17
GO:0019748	Secondary metabolic process	28	489	7.80E-17
GO:0044237	Cellular metabolic process	122	8722	4.00E-16
GO:0009987	Cellular process	143	11684	1.30E-14
GO:0050896	Response to stimulus	70	4057	1.30E-12
GO:0015977	Carbon fixation	8	16	3.40E-12
GO:0009409	Response to cold	19	328	7.00E-12
GO:0009628	Response to abiotic stimulus	38	1471	9.20E-12
GO:0009266	Response to temperature stimulus	22	485	1.40E-11
GO:0044262	Cellular carbohydrate metabolic process	20	417	4.90E-11
GO:0019760	Glucosinolate metabolic process	10	62	8.70E-11
GO:0016143	S-glycoside metabolic process	10	62	8.70E-11
GO:0019757	Glycosinolate metabolic process	10	62	8.70E-11
GO:0006869	Lipid transport	13	163	4.10E-10
GO:0016137	Glycoside metabolic process	11	104	6.00E-10
GO:0016144	S-glycoside biosynthetic process	8	41	1.80E-09
GO:0019761	Glucosinolate biosynthetic process	8	41	1.80E-09
GO:0019758	Glycosinolate biosynthetic process	8	41	1.80E-09
GO:0006950	Response to stress	43	2320	8.60E-09
GO:0016138	Glycoside biosynthetic process	9	79	1.20E-08
GO:0006790	Sulfur metabolic process	13	220	1.20E-08
GO:0009719	Response to endogenous stimulus	27	1068	1.80E-08
GO:0044272	Sulfur compound biosynthetic process	10	115	2.10E-08
GO:0009737	Response to abscisic acid stimulus	16	378	2.30E-08
GO:0009058	Biosynthetic process	70	5118	2.90E-08
GO:0055114	Oxidation reduction	12	203	4.50E-08
GO:0033036	Macromolecule localization	17	462	6.00E-08
GO:0009767	Photosynthetic electron transport chain	7	46	8.90E-08
GO:0005975	Carbohydrate metabolic process	23	866	9.30E-08
GO:0042221	Response to chemical stimulus	38	2085	1.10E-07
GO:0010033	Response to organic substance	29	1342	1.40E-07
GO:0016051	Carbohydrate biosynthetic process	13	277	1.60E-07
GO:0009664	Plant-type cell wall organization	8	79	1.90E-07
GO:0009725	Response to hormone stimulus	24	982	2.10E-07
GO:0022900	Electron transport chain	8	88	4.00E-07
GO:0034637	Cellular carbohydrate biosynthetic process	10	177	9.00E-07
GO:0034641	Cellular nitrogen compound metabolic process	16	506	1.00E-06
GO:0009813	Flavonoid biosynthetic process	7	69	1.10E-06
GO:0009699	Phenylpropanoid biosynthetic process	9	141	1.20E-06
GO:0042398	Cellular amino acid derivative biosynthetic process	11	233	1.40E-06
GO:0044249	Cellular biosynthetic process	63	4925	1.80E-06

Table 5-14: Genes upregulated by the constitutive overexpression of *AtIPCS2* under GO term: 0015979 (photosynthesis)

Gene ID	Gene annotation	At2 over-exp-2		At2 over-exp-9	
		Log₂ fold change	p-value	Log₂ fold-change	p-value
ATCG00520	ENCODES A PROTEIN REQUIRED FOR PHOTOSYSTEM I ASSEMBLY AND STABILITY	**2.4**	**3.56E-12**	**3.6**	**7.23E-26**
ATCG01090	ENCODES SUBUNIT OF THE CHLOROPLAST NAD(P)H DEHYDROGENASE COMPLEX NDHI	**2.7**	**1.01E-15**	**4.0**	**4.89E-33**
ATCG00540	ENCODES CYTOCHROME F APOPROTEIN	**2.2**	**1.72E-12**	**3.4**	**7.67E-27**
ATCG01010	CHLOROPLAST ENCODED NADH DEHYDROGENASE UNIT. NDHF	**2.2**	**2.06E-10**	**3.4**	**7.33E-22**
ATCG00280	CHLOROPLAST GENE ENCODING A CP43 SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	**1.9**	**2.22E-09**	**3.0**	**1.05E-20**
ATCG00680	ENCODES FOR CP47, SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	**1.5**	**8.93E-07**	**2.5**	**2.18E-15**
AT4G28660	PSB28 (PHOTOSYSTEM II REACTION CENTER PSB28 PROTEIN)	1.6	6.17E-59	1.2	8.53E-37
ATCG01100	NADH DEHYDROGENASE ND1 NDHA	**2.9**	**1.82E-18**	**4.0**	**3.49E-33**
AT3G04790	RIBOSE 5-PHOSPHATE ISOMERASE-RELATED	1.8	1.32E-101	1.3	6.95E-58
ATCG00270	PSII D2 PROTEIN PSBD	**2.4**	**7.28E-13**	**3.4**	**1.13E-25**
ATCG01060	ENCODES THE PSAC SUBUNIT OF PHOTOSYSTEM I	**2.6**	**8.05E-17**	**3.6**	**2.28E-30**
AT3G55800	SBPASE (SEDOHEPTULOSE-BISPHOSPHATASE);	1.2	1.64E-100	1.0	4.49E-67
AT4G09650	ATPD (ATP SYNTHASE DELTA-SUBUNIT GENE);	1.5	7.39E-119	1.2	3.91E-76
ATCG00730	A CHLOROPLAST GENE ENCODING SUBUNIT IV OF THE CYTOCHROME B6	**1.9**	**4.38E-09**	**2.8**	**6.43E-19**
ATCG01080	NADH DEHYDROGENASE ND6 NDHG	**1.9**	**9.92E-08**	**3.1**	**5.52E-19**
ATCG00710	ENCODES A 8 KD PHOSPHOPROTEIN	**1.7**	**5.64E-08**	**2.8**	**1.16E-17**
AT2G01590	CRR3 (CHLORORESPIRATORY REDUCTION 3)	1.2	1.06E-40	1.0	1.79E-31
ATCG00580	PSII CYTOCHROME B559	**1.8**	**8.75E-09**	**2.4**	**2.18E-13**
ATCG00720	ENCODES THE CYTOCHROME B(6) SUBUNIT OF THE CYTOCHROME B6F COMPLEX	**1.7**	**7.29E-08**	**2.9**	**4.62E-20**
AT3G01440	PSB-LIKE PROTEIN 2 (PQL2)	1.8	7.06E-107	1.5	2.74E-70
AT1G14150	PSB-LIKE PROTEIN 1 (PQL1)	1.5	3.82E-75	1.2	1.54E-44
ATCG01110	ENCODES THE 49KDA PLASTID NAD(P)H DEHYDROGENASE SUBUNIT H PROTEIN	**4.2**	**1.35E-40**	**4.7**	**5.23E-52**
AT1G19150	LHCA6; CHLOROPHYLL BINDING PSI TYPE II CHLOROPHYLL A/B-BINDING PROTEIN	2.0	2.50E-114	1.6	5.98E-73
AT1G60950	FED A; 2 IRON, 2 SULFUR CLUSTER BINDING	1.6	1.03E-63	1.5	1.68E-54
ATCG00300	ENCODES PSBZ, WHICH IS A SUBUNIT OF PHOTOSYSTEM II	**1.7**	**8.98E-07**	**3.2**	**1.04E-19**
ATCG00440	ENCODES NADH DEHYDROGENASE D3 SUBUNIT OF THE CHLOROPLAST NAD(P)H DEHYDROGENASE COMPLEX NDHC	**2.7**	**6.65E-16**	**3.4**	**4.14E-24**
ATCG00340	ENCODES THE D1 SUBUNIT OF PHOTOSYSTEM I AND II REACTION CENTERS. PSAB	**1.8**	**1.35E-08**	**2.9**	**4.71E-19**
ATCG00360	ENCODES A PROTEIN REQUIRED FOR PHOTOSYSTEM I ASSEMBLY AND STABILITY	**2.7**	**1.25E-15**	**3.8**	**1.03E-29**
AT3G16250	NDF4 (NDH-DEPENDENT CYCLIC ELECTRON FLOW 1);	1.4	1.02E-79	1.1	5.91E-52
AT3G62410	PROTEIN BINDING CP12-2 ENCODES A SMALL PEPTIDE FOUND IN THE CHLOROPLAST STROMA	1.8	5.62E-79	1.4	5.57E-49

Table 5-15: Genes upregulated by the constitutive overexpression of *AtIPCS2* under GO term: 0010876 (lipid localization)

Gene ID	Gene annotation	At2 over-exp-2		At2 over-exp-9	
		Log₂ fold change	p-value	Log₂ fold-change	p-value
AT5G59310	LTP4 (LIPID TRANSFER PROTEIN 4)	**3.7**	**2.07E-28**	**1.0**	**0.003587715**
AT5G48490	DIR1 LIPID TRANSFER PROTEIN (LTP)	1.2	4.21E-16	1.3	4.78E-18
AT3G08770	LTP6; LIPID BINDING PREDICTED TO ENCODE A PR (PATHOGENESIS-RELATED) PROTEIN	**2.0**	**3.96E-65**	**1.3**	**4.44E-27**
AT5G40420	OLEO2 (OLEOSIN 2)	2.5	9.95E-21	2.8	5.74E-25
AT4G25140	OLEO1 (OLEOSIN 1)	**4.6**	**2.85E-80**	**5.0**	**1.20E-91**
AT4G27140	2S SEED STORAGE PROTEIN 1	**4.3**	**1.28E-68**	**6.3**	**1.95E-148**
AT4G27150	2S SEED STORAGE PROTEIN 2	**4.7**	**1.09E-61**	**5.4**	**1.89E-81**
AT4G27160	2S SEED STORAGE PROTEIN 3	**4.2**	**4.18E-58**	**6.3**	**1.28E-129**
AT5G64080	LIPID TRANSFER PROTEIN (LTP)	1.6	1.28E-30	1.2	3.85E-18
AT5G05960	LIPID TRANSFER PROTEIN (LTP)	1.7	4.62E-51	1.2	6.31E-26
AT3G18280	LIPID TRANSFER PROTEIN (LTP)	1.3	7.33E-28	1.1	3.46E-19
AT5G54740	LIPID TRANSFER PROTEIN (LTP)	**5.8**	**1.57E-127**	**6.3**	**5.98E-153**
AT5G55410	LIPID TRANSFER PROTEIN (LTP)	**2.2**	**4.37E-12**	**2.0**	**1.50E-09**
AT4G27170	2S SEED STORAGE PROTEIN 4	**1.8**	**6.48E-09**	**3.9**	**2.71E-39**
AT3G27660	OLEO4 (OLEOSIN 4)	**2.8**	**6.81E-27**	**2.6**	**5.28E-23**
AT2G37870	LIPID TRANSFER PROTEIN (LTP)	**2.9**	**2.88E-27**	**1.5**	**4.73E-08**
AT3G01570	OLE05 (OLEOSIN 5)	1.5	8.26E-13	1.3	1.92E-08

Table 5-16: Genes upregulated by the constitutive overexpression of *AtIPCS2* under GO term: 0009791 (post embryonic development)

Gene ID	Gene annotation	At2-over-exp-2		At2-over-exp-9	
		Log$_2$ fold change	p-value	Log$_2$ fold change	p-value
AT3G22500	ATECP31 (LATE EMBRYOGENESIS PROTEIN ECP31)	**2.7**	**5.10E-20**	**2.2**	**6.09E-13**
AT4G24960	ATHVA22D (HVA22-LIKE PROTEIN)	1.0	2.41E-28	1.5	1.92E-57
AT3G27660	OLEO4 (OLEOSIN 4)	**2.8**	**6.81E-27**	**2.6**	**5.28E-23**
AT5G56030	HSP81-2 (HEAT SHOCK PROTEIN 81-2)	1.3	6.51E-80	1.0	3.92E-53
AT4G28520	CRU3 (CRUCIFERIN 3)	**6.0**	**5.56E-142**	**7.3**	**9.09E-207**
AT4G25140	OLEO1 (OLEOSIN 1)	**4.6**	**2.85E-80**	**5.0**	**1.20E-91**
AT4G26740	ATS1 (ARABIDOPSIS THALIANA SEED GENE 1)	**2.5**	**3.28E-18**	**2.9**	**4.39E-23**
AT2G27380	ATEPR1 (EXTENSIN PROLINE-RICH 1)	**2.9**	**4.27E-33**	**5.1**	**6.03E-100**
AT2G40170	GEA6 (LATE EMBRYOGENESIS ABUNDANT 6)	2.5	3.49E-22	1.2	2.59E-05
AT1G48130	ATPER1 (THIOREDOXIN PEROXIDASE)	**2.0**	**7.06E-23**	**2.3**	**2.32E-29**
AT3G20210	DELTA-VPE (CYSTEIN-TYPE ENDOPEPTIDASE)	1.2	0.000532902	1.4	8.21E-05
AT5G40420	OLEO2 (OLEOSIN 2)	**2.5**	**9.95E-21**	**2.8**	**5.74E-25**
AT1G03880	CRU2 (CRUCIFERIN 2)	**2.8**	**1.65E-22**	**4.4**	**2.05E-59**
AT3G58190	LBD29 (LATERAL ORGAN BOUNDARIES-DOMAIN 29)	1.5	2.69E-06	1.7	1.20E-07

and 2 were found to be upregulated in response to salt and drought stress in *Tritium aestivum*[77], showing that the plant LPTs expression are induced in response to biotic and abiotic stress. The function of most LTP are currently uncharacterized so it is not possible to hypothesize what the effects of the up regulation of the genes identified in response to *AtIPCS2* overexpression, will have on plant development. Genes upregulated under GO term post embryonic development with a $\log_2$ fold ≥ 2 in response to *AtIPCS2* transcript increase include *CRU2*, *CRU3*, *EPR1*, *S1* and *ECP31* (Figure 5-16).

For the *AtIPCS3* over-expressor there was a significant enrichment of upregulated genes associated with the GO term, GO: 0010876 (lipid localization) with 36.8% (7/19) of the genes associated with this term compared to 0.06% in the *Arabidopsis* transcriptome (Table 5-17). Other GO terms with siginificant enrighment include macromolcule localisation (7/19), lipid transport (5/7) and post embryonic development (5/19). Genes enriched under GO terms macromolecule localisation and lipid transport were also identified under the lipid localization GO term. Genes enriched under lipid localization and post-embryonic development that were upregulated in response to *AtIPCS3* overexpression, had already been identified in the responses to *AtIPCS1* and *2* overexpression. For the sake of completion the fold change in expression of these genes is shown in Table 5-18 and Table 5-19, with genes up-regulated by $\log_2$ fold-change ≥ 2 highlighed in bold. All genes identified showed a dose dependent increase in expression associated with higher transcript levels of *AtIPCS3* in At1 over-exp-7 compared to At1 over-exp-3.

Analyses of genes upregulated in response to increases in *AtIPCS1*, 2 and *3* expressions showed an expected significant enrichment for lipid localization as with 46.7% (7/15) of the genes associated with this term compared to 6.14% in the *Arabidopsis* transcriptome

(Table 5-20). These genes include *2S seed storgage protein 1-4*, *OLESIN 1* and *OLESIN 4*. This showed that genes enriched under lipid localization, and found to be involved in the early stage plant development (embyronic), are conserved with respect to the response to *AtIPCS1*, *2* and *3* overexpression indicating that modulation of *AtIPCS* expression is likely to be vital for survival, with a decrease in the genes identified likely to lead to non-viable seeds.

Following the analysis of the overexpressing transgenic lines, the transcriptome of the At1 t-DNA mutant was also analysed to find out to what extent the decrease in *AtIPCS1* expression relative to Col0, would have on the pathways already identified as affected by *AtIPCS* expression.

5.6 Genes downregulated in response to *AtIPCS1* homozygous knock down

Genes down regulated in response to the decrease in *AtIPCS1* expression were found to be significantly enriched for genes annotated with the GO term GO: 0050896 (response to stimulus), with 40.3% (52/129) of the genes associated with this term compared to 6.14% in the *Arabidopsis* transcriptome (Table 5-21). Other GO terms with siginificant enrichment include secondary metabolic process (21/129) and response to stress (34/129). For comparison with At1 over-expressors, genes annotated under the GO term response to stress were analysed. Of the 34 genes, 17 were associated with defence response and 8 down regulated by a $\log_2$ fold change ≥ 2 (highlighed in bold in Table 5-22). Some genes that were downregulated in response to *AtIPCS1* overexpression were also found to be down regulated in response to knockdown of *AtIPCS1* expresssion. These included, *MC8*, *WRKY 54*, *PR5*, *WRKY 70*, *VQ containg motif protein 10*, and ATP binding *BIP3*.

