How Does *Plasmodiophora brassicae* Manipulate Its Host Development and Metabolism?

Sylivere Habumugisha

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy

School of Biosciences

Faculty of Science

University of Sheffield

Submitted: 28th March 2025

Acknowledgements

It is with heartfelt gratitude that I first thank almighty God, for granting me good health, patience, and courage throughout the study period. I am genuinely thankful to my supervisor Prof Stephen Rolfe for accepting me as a PhD student. If it wasn't for your intervention when my scholarship was hanging in balance, I wouldn't have had the opportunity to do a PhD at the university of Sheffield. Thank you for your willingness to supervise my research, for your guidance and constructive criticisms throughout the study period. I highly appreciate the moral support you have provided to me throughout my PhD.

I would like to thank the Commonwealth Scholarship commission (CSC) for funding my PhD and offering valuable training and networking opportunities. I am grateful for this opportunity to study at university of Sheffield. Thank you, Dr Gregory J S Fowler, for your guidance and insightful discussions. Your encouragement and valuable comments along with penetrative questions from early stage of the project to the final thesis write up, taught me to ponder more deeply. Thank you for transforming my molecular biology skill set in the lab. Your continuous scientific and technical support allowed me to have such a fruitful and fulfilling PhD.

Thanks go to my fellow PhD students and researchers in the plant-biotic interactions group for their support from the moment I first stepped into the lab. My heartfelt thanks go to my friends in Sheffield, Rwanda, and beyond, for their constant moral support, especially the many jokes that helped me get through the toughest times.

I am deeply grateful to my wife, Sylvie, for her unwavering love, care, understanding, endless patience, and constant encouragement, especially when I needed it most. A special thanks also goes to my sons, Ian and Yannis, whose evening playtime provided a much-needed escape from the long, exhausting hours spent in the lab.

Finally, I would like to express my deepest appreciation to my parents for their invaluable guidance and support. I am deeply grateful for the sacrifices you've made to provide me with a supportive upbringing and the education that made this PhD possible. To my siblings, I am truly thankful for your unwavering belief in me, constant encouragement and support that kept me going in the challenging times of my PhD study.

Declaration

The work presented in this thesis is original and has not been submitted for examination at the University of Sheffield or any other institution. I acknowledge the university's guidance on the use of unfair means (www.sheffield.ac.uk/ssid/unfair-means) and confirms that all work included in this thesis is solely the my own. In chapter three, Figure 29 has been adapted from Dr Gregory J.S Fowler unpublished work on characterisation of PRO-EU transporters in which I contributed to during my PhD. I acknowledge the use of the generative AI tool Google Gemine for feedback on the phrasing and grammar of text in line with this Guidance.

Thesis summary

Plasmodiophora brassicae (Clubroot) is an intracellular biotrophic pathogen responsible for clubroot disease in Brassicaceae plants, such as Brussels sprouts and oilseed rape, which significantly impact crop production. However, the interactions between the pathogen and its host remain poorly understood. Little is known about the chemical nature of the nutrient compounds and the pathogen transporters involved in their uptake from host plants. While *P.brassicae* is believed to produce cytokinins (CK) during disease development, studies on phytohormone levels in infected plants have yielded mixed results.

In this thesis, I utilised bioinformatics approaches to identify potential *P.brassicae* transporters involved in nutrient acquisition from the host. I show that most existing *P.brassicae* gene models are inaccurate and require experimental validation before genomic data can be used for functional studies. I also demonstrate that a large portion of the *P.brassicae* genome is dedicated to encoding transporters belonging to various transport systems and families, particularly when compared to other obligate biotrophic pathogens. Moreover, I identify a novel class of transporters which is not common in eukaryotes, including two transporters in the SWEET family. These transporters are expressed at different stages of *P.brassicae* development and may provide the pathogen with a competitive advantage over its host.

Lastly, I employ genetic approaches to explore the role of *P.brassicae* IPT genes during disease development. I reveal that these IPT genes belong to the tRNA-IPT class and that both *P.brassicae* infection and the overexpression of IPT genes in Arabidopsis lead to CK-specific phenotypes. Together, these findings and the *P.brassicae* genome annotation resource developed in this thesis provide a solid foundation for future research on the role of *P.brassicae* nutrient transporters, the CK produced by the pathogen, and the broader function of