Table 5-17: GO enrichment of genes upregulated in response to the constitutive overexpression of *AtIPCS3*

GO term	Description	Number in input list	Number in Background	p-value
GO:0010876	Lipid localization	7	24	6.00E-18
GO:0033036	Macromolecule localization	7	462	1.90E-09
GO:0006869	Lipid transport	5	163	1.80E-08
GO:0051179	Localization	7	1922	2.60E-05
GO:0009791	Post-embryonic development	5	705	2.20E-05
GO:0032502	Developmental process	7	2304	8.30E-05
GO:0007275	Multicellular organismal development	6	2020	0.00035
GO:0032501	Multicellular organismal process	6	2094	0.00042
GO:0006810	Transport	5	1846	0.0018
GO:0051234	Establishment of localization	5	1851	0.0019
GO:0050896	Response to stimulus	7	4057	0.0026
GO:0045735	Nutrient reservoir activity	9	63	1.80E-20
GO:0008289	Lipid binding	5	323	5.00E-07

Table 5-18: Genes upregulated by the constitutive overexpression of *AtIPCS3* under GO term: 0010876 (lipid localization)

Gene ID	Gene annotation	At3 over-exp-4		At3 over-exp-5	
		Log$_2$ fold change	p-value	Log$_2$ fold-change	p-value
AT5G54740	PROTEASE INHIBITOR/SEED STORAGE/LIPID TRANSFER PROTEIN (LTP)	3.9	4.20E-57	7.7	9.99E-230
AT4G25140	ENCODES OLEOSIN1	2.5	1.01E-21	6.1	3.11E-141
AT4G27140	2S SEED STORAGE PROTEIN 1	3.0	4.86E-32	7.2	6.10E-193
AT4G27150	2S SEED STORAGE PROTEIN 2	2.3	8.24E-16	6.4	7.03E-115
AT4G27160	2S SEED STORAGE PROTEIN 3	3.6	4.03E-43	7.3	2.59E-176
AT4G27170	2S SEED STORAGE PROTEIN 4	1.8	6.80E-09	4.9	9.93E-64
AT3G27660	OLEO4 (OLEOSIN 4)	1.1	8.65E-05	4.0	2.28E-58

Table 5-19: Genes upregulated by the constitutive overexpression of *AtIPCS3* under GO term: 0009791 (post embryonic development)

		At3 over-exp-4		At3 over-exp-5	
Gene ID	**Gene annotation**	**Log$_2$ fold change**	**p-value**	**Log$_2$ fold-change**	**p-value**
AT4G28520	CRU3 (CRUCIFERIN 3)	**4.2**	**6.51E-67**	**8.3**	**2.55E-271**
AT4G25140	OLEO1 (OLEOSIN 1)	**2.5**	**1.01E-21**	**6.1**	**3.11E-141**
AT2G27380	ATEPR1	**1.5**	**4.85E-09**	**5.6**	**8.36E-120**
AT3G27660	OLEO4 (OLEOSIN 4)	**1.1**	**8.65E-05**	**4.0**	**2.28E-58**
AT1G03880	CRU2 (CRUCIFERIN 2)	**1.9**	**1.19E-10**	**5.5**	**6.22E-96**

Table 5-20: GO enrichment of genes upregulated in response to the constitutive overexpression of *AtIPCS1*, *2* and *3*

GO term	Description	Number in input list	Number in Background	p-value
GO:0010876	Lipid localization	7	24	7.70E-19
GO:0033036	Macromolecule localization	7	462	2.60E-10
GO:0006869	Lipid transport	5	163	4.70E-09
GO:0051179	Localization	7	1922	4.00E-06
GO:0009791	Post-embryonic development	5	705	5.90E-06
GO:0007275	Multicellular organismal development	6	2020	7.80E-05
GO:0032501	Multicellular organismal process	6	2094	9.50E-05
GO:0032502	Developmental process	6	2304	0.00016
GO:0051234	Establishment of localization	5	1851	0.00056
GO:0050896	Response to stimulus	7	4057	0.00049
GO:0006810	Transport	5	1846	0.00056
GO:0045735	Nutrient reservoir activity	9	63	9.80E-22
GO:0008289	Lipid binding	5	323	1.30E-07

The down regulation of these genes seems to be a basal response to any alteration in the expression of *AtIPCS1* as the level of down regulation of these genes is never > 2 in the knock-down and the over-expressor lines. Other genes which had already been identified as down regulated in response to overexpression of either *AtIPCS2* or *At*IPCS3 included, *MC8*, *RNS1*, *TPS04* and *VQ motif containing protein 10*. Again these genes had a $\log_2$ fold-change > 2 (apart from TPS04 down regulation in At2 over exp 2 which had a $\log_2$ fold-change of 2.6) indicating a basal response to an alteration in *At*IPCS expression. It is worth noting that none of the genes identified as enriched under the GO term response to stress were linked to cold, drough and salt stress. These had been previously identified as down regulated in response to *AtIPCS* overexpression.

NIMIN1, a defence related gene, found to be down regulated by a $\log_2$ fold-change of 3.2 in At1 t-DNA mutant line, encodes a protein that interacts with NPR1, a regulator of SARs in plants. Treatment of plants with SAR inducing compounds, SA and 2,6-dichloroisonicotinic acid, resulted in transient expression of NIMIN, suggesting that is a sensitive and selective inducible target for specific SAR signals[78]. It has also been suggested that the transient expression and accumulation of NIMIN1 indicates it is a regulator of late-SAR response by controlling the expression of PR1[79]. The down regulation of this gene in response to the decreased *At*IPCS1 which was not identified in any of the IPCS over-expressors, indicated that the plant was even more sensitive to decreases *At*IPCS1 than it was to increased expression. Any further induction of SAR is effectively switched off, thereby further supporting role of *At*IPCS as a negative regulator of plant defence response. Another gene down regulated by a $\log_2$ fold-change > 2, and unidentified as responding to overexpression of *AtIPCS*, was *LOX2*. This gene has been shown to be necessary for JA accumulation in the plant response to wounding[80]; its down

regulation indicated that the t-DNA mutant lacks the ability to induce the accumulation of JA and the associated stress response.

The camalexin biosynthesis enzyme, (*CYP71A13*) and flavin dependent monooxygenase (*FMO1*) were also down regulated in response to *AtIPCS1* knock-down. Previous studies have shown them both to be upregulated in plants infected with *Golovinomyces orontii*[81]. FMO1 is also a well characterized positive regulator of the programmed cell death effectors EDS1 (Enhanced disease susceptibility 1) and PADA4 (phytoalexin deficient 4) which regulate the accumulation of SA and are necessary for basal resistance to pathogens[82]. *PADA4* is also identified as down regulated (-1.1 fold) in this study, as are three TIR class of R (resistance) proteins which recognize a diverse set of pathogen effectors and are associated with the second line of plant defence against pathogens termed effector triggered immunity (ETI)[83]. Also downregulated are MYB28 and MYB 29 encoding positive regulators of methionine-derived glucosinolates[84], which are upregulated in response to herbivore attack. Glucosinolates are broken down by thioglucosidases, releasing toxic compounds such as nitriles and isothiocynates which are toxic to nematodes[85]. In addition, *VSP2* a well characterized storage protein which has acid phosphatase activity, was also down regulated. VSP2 expression has been shown to be enhanced in response to plant defence against the herbivory insect, *Sponoptera littoralis*[86].

The DIR genes which are responsible for the biosynthesis of lignin, have been shown in *Physcomitrella patens* to be induced upon inoculation with *B. cinerea* in order to reinforce the cell wall as a pathogen defence[87]. The DIR proteins *DIR7* and *DIR 21*, which have been identified in this study as down regulated in response to *At*IPCS1 knock-down, have not been characterized in relation to plant response to stress in *Arabidopsis*. However, based on the expression of the DIR proteins identified in *P. patens*, it is probable that the

Table 5-21: GO enrichment of genes upregulated in response to the knock-down in expression of _AtIPCS1_ in the t-DNA insertion mutant

GO term	Description	Number in input list	Number in Background	p-value
GO:0050896	Response to stimulus	52	4057	4.10E-18
GO:0019748	Secondary metabolic process	21	489	6.50E-17
GO:0009987	Cellular process	82	11684	2.80E-14
GO:0006950	Response to stress	34	2320	3.60E-13
GO:0065007	Biological regulation	44	4188	4.20E-12
GO:0009607	Response to biotic stimulus	17	638	1.10E-10
GO:0050794	Regulation of cellular process	37	3375	1.20E-10
GO:0016137	Glycoside metabolic process	9	104	2.00E-10
GO:0009753	Response to jasmonic acid stimulus	11	215	3.80E-10
GO:0051707	Response to other organism	16	599	3.70E-10
GO:0016138	Glycoside biosynthetic process	8	79	6.80E-10
GO:0050789	Regulation of biological process	37	3697	1.50E-09
GO:0006952	Defence response	17	766	1.60E-09
GO:0051704	Multi-organism process	17	776	1.90E-09
GO:0044249	Cellular biosynthetic process	42	4925	9.50E-09
GO:0009058	Biosynthetic process	43	5118	9.10E-09
GO:0034637	Cellular carbohydrate biosynthetic process	9	177	1.70E-08
GO:0042221	Response to chemical stimulus	25	2085	4.30E-08
GO:0009751	Response to salicylic acid stimulus	9	200	4.60E-08
GO:0016051	Carbohydrate biosynthetic process	10	277	5.90E-08
GO:0019760	Glucosinolate metabolic process	6	62	1.30E-07
GO:0016143	S-glycoside metabolic process	6	62	1.30E-07
GO:0019757	Glycosinolate metabolic process	6	62	1.30E-07
GO:0042398	Cellular amino acid derivative biosynthetic process	9	233	1.60E-07
GO:0031347	Regulation of defence response	6	66	1.80E-07
GO:0006575	Cellular amino acid derivative metabolic process	10	315	1.90E-07
GO:0019438	Aromatic compound biosynthetic process	9	237	1.80E-07
GO:0044272	Sulfur compound biosynthetic process	7	115	2.20E-07
GO:0009698	Phenylpropanoid metabolic process	8	175	2.30E-07
GO:0044262	Cellular carbohydrate metabolic process	11	417	2.70E-07
GO:0051726	Regulation of cell cycle	7	127	4.10E-07
GO:0009828	Plant-type cell wall loosening	5	40	4.60E-07
GO:0016144	S-glycoside biosynthetic process	5	41	5.20E-07
GO:0019761	Glucosinolate biosynthetic process	5	41	5.20E-07
GO:0019758	Glycosinolate biosynthetic process	5	41	5.20E-07
GO:0042440	Pigment metabolic process	7	134	5.80E-07
GO:0080134	Regulation of response to stress	6	83	6.40E-07
GO:0009699	Phenylpropanoid biosynthetic process	7	141	8.10E-07
GO:0010033	Response to organic substance	18	1342	8.90E-07

Table 5-22: Genes upregulated by the decrease in *AtIPCS1* expression under GO term: 0006950 (response to stress)

		At1 tDNA mutant	
Gene ID	**Gene annotation**	**Log₂ fold-change**	**p-value**
AT1G17600	DISEASE RESISTANCE PROTEIN (TIR-NBS-LRR CLASS)	-1.2	4.95E-06
AT1G16420	MC8 (METACASPASE 8)	-1.6	3.10E-06
AT2G33380	RD20 (RESPONSIVE TO DESSICATION 20)	-1.2	1.77E-05
AT3G13650	DIR7	-1.1	1.69E-06
AT2G30770	CYP71A13	**-2.1**	**3.61E-10**
AT2G40750	WRKY54	**-2.2**	**8.75E-31**
AT3G52430	PAD4 (PHYTOALEXIN DEFICIENT 4)	-1.1	6.69E-16
AT1G65870	DIR21	-1.4	9.42E-05
AT1G75040	PR5 (PATHOGENESIS-RELATED GENE 5)	**-2.4**	**1.80E-47**
AT1G72920	DISEASE RESISTANCE PROTEIN (TIR-NBS CLASS),	-1.1	1.30E-05
AT2G02990	RNS1 (RIBONUCLEASE 1)	-1.1	2.41E-09
AT1G61120	TPS04 (TERPENE SYNTHASE 04)	**-2.9**	**3.77E-19**
AT1G19670	ATCLH1 (ARABIDOPSIS THALIANA CORONATINE-INDUCED PROTEIN 1)	-1.7	1.14E-78
AT2G41240	BHLH100 (BASIC HELIX-LOOP-HELIX PROTEIN 100)	**-3.3**	**1.22E-44**
AT5G07690	ATMYB29 (ARABIDOPSIS THALIANA MYB DOMAIN PROTEIN 29)	**-2.1**	**1.56E-13**
AT1G47890	ATRLP7 (RECEPTOR LIKE PROTEIN 7)	-1.1	2.42E-05
AT2G35380	PEROXIDASE 20 (PER20)	-1.2	3.51E-18
AT5G24770	VSP2 (VEGETATIVE STORAGE PROTEIN 2)	-1.3	2.86E-13
AT1G65060	4-COUMARATE:COA LIGASE (4CL)	-1.1	3.64E-06
AT3G45140	LOX2 (LIPOXYGENASE 2)	-1.7	5.15E-51
AT4G17090	CT-BMY (CHLOROPLAST BETA-AMYLASE)	**-2.6**	**1.23E-243**
AT1G78410	VQ MOTIF-CONTAINING PROTEIN 10	-1.3	4.87E-05
AT3G56400	WRKY70	-1.6	6.27E-68
AT1G00450	NIMIN1 (NIM1-INTERACTING 1)	**-3.2**	**1.019475205**
AT1G18710	ATMYB47 (MYB DOMAIN PROTEIN 47)	-1.7	1.07E-19
AT1G19250	FMO1 (FLAVIN-DEPENDENT MONOOXYGENASE 1)	-1.6	3.19E-10
AT2G41480	PEROXIDASE 25	-1.7	1.39E-20
AT3G22840	ELIP1 (EARLY LIGHT-INDUCABLE PROTEIN)	-1.0	0.002586795
AT1G57630	DISEASE RESISTANCE PROTEIN (TIR CLASS)	-1.1	1.44E-05
AT1G68740	PHO1	-1.2	2.25E-09
AT1G55210	DISEASE RESISTANCE RESPONSE	-1.8	3.05E-45
AT5G61420	MYB28 (MYB DOMAIN PROTEIN 28)	**-2.1**	**2.52E-64**

DIR proteins are involved in the plant defence response to pathogens in *Arabidopsis* and *At*IPCS functions as a negative regulator. These data showed that diverse components of the plant defence machinery, regulating the response to pathogens, are down regulated in response to changes in *At*IPCS expression.

In relation to abiotic stress, two genes were notably down regulated in response to *At*IPCS1 knock-down: *CT-BMY* and *BHLH10*. *CT-BMY* encodes a beta amylase found in the chloroplast and induced in response to cold stress. In *Arabidopsis*, knock-down of CT-BMY was found to lead to reduced photosynthetic efficiency, suggesting that its accumulation might serve to increase plant freezing tolerance by the breakdown of glucans resulting in the accumulation of maltose and other oligosaccharides which act as protective solutes in response to stress[88]. These sugars act as 'compatible solutes' for the continued function of macromolecules such as proteins and enzymes when plants are exposed to freezing temperatures by aiding water retention in cells[89], thereby reducing the risk of cell lysis when the osmotic flow favours the movement of water molecules into the cell. *BHLH100* was downregulated by $\log_2$ fold-change of 3.3 in the insertion mutant; overexpression of the basic-helix-helix-loop has been shown to increase plant tolerance to high zinc (Zn) and nickel (Ni)[90]. This suggested that *AtIPCS1* expression levels may affect plant metal homeostasis by acting as a negative regulator of response to metal stress.

5.7 Genes upregulated in response to a decrease in *AtIPCS1* expression

Analyses of genes upregulated in response to the decrease in *AtIPCS1* expression found them to be significantly enriched under the GO terms photosynthesis and lipid localisation (Table 5-23, 5-24 and 5-25). Further examination of genes upregulated under the GO term

Table 5-23: GO enrichment of genes upregulated in response to the knock down in expression of _AtIPCS1_ tDNA insert muntant

GO term	Description	Number in input list	Number in Background	p-value
GO:0050896	Response to stimulus	95	4057	1.60E-36
GO:0010033	Response to organic substance	59	1342	3.60E-35
GO:0042221	Response to chemical stimulus	70	2085	5.40E-35
GO:0009719	Response to endogenous stimulus	47	1068	7.90E-28
GO:0009725	Response to hormone stimulus	44	982	2.50E-26
GO:0015979	Photosynthesis	20	162	2.90E-20
GO:0009628	Response to abiotic stimulus	38	1471	4.70E-15
GO:0006950	Response to stress	45	2320	2.30E-13
GO:0019684	Photosynthesis, light reaction	12	103	2.20E-12
GO:0070887	Cellular response to chemical stimulus	20	452	2.70E-12
GO:0010876	Lipid localization	8	24	7.20E-12
GO:0009743	Response to carbohydrate stimulus	15	240	1.60E-11
GO:0009873	Ethylene mediated signaling pathway	10	74	4.20E-11
GO:0009266	Response to temperature stimulus	19	485	7.10E-11
GO:0009755	Hormone-mediated signaling pathway	16	321	8.30E-11
GO:0032870	Cellular response to hormone stimulus	16	321	8.30E-11
GO:0006091	Generation of precursor metabolites and energy	15	285	1.60E-10
GO:0000160	Two-component signal transduction system (phosphor-relay)	10	92	2.90E-10
GO:0009987	Cellular process	107	11684	1.40E-09
GO:0010200	Response to chitin	11	151	2.00E-09
GO:0009733	Response to auxin stimulus	15	360	3.40E-09
GO:0009642	Response to light intensity	9	90	4.60E-09
GO:0044237	Cellular metabolic process	86	8722	7.40E-09
GO:0007242	Intracellular signaling cascade	19	659	9.30E-09
GO:0051716	Cellular response to stimulus	21	840	1.70E-08
GO:0009723	Response to ethylene stimulus	11	199	2.90E-08
GO:0009605	Response to external stimulus	15	429	3.20E-08
GO:0009416	Response to light stimulus	17	596	6.70E-08
GO:0009646	Response to absence of light	5	16	9.30E-08
GO:0009314	Response to radiation	17	613	9.90E-08
GO:0032774	RNA biosynthetic process	22	1040	1.30E-07
GO:0006351	Transcription, DNA-dependent	22	1039	1.30E-07
GO:0009611	Response to wounding	10	197	2.70E-07
GO:0009968	Negative regulation of signal transduction	6	43	3.10E-07
GO:0009737	Response to abscisic acid stimulus	13	378	3.30E-07
GO:0009409	Response to cold	12	328	5.00E-07

Table 5-24: Genes upregulated by the decrease in expression of _AtIPCS1_ under GO term: 0015979 (photosynthesis)

		At1 tDNA mutant	
Gene ID	**Gene annotation**	**Log$_2$ fold-change**	**p-value**
ATCG00730	SUBUNIT IV OF THE CYTOCHROME B6/F COMPLEX PETD	1.4	1.23E-05
AT1G29920	CAB2 (CHLOROPHYLL A/B-BINDING PROTEIN 2)	1.2	3.18E-12
ATCG00580	PSII CYTOCHROME B559	1.3	8.64E-05
ATCG01100	NADH DEHYDROGENASE ND1 NDHA	1.2	0.000267999
ATCG00520	PHOTOSYSTEM I ASSEMBLY PROTEIN YCF4	1.1	0.001040999
ATCG01090	SUBUNIT OF THE CHLOROPLAST NAD(P)H DEHYDROGENASE COMPLEX NDHI	1.1	0.001370433
ATCG00270	PSII D2 PROTEIN PSBD;	1.5	2.44E-06
ATCG00680	CP47, SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	1.2	
ATCG00710	PHOTOSYSTEM II REACTION CENTER PROTEIN H	1.2	0.000212549
ATCG00360	PHOTOSYSTEM I ASSEMBLY PROTEIN YCF3	1.1	0.001002877
ATCG00300	PSBZ, WHICH IS A SUBUNIT OF PHOTOSYSTEM II	1.1	0.002217157
ATCG00440	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT 3	1.0	0.002233798
ATCG00540	CYTOCHROME F APOPROTEIN	1.1	0.000424669
ATCG00340	ENCODES THE D1 SUBUNIT OF PHOTOSYSTEM I AND II REACTION CENTERS. PSAB	1.4	1.44E-05
ATCG01110	NAD(P)H-QUINONE OXIDOREDUCTASE SUBUNIT H	1.9	6.32E-10
AT5G01600	ATFER1	1.1	1.09E-67
ATCG01060	PSAC SUBUNIT OF PHOTOSYSTEM I	1.4	1.56E-05
ATCG00720	CYTOCHROME B(6) SUBUNIT OF THE CYTOCHROME B6F COMPLEX	1.2	9.76E-05
ATCG00280	CP43 SUBUNIT OF THE PHOTOSYSTEM II REACTION CENTER	1.4	8.36E-06
AT1G27730	STZ (SALT TOLERANCE ZINC FINGER)	1.2	4.52E-17

Table 5-25: Genes upregulated by the decrease in expression of _AtIPCS1_ under GO term: 0010876 (lipid localization)

| | | _At1 tDNA mutant_ | |
Gene ID	Gene annotation	Log$_2$ fold-change	p-value
AT5G40420	OLEO2 (OLEOSIN 2)	1.3	3.97E-06
AT4G25140	OLEO1 (OLEOSIN 1)	**2.1**	**4.96E-15**
AT4G27150	2S SEED STORAGE PROTEIN 2	1.6	1.52E-07
AT3G01570	OLE05 (OLEOSIN 5)	1.3	4.17E-09
AT5G54740	2S SEED STORAGE PROTEIN 5	**3.6**	**6.33E-48**
AT4G24230	ACBP3 (ACYL-COA-BINDING DOMAIN 3)	1.3	6.87E-101
AT3G27660	OLEO4 (OLEOSIN 4)	1.3	5.59E-06
AT4G27160	2S SEED STORAGE PROTEIN 3	1.1	0.000254456

photosynthesis found the same genes observed as up-regulated in response to the overexpression of *AtIPCS1*, *2* and *3*. Only one gene, *STZ*, was identified that had not been found in the overexpressing transgenic lines. STZ has a function in plant response to abiotic stress (Table 5-24) and have been shown to enhance tolerance to heat, salinity and osmotic stress in transgenic over-expressors and RNAi knock-downs[91]. Therefore, the downregulation of *STZ* in response to *AtIPCS1* indicates that *AtIPCS* affects the plant response to abiotic stress in a way that is advantageous for plant survival.

Genes up regulated under the GO enrichment term, lipid localization were found to encode the lipid transfer proteins and lipid storage proteins already identified as upregulated in response to *AtIPCS* overexpression. The only genes upregulated in response to overexpression of *AtIPCS1*, *2* and *3*, but not upregulated on *AtIPCS1* knock-down were, *2S seed storage protein 1* and *4*. These data indicated that *AtIPCS* acts as a positive and negative regulator of lipid content in plants.

5.8 Conclusion

In this Chapter, the role of *At*IPCS as a negative regulator of plant defence response against pathogens was indicated. The overexpression of each isoform resulted in the down regulation of genes involved in both JA and SA plant defence signalling, and these responded in a dose dependent manner to changes in the expression of each isoform. The PR genes are of particular interest because they are well characterized plant defence signals. In this study, it was found that overexpression of *AtIPCS1* and *AtIPCS3* down regulated these genes in a dose dependent manner. *AtIPCS2* did not show the same response; what was observed instead, was a heightened response resulting in an increase in the down regulation of these genes compared to over-expressors of *AtIPCS1* and *AtIPCS3*. The down regulation of these genes was also found to have reached a peak; the

same level of down regulation was observed for both At2 over-expressor transgenic lines, despite the differences in *AtIPCS2* transcript levels.

In addition, the overexpression of *AtIPCS2* levels affected more genes in the defence response pathway that the other isoforms. From these data, it was clear that *At*IPCS2 is one of the principle regulators of plant defence response against pathogens, and that the other two isoforms have similar roles but perhaps as rheostats. The knock-down in *AtIPCS1* expression was also found to affect the plant defence pathway, resulting in the down regulation of the disease resistance TIR class of proteins, DIR proteins, *WRKY 70* and *PADA*. Interestingly, only *PR5* (out of all the other PR genes: *PR1*, *PR2* and *PR4*) was down regulated in response to *At*IPCS1 transcript decrease. In contrast, *PR1* and *PR2* in particular were found to show a heightened level of response to the overexpression of *AtIPCS*, especially *AtIPCS2*. This indicated whilst *PR1* and *PR2* are specifically induced in response to an increase in *AtIPCS* expression, *PR5* is induced as a basal response to any change in expression.

*At*IPCS has a prominent role as a regulator of plant defence signalling, affecting not only pathways involved in response to fungal and bacterial pathogens, but also pathways involved in defence against parasites and herbivores. This showed that *At*IPCS is part of a network of plant defence genes that transcends different defence pathways. This is probably because any change in the levels of phytoceramide, acts as an indiscriminate pro-apoptotic signal that mobilizes all defence networks in the push for plant survival.

AtIPCS levels also affect the plant response to abiotic stress; overexpression of *AtIPCS1*, *2* and *3*, and knockdown of *AtIPCS1* resulted in the down regulation of genes involved in the response to salt, drought and cold stress. This is in alignment with the identification of transcription factor binding sites for genes involved in response to abiotic stress

upstream of the *At*IPCS1, 2 and 3 promotors (Chapter 2). The down regulation of these genes could result in plants that are tolerant to cold, salt and osmotic stress.

Similarly, in response to the overexpression of *AtIPCS1*, *2* and *3*, and knock-down of *AtIPCS1*, genes involved in photosynthesis and lipid localisation were upregulated. The link between IPCS expression and photosynthesis is unclear and could be an indirect effect. In relation to lipid localisation, the genes upregulated are those that have been identified as expressed specifically in seeds. The level of upregulation for these genes was similar in response to the overexpression of each isoform of *AtIPCS*. This indicated that the three isoforms have the same level of regulation of genes involved in lipid localization in seeds, and function to increase lipid content in seeds to ensure they reach maturity and germinate.

In conclusion, the RNASeq analyses have shown that changes in *AtIPCS* expression is a focal point for the induction of the plant defence response to biotic and abiotic stress, with other effects on plant photosynthesis efficiency and lipid content.

5.9 References

1. Liao, P.; Huang, J.; Tong, P.; Nie, W.; Yan, X.; Feng, Y.; Peng, H.; Peng, X.; Li, S., Characterization and expression analysis of inositolphosphorylceramide synthase family genes in rice (Oryza sativa L.). *Genes & Genomics* **2017,** *39* (5), 485-492.
2. Wang, W.; Yang, X.; Tangchaiburana, S.; Ndeh, R.; Markham, J. E.; Tsegaye, Y.; Dunn, T. M.; Wang, G.-L.; Bellizzi, M.; Parsons, J. F., An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in Arabidopsis. *The Plant Cell* **2008,** *20* (11), 3163-3179.
3. Mina, J.; Okada, Y.; Wansadhipathi-Kannangara, N.; Pratt, S.; Shams-Eldin, H.; Schwarz, R.; Steel, P.; Fawcett, T.; Denny, P., Functional analyses of differentially expressed isoforms of the Arabidopsis inositol phosphorylceramide synthase. *Plant molecular biology* **2010,** *73* (4-5), 399-407.
4. Ashburner, M.; Ball, C. A.; Blake, J. A.; Botstein, D.; Butler, H.; Cherry, J. M.; Davis, A. P.; Dolinski, K.; Dwight, S. S.; Eppig, J. T., Gene Ontology: tool for the unification of biology. *Nature genetics* **2000,** *25* (1), 25.
5. Tipney, H.; Hunter, L., An introduction to effective use of enrichment analysis software. *Human genomics* **2010,** *4* (3), 202.

6.	Penninckx, I.; Eggermont, K.; Terras, F.; Thomma, B.; De Samblanx, G. W.; Buchala, A.; Metraux, J.-P.; Manners, J. M.; Broekaert, W. F., Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. *The Plant Cell* **1996,** *8* (12), 2309-2323.

7.	Bonasera, J. M.; Kim, J. F.; Beer, S. V., PR genes of apple: identification and expression in response to elicitors and inoculation with Erwinia amylovora. *BMC Plant Biology* **2006,** *6* (1), 23.

8.	Iglesias, M. J.; Terrile, M. C.; Casalongué, C. A., Auxin and salicylic acid signalings counteract the regulation of adaptive responses to stress. *Plant signaling & behavior* **2011,** *6* (3), 452-454.

9.	Zhang, J.; Du, X.; Wang, Q.; Chen, X.; Lv, D.; Xu, K.; Qu, S.; Zhang, Z., Expression of pathogenesis related genes in response to salicylic acid, methyl jasmonate and 1-aminocyclopropane-1-carboxylic acid in Malus hupehensis (Pamp.) Rehd. *BMC research notes* **2010,** *3* (1), 208.

10.	Liu, W.-X.; Zhang, F.-C.; Zhang, W.-Z.; Song, L.-F.; Wu, W.-H.; Chen, Y.-F., Arabidopsis Di19 functions as a transcription factor and modulates PR1, PR2, and PR5 expression in response to drought stress. *Molecular plant* **2013,** *6* (5), 1487-1502.

11.	Lopukhina, A.; Dettenberg, M.; Weiler, E. W.; Holländer-Czytko, H., Cloning and characterization of a coronatine-regulated tyrosine aminotransferase from Arabidopsis. *Plant physiology* **2001,** *126* (4), 1678-1687.

12.	Yu, K.; Soares, J. M.; Mandal, M. K.; Wang, C.; Chanda, B.; Gifford, A. N.; Fowler, J. S.; Navarre, D.; Kachroo, A.; Kachroo, P., A feedback regulatory loop between G3P and lipid transfer proteins DIR1 and AZI1 mediates azelaic-acid-induced systemic immunity. *Cell Reports* **2013,** *3* (4), 1266-1278.

13.	Chanda, B.; Xia, Y.; Mandal, M. K.; Yu, K.; Sekine, K. T.; Gao, Q.-m.; Selote, D.; Hu, Y.; Stromberg, A.; Navarre, D., Glycerol-3-phosphate is a critical mobile inducer of systemic immunity in plants. *Nature genetics* **2011,** *43* (5), 421-427.

14.	Gupta, V.; Willits, M. G.; Glazebrook, J., Arabidopsis thaliana EDS4 contributes to salicylic acid (SA)-dependent expression of defense responses: evidence for inhibition of jasmonic acid signaling by SA. *Molecular Plant-Microbe Interactions* **2000,** *13* (5), 503-511.

15.	Petersen, M.; Brodersen, P.; Naested, H.; Andreasson, E.; Lindhart, U.; Johansen, B.; Nielsen, H. B.; Lacy, M.; Austin, M. J.; Parker, J. E., Arabidopsis MAP kinase 4 negatively regulates systemic acquired resistance. *Cell* **2000,** *103* (7), 1111-1120.

16.	Glazebrook, J., Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.* **2005,** *43*, 205-227.

17.	Yoshioka, K.; Kachroo, P.; Tsui, F.; Sharma, S. B.; Shah, J.; Klessig, D. F., Environmentally sensitive, SA- dependent defense responses in the cpr22 mutant of Arabidopsis. *The Plant Journal* **2001,** *26* (4), 447-459.

18.	Pandey, S. P.; Somssich, I. E., The role of WRKY transcription factors in plant immunity. *Plant physiology* **2009,** *150* (4), 1648-1655.

19.	Wang, D.; Amornsiripanitch, N.; Dong, X., A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS pathogens* **2006,** *2* (11), e123.

20.	Li, J.; Besseau, S.; Törönen, P.; Sipari, N.; Kollist, H.; Holm, L.; Palva, E. T., Defense- related transcription factors WRKY70 and WRKY54 modulate osmotic stress tolerance by regulating stomatal aperture in Arabidopsis. *New Phytologist* **2013,** *200* (2), 457-472.

21.	Besseau, S.; Li, J.; Palva, E. T., WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in Arabidopsis thaliana. *Journal of experimental botany* **2012,** *63* (7), 2667-2679.

22.   Greb, T.; Mylne, J. S.; Crevillen, P.; Geraldo, N.; An, H.; Gendall, A. R.; Dean, C., The PHD finger protein VRN5 functions in the epigenetic silencing of Arabidopsis FLC. *Current biology* **2007,** *17* (1), 73-78.

23.   Henderson, I. R.; Dean, C., Control of Arabidopsis flowering: the chill before the bloom. *Development* **2004,** *131* (16), 3829-3838.

24.   Berrocal-Lobo, M.; Molina, A., Arabidopsis defense response against Fusarium oxysporum. *Trends in plant science* **2008,** *13* (3), 145-150.

25.   Wolter, M.; Hollricher, K.; Salamini, F.; Schulze-Lefert, P., The mlo resistance alleles to powdery mildew infection in barley trigger a developmentally controlled defence mimic phenotype. *Molecular and General Genetics MGG* **1993,** *239* (1), 122-128.

26.   Matthes, M. C.; Bruce, T. J.; Ton, J.; Verrier, P. J.; Pickett, J. A.; Napier, J. A., The transcriptome of cis-jasmone-induced resistance in Arabidopsis thaliana and its role in indirect defence. *Planta* **2010,** *232* (5), 1163-1180.

27.   Hung, C.-Y.; Aspesi Jr, P.; Hunter, M. R.; Lomax, A. W.; Perera, I. Y., Phosphoinositide-signaling is one component of a robust plant defense response. *Frontiers in plant science* **2014,** *5*.

28.   La Camera, S.; Geoffroy, P.; Samaha, H.; Ndiaye, A.; Rahim, G.; Legrand, M.; Heitz, T., A pathogen- inducible patatin- like lipid acyl hydrolase facilitates fungal and bacterial host colonization in Arabidopsis. *The Plant Journal* **2005,** *44* (5), 810-825.

29.   Wisniewski, J.-P.; Cornille, P.; Agnel, J.-P.; Montillet, J.-L., The extensin multigene family responds differentially to superoxide or hydrogen peroxide in tomato cell cultures. *FEBS letters* **1999,** *447* (2-3), 264-268.

30.   Attaran, E.; Rostás, M.; Zeier, J., Pseudomonas syringae elicits emission of the terpenoid (E, E)-4, 8, 12-trimethyl-1, 3, 7, 11-tridecatetraene in Arabidopsis leaves via jasmonate signaling and expression of the terpene synthase TPS4. *Molecular plant-microbe interactions* **2008,** *21* (11), 1482-1497.

31.   Bu, Q.; Jiang, H.; Li, C.-B.; Zhai, Q.; Zhang, J.; Wu, X.; Sun, J.; Xie, Q.; Li, C., Role of the Arabidopsis thaliana NAC transcription factors ANAC019 and ANAC055 in regulating jasmonic acid-signaled defense responses. *Cell research* **2008,** *18* (7), 756.

32.   Tran, L.-S. P.; Nakashima, K.; Sakuma, Y.; Simpson, S. D.; Fujita, Y.; Maruyama, K.; Fujita, M.; Seki, M.; Shinozaki, K.; Yamaguchi-Shinozaki, K., Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *The Plant Cell* **2004,** *16* (9), 2481-2498.

33.   Lieberherr, D.; Wagner, U.; Dubuis, P.-H.; Métraux, J.-P.; Mauch, F., The rapid induction of glutathione S-transferases AtGSTF2 and AtGSTF6 by avirulent Pseudomonas syringae is the result of combined salicylic acid and ethylene signaling. *Plant and cell physiology* **2003,** *44* (7), 750-757.

34.   Acharya, B. R.; Raina, S.; Maqbool, S. B.; Jagadeeswaran, G.; Mosher, S. L.; Appel, H. M.; Schultz, J. C.; Klessig, D. F.; Raina, R., Overexpression of CRK13, an Arabidopsis cysteine- rich receptor- like kinase, results in enhanced resistance to Pseudomonas syringae. *The Plant Journal* **2007,** *50* (3), 488-499.

35.   Cao, S.; Ye, M.; Jiang, S., Involvement of GIGANTEA gene in the regulation of the cold stress response in Arabidopsis. *Plant cell reports* **2005,** *24* (11), 683-690.

36.   Ascenzi, R.; Gantt, J. S., A drought-stress-inducible histone gene in Arabidopsis thaliana is a member of a distinct class of plant linker histone variants. *Plant molecular biology* **1997,** *34* (4), 629-641.

37.   Mizoguchi, M.; Umezawa, T.; Nakashima, K.; Kidokoro, S.; Takasaki, H.; Fujita, Y.; Yamaguchi-Shinozaki, K.; Shinozaki, K., Two closely related subclass II SnRK2

protein kinases cooperatively regulate drought-inducible gene expression. *Plant and Cell Physiology* **2010,** *51* (5), 842-847.

38.	Chen, H.; Lai, Z.; Shi, J.; Xiao, Y.; Chen, Z.; Xu, X., Roles of Arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC plant biology* **2010,** *10* (1), 281.

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

40.	Kuno, N.; Møller, S. G.; Shinomura, T.; Xu, X.; Chua, N.-H.; Furuya, M., The novel MYB protein EARLY-PHYTOCHROME-RESPONSIVE1 is a component of a slave circadian oscillator in Arabidopsis. *The Plant Cell* **2003,** *15* (10), 2476-2488.

41.	Mizoguchi, T.; Wright, L.; Fujiwara, S.; Cremer, F.; Lee, K.; Onouchi, H.; Mouradov, A.; Fowler, S.; Kamada, H.; Putterill, J., Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. *The Plant Cell* **2005,** *17* (8), 2255-2270.

42.	Simpson, G. G., The autonomous pathway: epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. *Current opinion in plant biology* **2004,** *7* (5), 570-574.

43.	Sheldon, C. C.; Rouse, D. T.; Finnegan, E. J.; Peacock, W. J.; Dennis, E. S., The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences* **2000,** *97* (7), 3753-3758.

44.	Michaels, S. D.; Amasino, R. M., Loss of FLOWERING LOCUS C activity eliminates the late-flowering phenotype of FRIGIDA and autonomous pathway mutations but not responsiveness to vernalization. *The Plant Cell* **2001,** *13* (4), 935-941.

45.	Johanson, U.; West, J.; Lister, C.; Michaels, S.; Amasino, R.; Dean, C., Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science* **2000,** *290* (5490), 344-347.

46.	Michaels, S. D.; Amasino, R. M., FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell* **1999,** *11* (5), 949-956.

47.	Ambawat, S.; Sharma, P.; Yadav, N. R.; Yadav, R. C., MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants* **2013,** *19* (3), 307-321.

48.	Knoth, C.; Eulgem, T., The oomycete response gene LURP1 is required for defense against Hyaloperonospora parasitica in Arabidopsis thaliana. *The Plant Journal* **2008,** *55* (1), 53-64.

49.	Kim, S. A.; Kwak, J.; Jae, S.-K.; Wang, M.-H.; Nam, H., Overexpression of the AtGluR2 Gene Encoding anArabidopsis Homolog of Mammalian Glutamate Receptors Impairs CalciumUtilization and Sensitivity to Ionic Stress in TransgenicPlants. *Plant and Cell Physiology* **2001,** *42* (1), 74-84.

50.	Yang, Y.; Zhang, Y.; Ding, P.; Johnson, K.; Li, X.; Zhang, Y., The ankyrin-repeat transmembrane protein BDA1 functions downstream of the receptor-like protein SNC2 to regulate plant immunity. *Plant physiology* **2012,** *159* (4), 1857-1865.

51.	Holm, M.; Ma, L.-G.; Qu, L.-J.; Deng, X.-W., Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. *Genes & Development* **2002,** *16* (10), 1247-1259.

52.	Liu, C.; Wang, T.; Zhang, W.; Li, X., Computational identification and analysis of immune-associated nucleotide gene family in Arabidopsis thaliana. *Journal of plant physiology* **2008,** *165* (7), 777-787.

53.	Li, J.; Brader, G.; Kariola, T.; Tapio Palva, E., WRKY70 modulates the selection of signaling pathways in plant defense. *The Plant Journal* **2006,** *46* (3), 477-491.

54.   Lu, H.; Rate, D. N.; Song, J. T.; Greenberg, J. T., ACD6, a novel ankyrin protein, is a regulator and an effector of salicylic acid signaling in the Arabidopsis defense response. *The Plant Cell* **2003,** *15* (10), 2408-2420.

55.   Peng, L.; Fukao, Y.; Myouga, F.; Motohashi, R.; Shinozaki, K.; Shikanai, T., A chaperonin subunit with unique structures is essential for folding of a specific substrate. *PLoS biology* **2011,** *9* (4), e1001040.

56.   Becerra, C.; Puigdomenech, P.; Vicient, C. M., Computational and experimental analysis identifies Arabidopsis genes specifically expressed during early seed development. *BMC genomics* **2006,** *7* (1), 38.

57.   Shimada, T. L.; Shimada, T.; Takahashi, H.; Fukao, Y.; Hara- Nishimura, I., A novel role for oleosins in freezing tolerance of oilseeds in Arabidopsis thaliana. *The Plant Journal* **2008,** *55* (5), 798-809.

58.   Siloto, R. M.; Findlay, K.; Lopez-Villalobos, A.; Yeung, E. C.; Nykiforuk, C. L.; Moloney, M. M., The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. *The Plant Cell* **2006,** *18* (8), 1961-1974.

59.   Youle, R. J.; Huang, A. H., Occurrence of low molecular weight and high cysteine containing albumin storage proteins in oilseeds of diverse species. *American Journal of Botany* **1981**, 44-48.

60.   Krebbers, E.; Herdies, L.; De Clercq, A.; Seurinck, J.; Leemans, J.; Van Damme, J.; Segura, M.; Gheysen, G.; Van Montagu, M.; Vandekerckhove, J., Determination of the processing sites of an Arabidopsis 2S albumin and characterization of the complete gene family. *Plant Physiology* **1988,** *87* (4), 859-866.

61.   Hundertmark, M.; Hincha, D. K., LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC genomics* **2008,** *9* (1), 118.

62.   Nguyen, T.-P.; Cueff, G.; Hegedus, D. D.; Rajjou, L.; Bentsink, L., A role for seed storage proteins in Arabidopsis seed longevity. *Journal of experimental botany* **2015,** *66* (20), 6399-6413.

63.   Poxleitner, M.; Rogers, S. W.; Lacey Samuels, A.; Browse, J.; Rogers, J. C., A role for caleosin in degradation of oil- body storage lipid during seed germination. *The Plant Journal* **2006,** *47* (6), 917-933.

64.   Dubreucq, B.; Berger, N.; Vincent, E.; Boisson, M.; Pelletier, G.; Caboche, M.; Lepiniec, L., The Arabidopsis AtEPR1 extensin- like gene is specifically expressed in endosperm during seed germination. *The Plant Journal* **2000,** *23* (5), 643-652.

65.   Manfre, A. J.; Lanni, L. M.; Marcotte, W. R., The Arabidopsis group 1 LATE EMBRYOGENESIS ABUNDANT protein ATEM6 is required for normal seed development. *Plant physiology* **2006,** *140* (1), 140-149.

66.   Haslekås, C.; Viken, M. K.; Grini, P. E.; Nygaard, V.; Nordgard, S. H.; Meza, T. J.; Aalen, R. B., Seed 1-cysteine peroxiredoxin antioxidants are not involved in dormancy, but contribute to inhibition of germination during stress. *Plant physiology* **2003,** *133* (3), 1148-1157.

67.   Murgia, I.; Vazzola, V.; Tarantino, D.; Cellier, F.; Ravet, K.; Briat, J.-F.; Soave, C., Knock-out of ferritin AtFer1 causes earlier onset of age-dependent leaf senescence in Arabidopsis. *Plant Physiology and Biochemistry* **2007,** *45* (12), 898-907.

68.   Raynal, M.; Guilleminot, J.; Gueguen, C.; Cooke, R.; Delseny, M.; Gruber, V., Structure, organization and expression of two closely related novel Lea (late-embryogenesis-abundant) genes in Arabidopsis thaliana. *Plant molecular biology* **1999,** *40* (1), 153-165.

69.   Favory, J.-J.; Kobayshi, M.; Tanaka, K.; Peltier, G.; Kreis, M.; Valay, J.-G.; Lerbs-Mache, S., Specific function of a plastid sigma factor for ndh F gene transcription. *Nucleic acids research* **2005,** *33* (18), 5991-5999.

70.   Dobáková, M.; Sobotka, R.; Tichý, M.; Komenda, J., Psb28 protein is involved in the biogenesis of the photosystem II inner antenna CP47 (PsbB) in the cyanobacterium Synechocystis sp. PCC 6803. *Plant physiology* **2009,** *149* (2), 1076-1086.

71.   Lefebvre, S.; Lawson, T.; Fryer, M.; Zakhleniuk, O. V.; Lloyd, J. C.; Raines, C. A., Increased sedoheptulose-1, 7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. *Plant Physiology* **2005,** *138* (1), 451-460.

72.   Yabuta, S.; Ifuku, K.; Takabayashi, A.; Ishihara, S.; Ido, K.; Ishikawa, N.; Endo, T.; Sato, F., Three PsbQ-like proteins are required for the function of the chloroplast NAD (P) H dehydrogenase complex in Arabidopsis. *Plant and cell physiology* **2010,** *51* (6), 866-876.

73.   Peng, L.; Fukao, Y.; Fujiwara, M.; Takami, T.; Shikanai, T., Efficient operation of NAD (P) H dehydrogenase requires supercomplex formation with photosystem I via minor LHCI in Arabidopsis. *The Plant Cell* **2009,** *21* (11), 3623-3640.

74.   Jouannic, N.; Lepetit, M.; Vergnolle, C.; Cantrel, C.; Gardies, A. M.; Kader, J. C.; Arondel, V., Isolation of a cDNA from Arabidopsis thaliana that complements the sec14 mutant of yeast. *The FEBS Journal* **1998,** *258* (2), 402-410.

75.   Maldonado, A. M.; Doerner, P.; Dixon, R. A.; Lamb, C. J.; Cameron, R. K., A putative lipid transfer protein involved in systemic resistance signalling in Arabidopsis. *Nature* **2002,** *419* (6905), 399.

76.   Sottosanto, J. B.; Saranga, Y.; Blumwald, E., Impact of AtNHX1, a vacuolar Na+/H+ antiporter, upon gene expression during short-and long-term salt stress in Arabidopsis thaliana. *BMC Plant Biology* **2007,** *7* (1), 18.

77.   Jang, C. S.; Lee, H. J.; Chang, S. J.; Seo, Y. W., Expression and promoter analysis of the TaLTP1 gene induced by drought and salt stress in wheat (Triticum aestivum L.). *Plant Science* **2004,** *167* (5), 995-1001.

78.   Weigel, R. R.; Bäuscher, C.; Pfitzner, A. J.; Pfitzner, U. M., NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from Arabidopsis that interact with NPR1/NIM1, a key regulator of systemic acquired resistance in plants. *Plant molecular biology* **2001,** *46* (2), 143-160.

79.   Hermann, M.; Maier, F.; Masroor, A.; Hirth, S.; Pfitzner, A. J.; Pfitzner, U. M., The Arabidopsis NIMIN proteins affect NPR1 differentially. *Frontiers in plant science* **2013,** *4*.

80.   Bell, E.; Creelman, R. A.; Mullet, J. E., A chloroplast lipoxygenase is required for wound-induced jasmonic acid accumulation in Arabidopsis. *Proceedings of the National Academy of Sciences* **1995,** *92* (19), 8675-8679.

81.   Pandey, S. P.; Roccaro, M.; Schön, M.; Logemann, E.; Somssich, I. E., Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *The Plant Journal* **2010,** *64* (6), 912-923.

82.   Bartsch, M.; Gobbato, E.; Bednarek, P.; Debey, S.; Schultze, J. L.; Bautor, J.; Parker, J. E., Salicylic acid–independent ENHANCED DISEASE SUSCEPTIBILITY1 signaling in Arabidopsis immunity and cell death is regulated by the monooxygenase FMO1 and the nudix hydrolase NUDT7. *The Plant Cell* **2006,** *18* (4), 1038-1051.

83.   DeYoung, B. J.; Innes, R. W., Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature immunology* **2006,** *7* (12), 1243.

84.   Sønderby, I. E.; Burow, M.; Rowe, H. C.; Kliebenstein, D. J.; Halkier, B. A., A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in Arabidopsis. *Plant Physiology* **2010,** *153* (1), 348-363.

85.   Lazzeri, L.; Curto, G.; Leoni, O.; Dallavalle, E., Effects of glucosinolates and their enzymatic hydrolysis products via myrosinase on the root-knot nematode Meloidogyne

incognita (Kofoid et White) Chitw. *Journal of Agricultural and Food Chemistry* **2004,** *52* (22), 6703-6707.

86.   Berger, S.; Mitchell- Olds, T.; Stotz, H. U., Local and differential control of vegetative storage protein expression in response to herbivore damage in Arabidopsis thaliana. *Physiologia Plantarum* **2002,** *114* (1), 85-91.

87.   Ponce De Leon, I.; Schmelz, E. A.; Gaggero, C.; Castro, A.; Alvarez, A.; Montesano, M., Physcomitrella patens activates reinforcement of the cell wall, programmed cell death and accumulation of evolutionary conserved defence signals, such as salicylic acid and 12- oxo- phytodienoic acid, but not jasmonic acid, upon Botrytis cinerea infection. *Molecular plant pathology* **2012,** *13* (8), 960-974.

88.   Kaplan, F.; Guy, C. L., RNA interference of Arabidopsis beta- amylase8 prevents maltose accumulation upon cold shock and increases sensitivity of PSII photochemical efficiency to freezing stress. *The Plant Journal* **2005,** *44* (5), 730-743.

89.   Yancey, P. H.; Clark, M. E.; Hand, S. C.; Bowlus, R. D.; Somero, G. N., Living with water stress: evolution of osmolyte systems. *Science* **1982,** *217* (4566), 1214-1222.

90.   van de Mortel, J. E.; Schat, H.; Moerland, P. D.; van Themaat, E. V. L.; Van Der Ent, S.; Blankestijn, H.; Ghandilyan, A.; Tsiatsiani, S.; AARTS, M. G., Expression differences for genes involved in lignin, glutathione and sulphate metabolism in response to cadmium in Arabidopsis thaliana and the related Zn/Cd- hyperaccumulator Thlaspi caerulescens. *Plant, Cell & Environment* **2008,** *31* (3), 301-324.

91.   Mittler, R.; Kim, Y.; Song, L.; Coutu, J.; Coutu, A.; Ciftci-Yilmaz, S.; Lee, H.; Stevenson, B.; Zhu, J.-K., Gain- and loss- of- function mutations in Zat10 enhance the tolerance of plants to abiotic stress. *FEBS letters* **2006,** *580* (28-29), 6537-6542.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this study two approaches have been utilised to understand the role of IPCSs in plants: a chemical approach involving the screen for inhibitors of *At*IPCS2 and *Os*IPCS2 combined with a genetic approach involving the analyses of transcriptome changes in response to the over expression and the knock down in expression of *At*IPCS.

Inhibitors of IPCS2 were identified which showed inhibitory activity at low µM, and exhibited herbicidal phenotypes associated with stress (purple patches) and PCD (chlorotic leaves) in an *in vivo* screen. The phenylamidine and triazinone classes of compounds have both been shown to be inhibitors of *At*IPCS2 and *Os*IPCS2, but showed differential selectivity for *At*IPCS2. Both were found to have little or no effect on AUR1, the yeast orthologue.

The identification of these two classes of compounds was interesting, in that the triazinone are known antifungals and the phenylamidines have herbicidal activity. These data also indicated that *At*IPCS2 could be both a herbicidal and antifungal target. The use of the phenylamidines as herbicides is quite promising in that they show obvious phenotypes of stress resulting from the application of the compounds on wildtype plants. The two phenylamidines tested show phenotypes associated with oxidative stress and PCD which result in the inhibition of plant growth. In the case of seedlings treated with compound 1, the PCD phenotype leads to the eventual death of the plant. Further studies will need to be conducted determine if the *in vivo* results can be replicated in mature plants, and whether constant contact between the compound and the plant is necessary to observe the phenotypes reported in this study.

The ability of *At*IPCS2 inhibitors to also have fungicidal effects can be explained by its role as a negative regulator of plant defence against pathogens as discovered from the

RNASeq transcriptomic analyses. *Arabidopsis* transgenic lines over-expressing *AtIPCS1*, *2* and *3*, showed a marked decrease in the expression of genes involved in plant defence; this was heightened in response to *At*IPCS2 overexpression. The knock-down of *At*IPCS1 in a t-DNA insertion line mutant line also resulted in a decrease in genes associated with plant defence, however this did not include the PR genes. *PR1*, *PR2* and *PR5* were highly downregulated in *AtIPCS1*, *2* and 3 over-expressing plants and are well characterised as markers of SAR in plants. These results indicated the prominent role of *AtIPCS1*, *2* and *3* as negative regulators of plant defence. Perhaps these plant defence genes would be upregulated in a homozygous knockout of *AtIPCS2*, resulting in a less constrained response to pathogen attack and heightened resistance.

Another finding of this work was the observation of an early flowering phenotype in *AtIPCS* RNAi knock-down and over-expressers lines; this resembles the phenotype previously linked to vernalization in plants[1, 2]. Combined with RNASeq data which identified the downregulation of several genes involved in abiotic stress in plants over expressing *At*IPCS, there was a strong indication that the enzyme is also involved in abiotic stress response. As such manipulation of expression could increase plant stress tolerance, this demands further investigation.

In summary, this study has identified inhibitors of plant IPCS that have in *vivo* activity and may be tractable for further development towards herbicidal compounds. These compounds show differential selectivity for the dicot IPCS in *A. thaliana*, over the monocot IPCS in *O. sativa*; this indicated that inhibitors selective for a weed versus crop IPCS could be identified. In parallel, transcriptomic analyses of RNASeq analyses revealed that *AtIPCS1, 2* and *3* are involved in the response to abiotic and biotic stress, with a prominent role as a negative regulator of the plant defence network. *AtIPCS1*, *2*

and *3* also seems to have a role in enhancing plant stress tolerance, which could potentially be exploited in plant breeding strategies to create stress tolerant crops, increase yields and meet the challenge of increasing global agricultural output to sustain the projected 2-3 billion population increase by the year 2050.

6.2 Future work

Although these compounds do inhibit *At*IPCS2 still remains a chance that they also affect other targets. To test that these compounds are on target, a dose response assay should be carried out against *Arabidopsis AtIPCS2* over-expressers, knock-downs and wildtype. Theoretically, compounds targeting *At*IPCS2 should be less effective against the over-expressers, due to excess target, and more effective against knock-downs for the opposite reason. These identified inhibitors could also be tested against IPCS from crop and weed species to see if the observed differential selectivity is maintained.

RNASeq could also be used to analyses of wildtype plants and plants with *At*IPCS2 deleted or knocked-down, treated with the identified inhibitors. If the compounds are completely on target, an absence of *At*IPCS2 would preclude changes due to the inhibitor. The homozygous RNAi lines created in this study could also be used in transcriptomic analyses to see if the changes observed in the *AtIPCS1* t-DNA mutant also occur in response to the knock-down.

Work done by Owen et al. showed that the activity of triazinone possessing an R_2 haloalkyl group is enhanced against the ascomycetes, *Leptosphaeria nodorum*, *Magnaporthe grisea* and *Zymoseptoria tritici*[3]. Further activity enhancement was achieved by the replacement of the haloalkyl group with a benzyl or alkyl group. These triazinone derivatives are quite different from the triazinone identified as inhibitors of

*At*IPCS2, in that they lack the thio linker and possess two benzyl groups that have halogens attached to them. Therefore, it is possible that these could be selective for plant over fungal IPCS. Different derivatives of triazinone may be able to protect plants from fungal infection by direct inhibition of pathogenic fungi IPCS, by inhibition of the plant IPCS, or by both. Inhibition of a negative regulator of PCD result in an enhanced plant defence response able to stop fungal growth. This latter effect, has been indicated by the previously observed, slightly lower susceptibility of *AtIPCS2* knockout plants to powdery mildew[4].

To further test this hypothesis, IPCS from different ascomycetes species would need to be isolated, purified and assayed against the identified triazinone to see if the fungal IPCS is also inhibited if compounds selective for the *At*IPCS2 could be identified, a pathogen assay could also be designed to test the susceptibility of *Arabidopsis* to infection on triazinone treatment. The susceptibility of *AtIPCS2* knock-out or knock-down could be similarly analysed.

If *At*IPCS2 acts as a prophylactic to fungal infection, these experiments should show a reduction in fungal mass on chemical inhibition, genetic knock-out/knock-down of *At*IPCS2. In addition, *AtIPCS2* over-expressor lines could also be assayed to see if the increase in expression has the opposite effect, making plants more susceptible to fungal infection. My results show an overall decrease in the expression of plant defense genes, on overexpression of *AtIPCS1*, *2* and *3*, suggesting that this would be the case.

Similar to the proposal above, regarding the further analyses of plants treated with the identified compounds, RNASeq could be utilized to identify if genes associated with plant defense are upregulated to higher degree in *AtIPCS2* knockout/down or compound treated *Arabidopsis* during infection. Work done by Weeda *et al.* using this method identified the

role of melatonin in plant defense by treating plants with low (100 pM) and high (1 mM) concentrations and then comparing the genes upregulated in response to both treatments[5]. This experimental approach when applied to the study of *At*IPCS2 inhibition, would serve to further probe the hypothesized prophylactic effect of *At*IPCS2 inhibition in fungal infection by upregulation of plant defense genes.

One of the intriguing findings from this study was that an increase in the expression of *AtIPCS1*, *2* and *3* resulted in the upregulation of genes involved in the photosynthetic pathway. To test if this has an effect on the plants' ability to photosynthesize, the photosynthetic efficiency of PSII could be quantified by measuring the florescence of chlorophyll[6]. In addition, plant biomass could also be measured to see if there is an increase, correlated with *AtIPCS1*, *2* and *3* overexpression and upregulation of genes involved in the photosynthetic pathway.

The lipid storage and lipid transfer proteins were also upregulated in response to *At*IPCS1-3 overexpression. To probe the effects of this on plant lipid content, lipidomic experiments could be performed; it would also be of interest to investigate how the sphingolipid and long chain base (LCB) content is altered in response to the overexpression and knock-down of *AtIPCS1*, *2* and *3*. Finally, the over-expressor lines could also be subjected to abiotic stress such as cold, salinity and drought to find out if they have higher stress tolerance compared to Col-0 and if there is a link between higher levels of *AtIPCS1*, *2* and *3* expression, and an increase in stress tolerance. The experiments outlined could confirm the tractability of IPCS inhibition in crop protection, which would be advantageous to increase crop yields.

6.3 References

1.   Henderson, I. R.; Dean, C., Control of Arabidopsis flowering: the chill before the bloom. *Development* **2004,** *131* (16), 3829-3838.
2.   Sheldon, C. C.; Rouse, D. T.; Finnegan, E. J.; Peacock, W. J.; Dennis, E. S., The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences* **2000,** *97* (7), 3753-3758.
3.   Owen, W. J.; Sullenberger, M. T.; Loso, M. R.; Meyer, K. G.; Slanec, T. J., Synthesis and antifungal activity of 3- aryl- 1, 2, 4- triazin- 6- one derivatives. *Pest management science* **2015,** *71* (1), 83-90.
4.   Wang, W.; Yang, X.; Tangchaiburana, S.; Ndeh, R.; Markham, J. E.; Tsegaye, Y.; Dunn, T. M.; Wang, G.-L.; Bellizzi, M.; Parsons, J. F., An inositolphosphorylceramide synthase is involved in regulation of plant programmed cell death associated with defense in Arabidopsis. *The Plant Cell* **2008,** *20* (11), 3163-3179.
5.   Weeda, S.; Zhang, N.; Zhao, X.; Ndip, G.; Guo, Y.; Buck, G. A.; Fu, C.; Ren, S., Arabidopsis transcriptome analysis reveals key roles of melatonin in plant defense systems. *PloS one* **2014,** *9* (3), e93462.
6.   Baker, N. R., Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.* **2008,** *59*, 89-113.

CHAPTER 7

MATERIALS AND METHODS

7.1 Experimental section

7.1.1 Materials

Forward and reverse primers for the cloning of *At*IPCS1-3 and *Os*IPCS1-3 into the pESC-LEU plasmid and the cloning of *At*IPCS1-3 in to the pENTR/D-TOPO vector were purchased from Sigma; Quick-Load 1kb DNA ladder (NO4685, Dorset, UK) and Gel Loading Dye Blue (6X) (B70215, Massachusetts, USA) were purchased from BioLabs; alamarBlueR (DAL1025, Loughborough, UK) was purchased from Invitrogen; PCR high fidelity Master Mix kit (F531S, Loughborough, UK) was purchased from Thermoscientific; QIAprep DNA Miniprep kit was purchased from Qiagen (27104, Manchester, UK); High capacity cDNA reverse transcription kit were purchased from Thermoscientific (00539510, Loughborough, UK); Go-taq qPCR Master mix was purchased from promega (A600A, Southampton, UK); Reliaprep RNA tissue miniprep system was purchased from promega (Z6112, Southampton, UK); . In-Fusion® HD Cloning Kit with Cloning Enhancer was purchased from Clontech (638909, Saint-Germain-en-Laye, France); complete® EDTA-free Protease Inhibitor Cocktail Tablets were purchased from Roche Applied Science (05892791001, Sussex, UK). D (+) raffinose (RAF02, Norfolk, UK); Complete Supplement Mix (CSM) drop-out: -His –Leu –Trp –Ura (DCS1389, Norfolk, UK); yeast nitrogen base low fluorescence without amino acids, riboflavin (CYN6510, Norfolk, UK); D (+) galactose (GAL03, Norfolk, UK) and 5-FOA (5FOA01, Norfolk, UK) were purchased from Formedium; leucine (L8912), uracil (U0750), tryptophan (T0254), D (+) glucose (G8270), ammonium sulphate (A5132), tris base (tris hydroxymethyl amino methane) (252859), myriocin from *Mycelia sterilia* (M1177), L-α phosphatidylinositol ammonium salt from soybean (P0639), L-α phosphatidylinositol ammonium salt from bovine liver (P5766), cycloheximide (C7698), dimethyl sulfoxide (D8418), potassium formate (294454), potassium phosphate

monobasic (P9791), potassium phosphate dibasic (P3786), agar for plant cell culture (P9791), rifampicin (R3501) and ampicillin (A1593) were all purchased from Sigma Aldrich (Norfolk, UK); N-Hexanoyl NBD-phytoceramide (1628, Cambridge, UK) was purchased from Universal Biologicals; Aureobasidin A was purchased from AureoGen (Kalamazoo, USA); NBD-C6-ceramide (N1154, Dorset, UK) was purchased from Invitrogen; AG4-X4 resin (140-4341, Hertfordshire, UK) was purchased from BioRad; Molecular Biology Agarose Electrophoresis Grade (MB1200), kanamycin (K0126), spectinomycin (S0188) and timentin (T0190) were purchased from Melford (Suffolk, UK); Murashige and Skoog medium (M0222, Haarlem, Netherlands) was purchased from Duchefa Biochemie; bacto-Agar (214010, Surrey, UK) was purchased from Becton Dickinson and Company. Peat plugs were purchased from Jiffy Products (3825451, Kristiansand, Norway); ethidium bromide solution (46067, Buchs, Switzerland) was purchased from Fluka BioChemika; LoBind tubes for microsome work were purchased from Eppendorf (10708704, Loughborough, UK). The acid-washed glass beads 425-600 µm (G8772), glycerol (G5516), sucrose (S0389), polyethylene glycol 3350 (1546547), $MgCl_2$ (M2670), lithium acetate (517992), were all purchased from Sigma Aldrich (Dorset, UK). Yeast strains, lacking drug efflux pumps (*pdr1Δ::KanMX4 pdr3Δ::KanMX4 pdr16Δ::KanMX4 pdr17Δ::KanMX4)* were provided by Prof. Stark, University of Dundee: MSY-23-3C (Parent Strain), MSY-23-3C-1A (pESC-Leu-*Lmj*IPCS), MSY-23-3C-3A (pESC-Leu-AUR1), and MSY-23-3C-6A (pESC-Leu-*Os*IPCS1).

7.1.2 Instruments and Equipment

Centrifugation steps were carried out using Beckman Coulter centrifuges or ultracentrifuges. Eppendorf tubes were spun down using a Sigma 1-14 microfuge. Eppendorf contents were dried using an Eppendorf Vacuum Concentrator 5301 from Brinkmann. Disruption of cells was performed using an IKA® Vortex Genius 3. For protein and cell quantification carried out using BOECO S-30 spectrophotometer. Quantitative PCR was carried out using the applied Biosystems 7300 machine. For the enzyme assay, the filtration step was carried out using multiscreen® solvient filter plates purchased from Millipore and OptiPlate-96 F black plates used for reading the fluorescence of NBD-IPCS from Perkin Elmer. V-shaped and flat bottom 96 well plates were purchased from Sigma Aldrich. HPTLC plates were from Merck and were analysed using a Fuji FLA−3000 plate reader and AIDA Image Analyser® software (version 3.52). The localisation experiments of 35s::GFP.*At*IPCS1-3 were performed with the Leica Stereo Fluor M165FC and Leica SP5 CLSM FLIM FCCS microscopes. Details of the buffers and solutions used are given in Table 7-1 and details of media composition is given in Table 7-2.

7.2 Table 7-1: Buffer composition

Buffers	Final concentration	Volume/Weight	Stock solution
Sodium chloride-Tris-EDTA (STE) buffer	25 mM	1.25 ml	Tris/HCl (1 M)
	250 mM	12.5 ml	Sucrose (1 M)
	1 mM	100 ul	EDTA (0.5 M)
		1 tab	Complete® Protease Inhibitor Cocktail
		50 ml	Water
Storage buffer	50 mM	2.5 ml	Tris/HCl (1 M)
	10 w/v	6.25 ml	Glycerol (80 % w/v)
	5 mM	0.25 ml	$MgCl_2$ (1M)
		1 tab	Complete® Protease Inhibitor Cocktail
		50 ml	Water
Tris/EDTA/BSA buffer	250 mM	12.5 ml	Tris/HCl (1 M)
	25 mM	2.5 ml	EDTA (0.5 M)
	15 mg/mL	750 mg	Free fatty-acid BSA
		50 ml	Water
TE buffer		100 ml	1 M Tris/HCl (pH7.5)
		20 ml	500 Mm EDTA (pH8)
		880 ml	Ultrapure water
PEG 3350 (50%)		100 mg	PEG 3350
		200 ml	Ultrapure water
TE/Li Acetate		5 ml	1 M Li Acetate
		5 ml	10x TE buffer

	50 ml	Ultrapure water
PEG/Li Acetate	5 ml	1 M Li Acetate
	5 ml	10x TE buffer
	50 ml	50% PEG 3350

7.3 Table 7-2: Media composition

Growth media	Volume/Weight	Stock
Lysogeny broth	25 g	Miller's LB broth
	20 g	Bacto agar
	1 ml	Ampicillin (100 mg/ml)
	1 L	Distilled water
Murashige and Skoog (0.8% w/v) agar (pH 5.6)	4.4 g	MS medium with vitamins
	8 g	Plant tissue agar
	1L	Distilled water
SD-H-L agar	20 g	Glucose
	1.93 g	Yeast Nitrogen base without amino acids and without ammonium sulphate
	5 g	Ammonium Sulphate
	15 g	Bacto agar
	580 mg	Amino acid drop out supplement -His-Leu-Ura-Ttp
	20 mg	Uracil
	40 mg	Tryptophan
	1 L	Distilled water

SGR-H-L + FOA agar	**1 g**	**Galactose**
	10 g	Raffinose penhydrate
	1.93 g	Yeast Nitrogen base without amino acids and without ammonium sulphate
	5 g	Ammonium Sulphate
	15 g	Bacto agar
	1 g	5-Fluoroortoic acid monohydrate
	580 mg	Amino acid drop out supplement - His-Leu-Ura-Ttp
	20 mg	Uracil
	40 mg	Tryptophan
	1 L	Distilled water
SGR-H-L media	1 g	Galactose
	10 g	Raffinose penhydrate
	1.93 g	Yeast Nitrogen base without amino acids and without ammonium sulphate
	5 g	Ammonium Sulphate
	580 mg	Amino acid drop out supplement -His-Leu-Ura-Tryp
	20 mg	Uracil
	40 mg	Tryptophan
	1 L	Distilled water
Phosphate buffer (0.25 mM, pH 7)	13.1g	KH_2PO_4
	26.8g	K_2HPO_4
	1L	Ultrapure water

7.4 Biological Procedures

All of the following biological procedures were carried out under sterile conditions unless stated otherwise. Plant were gown under 150–200 µmol m^{-1} s^{-1} lightening at 20°C and under a 16h photoperiod.

7.4.1 DNA extraction from bacteria

The QIAprep Spin Miniprep Kit was used to extract DNA from bacterial cultures by following manufacturer's instructions.

7.4.2 RNA extraction from plant

The ReliaPrep^{TM} RNA tissue miniprep system kit was used to extract RNA from 1mg of plant material by following manufacturer's instructions.

7.4.3 cDNA synthesis

The RNA concentrations were quantified using nanodrop and all samples diluted to 2µg in 10µl total volume. To the RNA mixture was added 10µl of Master mix of:

Component	Volume per reaction (µl)
10x RT buffer	2
25x dNTP Mix (100 mM)	0.8
10x RT Random Primers	2
Nuclease-free water	4.2
Multiscribe^{TM} Reverse transcriptase	1

The mixture was mixed and spun down briefly, then placed in thermal cycler using the program below:

	Step 1	Step 2	Step 3	Step 4
Temperature (°C)	25	37	85	4
Time (secs)	160	7200	5	∞

7.4.4 Quantitative reverse-transcription real-time PCR

cDNA samples were diluted (1:50) and 5ul of the diluted sample was added to 10µl of the real-time PCR master mix, which consisted of:

Component	Volume per reaction (µl)
2x SYBR mix (ROX)	7.5
Forward primer (5µM)	0.9
Reverse primer (5µM)	0.9
Nuclease-free water	0.7

Real-time PCR reaction was carried out using the following program:

Temperature (°C)	Time (seconds)	Number of cycles
50	120	1
95	600	1
95	15	40
60	60	1

7.4.5 Polymerase chain reaction protocol

The PCR mixture was prepared using 2x phusion HF master mix (10 µL) and (1 µL) of the forward and reverse primers with plasmid DNA (10 ng); final volume was adjusted to 20 µL using distilled water. The PCR reaction was run on a PCR machine under the conditions described below. The DNA was stored at -20ºC.

Cycle step	Temperature (ºC)	Time (seconds)	Number of cycles
Initial denaturation	98	30	1
Denaturation	98	10	30
Annealing	56	30	
Extension	72	100	
Final extension	72	600	1
	4	hold	

7.4.6 T4 ligase cloning

Agarose gel electrophoresis was used to quantify DNA concentration and for the DNA moles calculated. The molar ratio of insert to vector in the reaction was 1:3. T4 ligase cloning reaction was set up as follows:

Component	Volume per reaction (µl)
T4 ligase enzyme	0.5
T4 ligase buffer	1.5
Vector	X µl (~ 30ng)
Insert	Y µl to give 1:3 molar ratio
Nuclease-free water	20-X-Y

The reaction mixture was incubated at 25°C for one hour; 5 µl of the reaction mixture was subsequently used for bacterial transformation. *At*IPCS1-3 were cloned into pENTR/D-TOPO vector using T4-ligase cloning method.

7.4.7 Gateway cloning

Creating multiple expression vector constructs for the target gene of interest (*At*IPCS1-3) using traditional ligase-mediated cloning is time-consuming and laborious, hence the use of gateway cloning. Gateway cloning exploits the bacteriophage lambda recombination system, bypassing the use of ligase-mediated cloning[1]. PCR amplification of the target sequence followed by T4 ligase cloning into the pENTR/D-TOPO 'entry vector' creates a plasmid that has the target DNA sequences flanked by *att*L recombination sequences. The attL sequence recombines with the *att*R sites using the LR clonase reaction mix which transfers the target sequence into a desired 'destination vector' (Figure 7-1). This destination vector usually carries the chloramphenicol resistance gene (*CmR*) and a *ccdB* gene which is lethal to most strains of *Escherichia coli*. the (*ccdB*) as a result it is possible to negatively select for the vector possessing the recombinant sites upon transformation of *E. coli* cells. This selection is combined with positive selection for an antibiotic resistance marker thereby making certain that only colonies harboring the 'expression vector' are selected.

1) PCR amplification of insert

2) Clone into entry vector pENTR/D-TOPO

attL1   attL2

3) Recombination

attL1   attL2   Entry vector

promoter   attR1   CmR   ccdB   attR2   Destination vector

4) Expression vector

promoter attB1   attB2
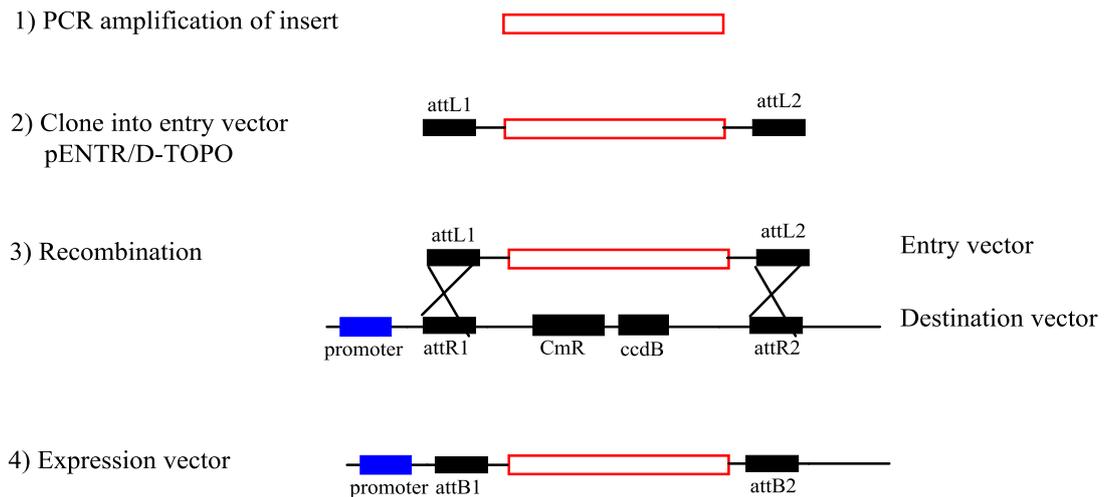
Figure 7-1: Overview of Gateway cloning (1) the target of interest is amplified by PCR (2) PCR products are mixed with the linearized pENTR/D-TOPO vector and ligation mediated by T4 ligase (3) Recombination between *att*L and *att*R sites is mediated by LR clonase with the target sequence recombined into a destination vector of choice (4) Resulting expression vector containing the target gene.

Cloning into the destination vector pK7WG2, carried out via Gateway LR clonase mediated cloning. The following components were added to a PCR tube:

Component	Volume per reaction (µl)
Entry clone (50-150 ng/reaction)	1-7
Destination vector (150 ng/µl)	1
TE buffer, pH 8	to a volume of 8 µl
LR Clonase^{TM} ll enzyme mix	2

Reaction was incubated at 25°C for 1 hour and 4ul of mixture used for transformation.

7.4.8 In-Fusion cloning

Agarose gel electrophoresis was used to quantify the insert and plasmid DNA. Using the commercial Infusion cloning kit, the linearized vector was mixed with the insert DNA (1:3 ratios) and 5× In-Fusion HD enzyme premix (2 µl) added with the final volume

adjusted to 10 μl with distilled water. The mixture was then incubated at 50 °C for 15 mins and then used immediately for the transformation of *E. coli*.

7.5 Yeast culture

The *S. cerevisiae* strains lacking the drug efflux genes *pdr1Δ::KanMX4, pdr3Δ::KanMX4, pdr16Δ::KanMX4 and pdr17Δ::KanMX4* which had been stored at -80oC in 80% were inoculated onto SGR-H-L agar and incubated at 30 °C for 3 days, then stored at 4oC. This was then used to inoculate SGR-H-L media (5 ml) which was incubated at 30 °C with 180 rpm shaking.

7.5.1 Frozen stocks

400 μl of 80% glycerol was mixed with 800 μl of yeast culture grown in SGR-H-L media and stored at -80°C.

7.5.2 Yeast transformation

Cells grown overnight in SGR-H-U media at 30°C were harvested by centrifugation (1,000 ×g for 5 minutes at room temperature). The supernatant was discarded, and the cells were resuspended in sterile 1x TE buffer (5 ml). The cells were spun down again (1,000 × g for 5 minutes at room temperature) and the supernatant discarded. The cell pellets were then resuspended in freshly prepared, sterile $1 \times$ TE/LiAc buffer (1 ml) to form the competent yeast cells.

Purified plasmid (1 μg) and competent yeast cells ($OD_{600} = 0.6$-0.8, 0.1 ml) was added to Herring testes carrier DNA (10 μl) which had previously been boiled at 95 °C for 10 minutes; and mixed thoroughly. Freshly prepared, sterile PEG/LiAc (0.6 ml) was added and the mixture thoroughly mixed, followed by incubation with shaking at 30 °C for 30

minutes. After which DMSO (70 µl) was added and mixed by gentle swirling. The competent yeast cells then underwent heat shocked at 42 °C for 15 minutes followed by chilling on ice for 2 minutes. The cells were spun down (14 000 rpm for 5 seconds at room temperature), the supernatant discarded. The cell pellets were resuspended in $1\times$ TE buffer (0.2 ml) and the cells were plated onto SGR-H-L-U agar plates and incubated at 30 °C for 5 days.

7.5.3 Plasmid shuffle

SGR –H –L medium (0.1%:1%) was inoculated with yeast strains and incubated with shaking at 30 °C overnight and 200 µl plated onto SGR –H –L +5FOA agar medium (0.1% galactose: 1% Raffinose) which was incubated for 4 days at 30 °C.

7.5.4 Yeast culture scale-up

Plasmid-shuffled yeast cultures were propagated in SGR –H –L medium (0.1% galactose: 1% Raffinose). Liquid medium (5 ml) was inoculated and incubated overnight at 30 °C until OD_{600} had reached ≥ 0.8. Fresh medium (250 ml) was added and the culture incubated at 30 °C until OD_{600} had reached ≥ 0.8. The culture was then diluted into fresh medium (1L) and incubated at 30 °C overnight until OD_{600} =0.8-1.0.

7.5.5 Preparation of crude microsomal membranes

The yeast cells were harvested by centrifugation (4,000 $\times$ g for 10 minutes at 4 °C) and resuspended in cold PBS (20 ml per litre). The cell suspension was then transferred into 50 mL falcon tubes (no more than 20 ml to each falcon tube) and the cells were pelleted by centrifugation (4000 rpm for 10 mins at 4ºC). The centrifuge bottles were washed with 5 mL per litre of PBS and the washing transferred to falcon tubes and spun down (4 000

rpm for 10 mins at 4°C) supernatant from both centrifugation steps were discarded. The cell pellets were weighed and pre-chilled glass acid-washed glass beads (425-600 μm) to disrupt the cells were added (mass equal to 1.5 x mass of the pellet) and STE buffer was also added (volume in ml equal to the mass of the pellet in g). The cells were disrupted using pre-chilled, acid-washed glass beads (mass equal to 1.5x wet cell mass) for 30 cycles of 1 minute on the vortex machine, followed by 1 minute rest on ice. This mixture was then spun down (5 000 rpm for 15 minutes at 4 °C) and the supernatant was stored. STE buffer was added to cell pellets (volume in ml equal to 0.5x mass of wet cell mass) and the cell disruption cycle repeated for a further 20 cycles, followed by centrifugation (5500 rpm for 15 minutes at 4°C). Supernatant was transferred to ultracentrifuge tubes which were balanced to within 10 mg of each other using STE buffer and spun down (12500 rpm for 30 mins at 4°C); the supernatant was transferred to new ultracentrifuge tubes and spun down again (32 000 rpm for 90 mins at 4°C). Pellets were resuspended in the minimum amount of storage buffer and stored at -80°C. The protein content for each pellet and the supernatant was determined by Bradford assay[2].

7.5.6 Diffusion assay

Agarose was added to SGR –H –L medium (0.8%, 200 mL) and sterilised. To 15 mL of agarose SGR-H-L medium, $OD_{600} = 2$ of liquid yeast culture were added, which was then mixed by gentle inversion. This was poured into a square plate and left to set. Aureobasidin A, myriocin and DMSO were each spotted onto the agarose surface in aliquots of 1 μl, 2 μl and 3 μl; the plate was then incubated at 30 °C for 3 days.

7.5.7 High performance thin layer chromatography (HPTLC) assay

Phosphatidyl inositol solution (50 mM) was dried in a rotavapor and a predetermined Unit/µl of active IPCS added. Tris/EDTA/BSA buffer (20 µl) was added and the volume adjusted to 48 µl with distilled water. This mixture was then incubated at 30 °C for 30 minutes, after which NBD-C6-ceramide (200 µM) was added and then incubated at 30 °C for 60 minutes. The reaction was then quenched with 150 µL of chloroform:methanol:water mixture (10:10:3). The mixture was spun down (14 000 rpm for 5 minutes at room temperature) and the aqueous layer was extracted. Upon the addition of chloroform (50 µl), the aqueous layer was re-extracted and the organic extracts were combined and solvent dried in the rotavapor at 30 °C for 15 minutes. The pellets were resuspended in 20 µl of chloroform:methanol:water mixture (10:10:3). On the HPTLC plate 5µL of the resuspended pellet was loaded and this was run using the solvent system chloroform: methanol: 0.25% (w/v) aqueous KCl mixture (55:45:10). NBD-IPC turnover was quantified and analysed on a TLC reader.

7.6 Experiments for high throughput screen

7.6.1 Frozen stocks for screening

Yeast strains were grown to an OD_{600} of 0.5-0.7 and then spun down at 1000 rpm. The pellets were resuspended in a one to one mixture of SGR-H-L media and 20% glycerol, then frozen and stored at -80°C.

7.6.2 Primary screen, cellular assay

Cells were thawed on ice and diluted with SGR-H-L media (1:20) to obtain cells in the lag phase of growth with a starting OD_{600} of 0.06 corresponding to 0.5×10^7 cell per ml.

The yeast cell culture was aliquoted (198 µl) into 96-well plates using the Biomek FX^P automated workstation with DMSO (2ul) and cycloheximide (2ul, 0.01 mM) dispensed into eight wells in the first column of the 96 well plates; the dispension of DMSO and cycloheximide was alternated between a row of wells. Following this, these compound (10 µM) were aliquoted into the remaining wells. Plates were sealed with a breathable membrane and incubated at 30°C for 24 hours After 24 hours the OD_{600} reading was obtained. Readings were carried out in duplicate for each strain and percentage inhibition calculated with this formula:

$$I\,(\%) = \frac{\Delta OD600nm(NC - Compound)\,x\,100}{\Delta OD600nm\,(NC - PC)}$$

Where NC is the average of the negative control (DMSO) and PC is the average of the positive control (cycloheximide)

For dose response, hit compounds were diluted (1:3) to obtain a range of concentrations from 68nM-50µM. Frozen yeast strains were thawed on ice and diluted with SGR-H-L media (1:20) and aliquoted into 96 well plates, then sealed with a breathable membrane and incubated for 24 hours at 30°C. Cell density was measured at OD_{600}. IC_{50} values were determined using Graphpad Prism.

7.6.3 Secondary screen, enzyme assay

Filter plate wells are coated with AG4-X4 resin (10% w/v in ethanol) and prepared according to the protocol described by Mina et al[3]. For a 96 well plate assay set up, a mixture consisting of dried soy phosphoinositol solution (200 µM), NBD phytoceramide (30 µM) and phosphate buffer (pH7, 71 mM) was prepared. A second mixture made up of *At*IPCS2 microsomes (0.44 Unit/µl), *Os*IPCS2 microsomes (0.07 Unit/µl) and AUR1 microsomes (0.18 Unit/µl) respectively, were prepared with the addition of storage buffer

and phosphate buffer (pH7, 71 mM). The mixture consisting of PI and NBD-phytoceramide was combined with the microsome mix to get a final volume of 40 µl aliquoted into a v-shaped 96-well plate. The hit compounds were assayed at 10 µM in duplicate. Mixture was incubated for 60 mins at 30°C and quenched with 200 µl of HPLC grade methanol. IPC was eluted from the resin using 4x 50 µl potassium formate solution (1 M) in methanol and fluorescence quantified at λ_{Ex} 480 and λ_{Em} 540. Compounds showing an average of $\geq$ 30% inhibition were taken on to the next stage for IC_{50} determination. For IC_{50} determination compounds were diluted (1:3) to obtain compound concentrations ranging from 68 nM-100 µM, and screening carried out in triplicate for each hit. IC_{50} was determined using GraphPad Prism.

7.7 Method for plant-based experiments

7.7.1 Agrobacterium transformation

Competent cells of agrobacterium strains GV3101 (for localisation experiments in tobacco) and C58C1 (for agrobacterium mediated plant transformation) were transformed with 1 µl of expression vectors (pK7FWG2, pK7WG2 and pK7GWIWG2 (I) respectively) harbouring *At*IPCS1-3. Cell were heat shocked for 5 mins at 37°C and 1 ml of lysogeny broth (LB) added and then incubated at 28°C for 3 hours at 140 rpm. Transformants were plated on LB (100 µg/µl Rifampicin and 100 µg/µl Spectinomycin) and incubated for 3 days at 28°C.

7.7.2 Agrobacterium mediated plant transformation

Following transformation of agrobacterium strain C58C1 with binary constructs, agrobacterium cells were cultured overnight in Lb (10 ml, 100 µg/µl Rifampicin and 100 µg/µl Spectinomycin) and then scaled up for growth in 200ml LB to use for plant

transformation. Wildtype *A .thaliana* plants were transformed using the floral dipping method[4].

7.7.3 Plant growth conditions

Arabidopsis thaliana plants were grown on 0.8% Murashige and Skoog (MS) agar and transferred on to peat plugs after 10 days. Plants were grown at 20°C under 16h day/8 h night photoperiod.

7.7.4 Plant transformants selection

Transformed plants were grown to maturity, and the seeds collected. Selection of positive transformants was carried out on MS agar (50 µg/µl Timentin and 50 µg/µl Kanamycin) which were incubated for 6 days.

7.7.5 Tobacco infiltration

Agrobacterium GV1310 strain harbouring GFP::AtIPCS1-3 and ST RFP Golgi marker respectively were grown in Lb media (100 µg/µl Rifampicin and 100 µg/µl Spectinomycin). GV1310 agrobacterium strain harbouring a vector encoding a p35S cytosolic GFP marker was also grown in Lb media (100 µg/µl Rifampicin and 100 µg/µl Kanamycin) overnight at 28°C. cell density at OD_{600} was quantified and diluted to an $OD_{600} = 0.1$ for GFP::*At*IPCS1-3 and ST RFP, $OD_{600} = 0.3$ for p35S cytosolic GFP marker with a solution of $MgCl_2$ (10 mM). *Nicotiana benthamiana* was co-infiltrated with agrobacterium harbouring vectors encoding ST RFP and GFP::*At*IPCS1-3 respectively. *N. benthamiana was also infiltrated with agrobacterium encoding the* p35S GFP cytosolic marker separately. After 48 hours, plants were viewed using the Leica Stereo Fluor M165FC and Leica SP5 CLSM FLIM FCCS microscopes.

7.7.6 *In-vivo* screening

A. thaliana seedling were grown for 10 days on 0.8% MS agar and then transferred to 1.2% MS agar containing compounds at a final concentration of 10 µM and 40 µM.

7.8 Bioinformatics analysis

Opening reading frames that encode for IPCS homologues in different plant species were identified by running BLAST search of *At*IPCS1 nucleotide coding sequence against each species genome using the phytozyme database (https://phytozome.jgi.doe.gov/pz/portal.html). Multiple amino acid sequence alignments of *At*IPCS1-3, *Os*IPC1-3 and *Ench*IPCS was carried out using ClaustalW[5] and shading done using Boxshade. Phylogenetic tree was constructed using MEGA 7.0[6] based on the unweighted pair group method with arithmetic mean (UPGMA) and the evolutionary distances calculated using the maximum composite likelihood method[7]. Transmembrane topology was predicted using the transmembrane hidden Markov model (TMHMH) server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/)[8]. Conserved motifs in the promoter region, 1-1.5 kb upstream of the initiation codon were identified using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)[9] and PlantPAN 2.0 (http://plantpan2.itps.ncku.edu.tw/)[10] .

7.9 Materials preparation for RNASeq experiment

For RNASeq experiment, Col-0 and over-expressors of AtIPCS1, 2 and 3 were grown for 10 days and flash frozen in nitrogen. RNA extraction was carried out using the Plant RNeasy kit from Qiagen according to manufacturer's protocol, and the integrity of the RNA determined by running the samples on an Agilent 2100 Bioanalyzer to obtain a RIN

score. The RIN score was determined by measuring the ratio of 28S to 18S rRNA in the RNA sample, with a 1:2 ratio of 18S:28S being ideal for a RIN score of 10[11]. Following DNAse treatment, quality control of the samples were carried out using the Agilent 2100 Bioanalyzer before proceeding to sample preparation for RNASeq experiment.

7.9.1 RNASeq Experiment

Single end libraries for RNASeq were generated from DNAse treated total RNA using Illumina Truseq stranded mRNA sample preparation kit and instruction manual. Briefly, mRNAs were fragmented and purified for use as template for the synthesis of double stranded cDNA. End repair of the double stranded cDNA was carried out and the 3' end adenylated. Sample specific indexing adapters were ligated to the ends of double stranded cDNA samples and amplified by PCR, then purified. Samples were normalized, pooled and then sequenced using Nextseq500 instrument to obtain 150 base pair single reads.

7.9.2 RNASeq analysis

The RNA sequence in Fastq format[12] were filtered and trimmed (sliding window 4:15 and 50 bp minimum) to remove low quality reads using Trimomatic[13]. Reads were aligned to the *Arabidopsis* genome (*Arabidopsis* Araport 2017) using STAR[14]. The sequence alignment files were sorted by name[15] for HTSeq-count and indexed using SAMtools[15]. Files were then converted to BAM files and number of read mapped onto a genes calculated using HTSeq package[16].Gene counts were normalized and compared sample by sample using DESeq2[17] from Bioconductor[18] in R[19]. Principle component analyses of the normalized counts was done in R using the DESeq package. Differential expression between over-expressors and Col-0 was determined using the DESeq package with a Log_2 fold-change output. GO term enrichment was performed for analysis of genes upregulated and downregulated in both biological replicates of each *At*IPCS isoform over-expressor

using the agriGO analysis tool (http://bioinfo.cau.edu.cn/agriGO/analysis.php) with the

default settings[20].

8.0 References

1.   Earley, K. W.; Haag, J. R.; Pontes, O.; Opper, K.; Juehne, T.; Song, K.; Pikaard, C. S., Gateway- compatible vectors for plant functional genomics and proteomics. *The Plant Journal* **2006,** *45* (4), 616-629.

2.   Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **1976,** *72* (1-2), 248-254.

3.   Mina, J., Thesis: Exploring IPC synthase: a potential anti-leischmanial drug target. **2010**.

4.   Clough, S. J.; Bent, A. F., Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. *The plant journal* **1998,** *16* (6), 735-743.

5.   McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R., Analysis tool web services from the EMBL-EBI. *Nucleic acids research* **2013,** *41* (W1), W597-W600.

6.   Kumar, S.; Stecher, G.; Tamura, K., MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution* **2016,** *33* (7), 1870-1874.

7.   Tamura, K.; Nei, M.; Kumar, S., Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (30), 11030-11035.

8.   Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E. L., Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology* **2001,** *305* (3), 567-580.

9.   Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van de Peer, Y.; Rouzé, P.; Rombauts, S., PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research* **2002,** *30* (1), 325-327.

10.   Chow, C.-N.; Zheng, H.-Q.; Wu, N.-Y.; Chien, C.-H.; Huang, H.-D.; Lee, T.-Y.; Chiang-Hsieh, Y.-F.; Hou, P.-F.; Yang, T.-Y.; Chang, W.-C., PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic acids research* **2016,** *44* (D1), D1154-D1160.

11.   Schroeder, A.; Mueller, O.; Stocker, S.; Salowsky, R.; Leiber, M.; Gassmann, M.; Lightfoot, S.; Menzel, W.; Granzow, M.; Ragg, T., The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology* **2006,** *7* (1), 3.

12.   Cock, P. J.; Fields, C. J.; Goto, N.; Heuer, M. L.; Rice, P. M., The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic acids research* **2009,** *38* (6), 1767-1771.

13.   Bolger, A. M.; Lohse, M.; Usadel, B., Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014,** *30* (15), 2114-2120.

14.   Dobin, A.; Gingeras, T. R., Mapping RNA- seq Reads with STAR. *Current protocols in bioinformatics* **2015**, 11.14. 1-11.14. 19.

15.   Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R., The sequence alignment/map format and SAMtools. *Bioinformatics* **2009,** *25* (16), 2078-2079.

16.   Anders, S.; Pyl, P. T.; Huber, W., HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015,** *31* (2), 166-169.

17.   Love, M. I.; Huber, W.; Anders, S., Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **2014,** *15* (12), 550.

18.   Huber, W.; Carey, V. J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B. S.; Bravo, H. C.; Davis, S.; Gatto, L.; Girke, T., Orchestrating high-throughput genomic analysis with Bioconductor. *Nature methods* **2015,** *12* (2), 115-121.

19.   Ihaka, R.; Gentleman, R., R: a language for data analysis and graphics. *Journal of computational and graphical statistics* **1996,** *5* (3), 299-314.

20.   Du, Z.; Zhou, X.; Ling, Y.; Zhang, Z.; Su, Z., agriGO: a GO analysis toolkit for the agricultural community. *Nucleic acids research* **2010,** *38* (suppl_2), W64-W70.

APPENDICES

APPENDIX A – Transcription factor binding site motif sequence

Transcription factor	Motif sequence
LS5ATPR1	TCTACGTCAC
LS7ATPR1	ACGTCATAGA
MYB1LEPR	GTTAGTT
WRKY 4	TTGACCT
WRKY 18	TTGACCT
WRKY 60	TTGACCT
TGA 1	TGACG

APPENDIX B – pESC-LEU vector map
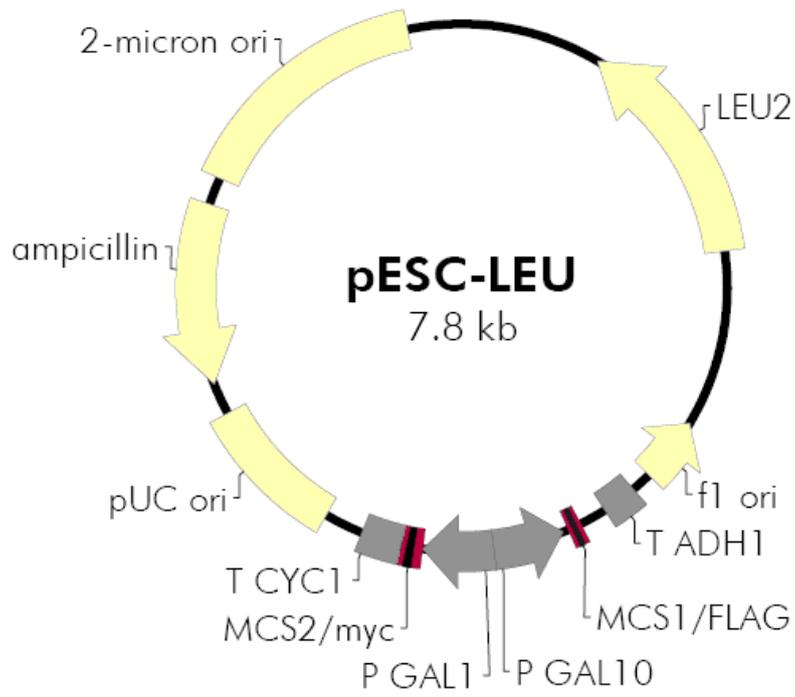
APPENDIX C – vector maps of expression plasmids used for plant genetic work

pK7WG2,0
11159 bp

Xbal (11150)
Smal (11141)
Xmal (11139)
SacII (11015)
Apal (1)
AatII (7)
T35S
SacII (245)
attR2
Xmal (638)
Smal (640)
ccdB
EcoRI (1507)
attR1
Spel (1965)
p35S
SacI (3030)
HindIII (3032)
RB
Kan
LB
Sm/SpR
ClaI (6286)
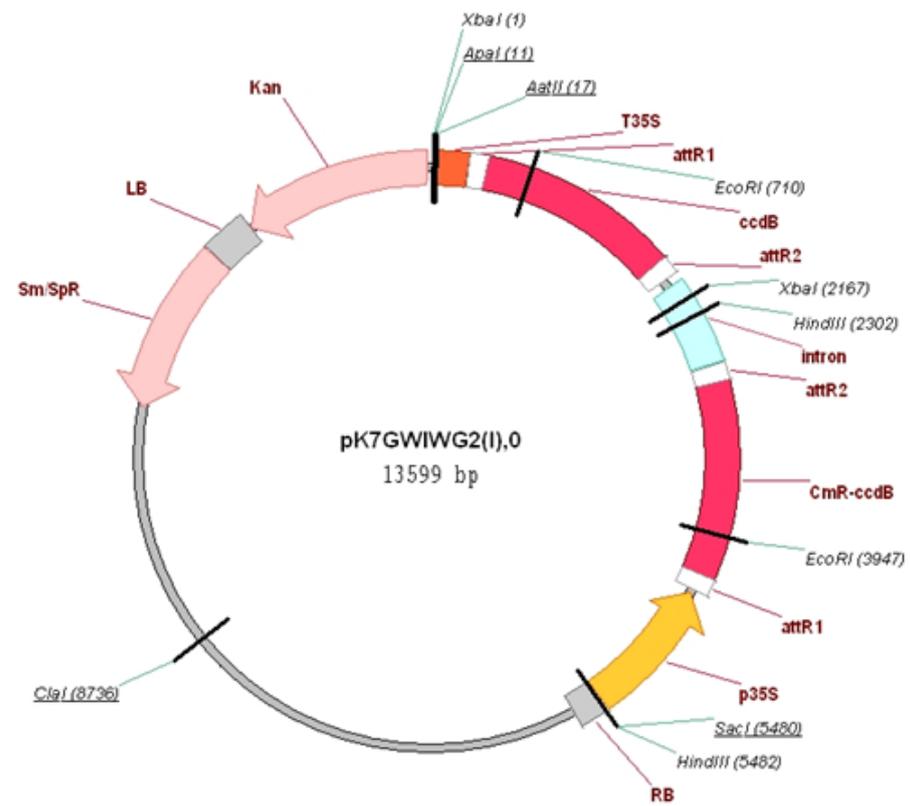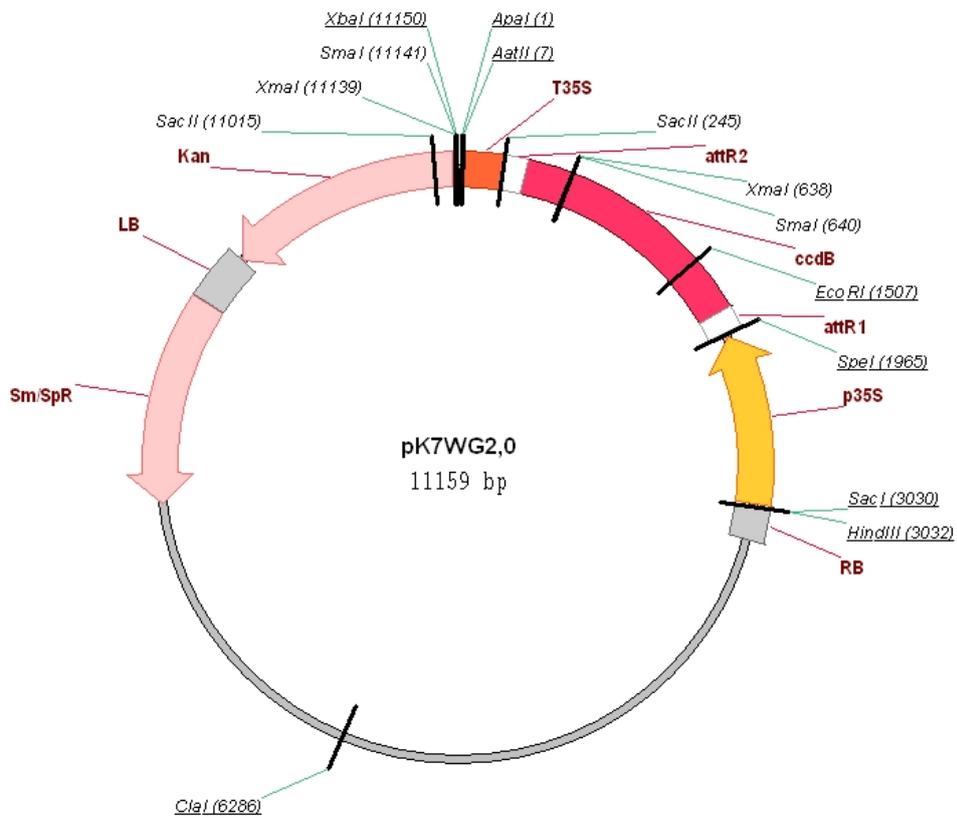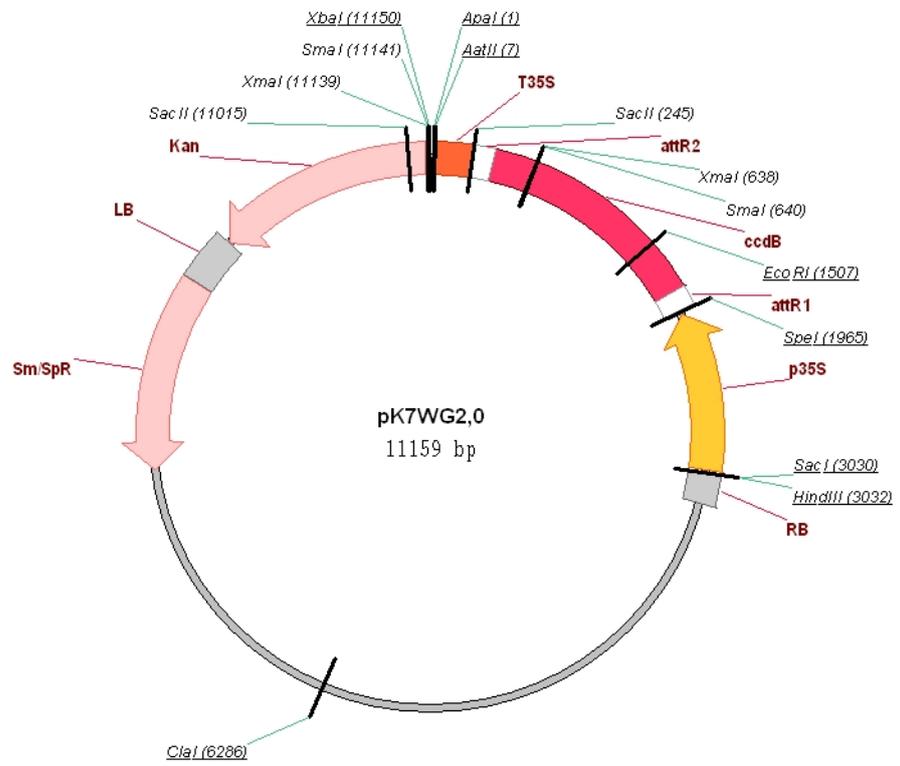
APPENDIX D – Structure of compounds tested against *At*IPCS1 complemented yeast strains

W2345

ND8

Ex2277

Ex2284

1.9

1.25

1.76

2.38

2.42

2.64

1.117
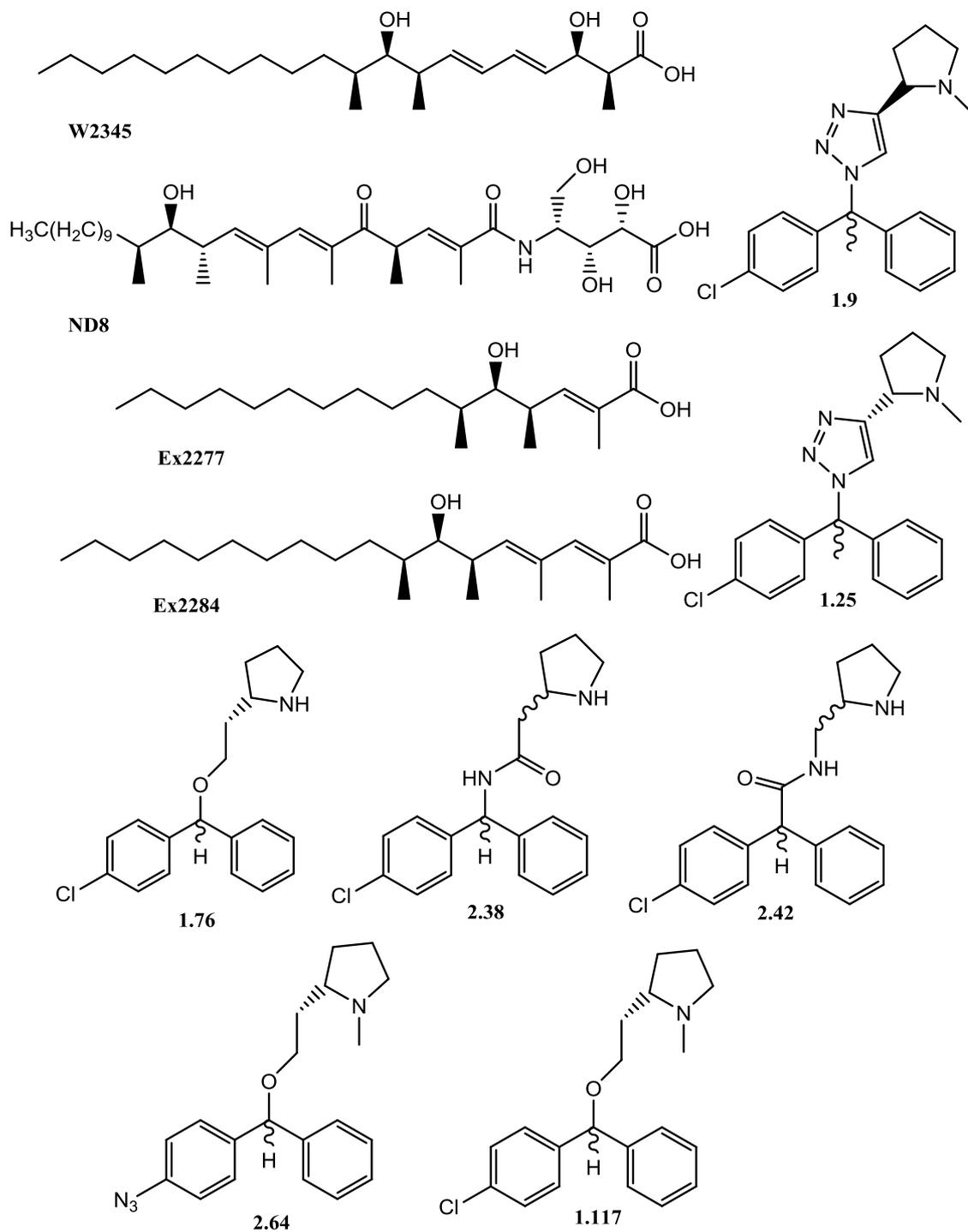
APPENDIX E – Primer sequences

Primers for infusion cloning of *At*IPCS 1,2, and 3 into pESC-LEU vector:

*At*IPCS1 FP: CTCACTATAGGGCCCGGGATGACGCTTTATATTCGCCGC

*At*IPCS1 RP: TCCATGTCGACGCCCGGGTCATGTGCCATTAGTAGCATT

*At*IPCS2 FP: CTCACTATAGGGCCCGGGATGACACTTTATATTCGTCGT

*At*IPCS2 RP: TCCATGTCGACGCCCGGGTCACGCGCCATTCATTGTGTT

AtIPCS3 FP: CTCACTATAGGGCCCATGCCGGTTTACGTTGATCGC

AtIPCS3 FP: TCCATGTCGACGCCCGGGTCAATGATCATCTGCTACATT

*Os*IPCS1 FP: CTCACTATAGGGCCCGGGATGGCGGTTTACATCGCTCG

*Os*IPCS1 RP: TCCATGTCGACGCCCGGGTCATGTGCCATTGGGAGTGGC

*Os*IPCS2 FP: CTCACTATAGGGCCCGGGATGACGATTTACATAGCGCGG

*Os*IPCS2 RP: TCCATGTCGACGCCCGGGTCATGTACCACCATTGACAGC

*Os*IPCS3 FP: CTCACTATAGGGCCCGGGATGTCGGCCTTGTACCTTGCT

*Os*IPCS3 RP: TCCATGTCGACGCCCGGGTCATGTTTTTGTTGTTTCGGT

Internal primers for validation of IPCS expression in yeast mutants by colony PCR:

*At*IPCS1 FP: GGTTCTGAGCTTGCCAGG

*At*IPCS1 RP: TGCCGTTGTTCGATCAGG

*At*IPCS2 FP: CCTAACTATCACTGCCGT

*At*IPCS2 RP: GAGCAACACAGCAGTCCG

*At*IPCS3 FP: CGAGAAGGCTCCAAGCTC

*At*IPCS3 RP: GGTCCACTAGAACGTTCT

*Os*IPCS1 FP: TCTTGCTTGGCTGATGGC

*Os*IPCS1 RP: GTGGTTGATGTCCTCGCC

*Os*IPCS2 FP: GCTTCTTGCTTGGTTCAT

*Os*IPCS2 RP: GGCGTGATTGATGTCTTC

*Os*IPCS3 FP: GGTGCGGTGACTTGATAT

OsIPCS3 RP: CAATGTGCTTCCCATTCA

AUR1 FP: GGTCCACCAACTGTTTTGC

AUR1 RP: CTCAATTGAAGTGTATGACC

Primers for *At*IPCS 1,2, and 3 cloning into pENTR/D-TOPO:

*At*IPCS1-NotI-F: GCGCGCGGCCGCCACAATGACGCTTTATATTCGCCGCG

*At*IPCS1-AscI-R: GCGCGGCGCGCCTCATGTGCCATTAGTAGCATTATCAGTGTG

*At*IPCS2-NotI-F: GCGCGCGGCCGCCACAATGACACTTTATATTCGTCGTGAATCTTCCAAG

*At*IPCS2-AscI-R: GCGCGGCGCGCCTCACGCGCCATTCATTGTGTTATC

*At*IPCS3-NotI-F: GCGCGCGGCCGCCACAATGCCGGTTTACGTTGATCGC

*At*IPCS3-AscI-R: GCGCGGCGCGCCTCAATGATCATCTGCTACATTGTTCTCGTTT

Primers to quantify *At*IPCS 1,2, and 3 transcripts in the overexpression transgenic lines using real-time PCR:

*At*IPCS1_F: TGCGTCCCGTAAACATTACA
*At*IPCS1_R: ACACCGTTCCCATTCAAGAG
*At*IPCS2_F: TACCAGATCGGACTGCTGTG
*At*IPCS2_F: GTGAACTCCGTTGCTGTCAA
*At*IPCS3_F: CTGGGCCGAATTATCATTGT
*At*IPCS3_R: CCTTCGTGTGCCGTATCTTT

Primers to quantify *At*IPCS 1,2, and 3 transcripts in the RNAi transgenic lines using real-time PCR:

*At*IPCS1_F: TGGTTGTCTTCTTCCTCGAC
*At*IPCS1_R: GAGCAGAGATCTCATGTGCC
*At*IPCS2_F: TGGTGGTGTTCTGTCTAGAC
*At*IPCS2_F: GATCTTTTCACGCGCCATTC
*At*IPCS3_F: GCATGGTATACTGTGAACCT
*At*IPCS3_R: GGTGATCGGAGACATAAG

APPENDIX F – RNA quality control for RNASeq experiment

Col 0 (1) — RIN: 6.90
Col 0 (2) — RIN: 7.70
Col 0 (2) — RIN: 7.50
At2 over exp 9 (1) — RIN: 7.50
At2 over exp 9 (2) — RIN: 7.60
At2 over exp 9 (3) — RIN: 7.70
At1 tDNA insert mutant (1) — RIN: 7.70
At1 tDNA insert mutant (2) — RIN: 7.30
At1 tDNA insert mutant (3) — RIN: 7.70
At3 over exp 4 (1) — RIN: 7.60
At3 over exp 4 (2) — RIN: 7.50
At3 over exp 4 (3) — RIN: 7.60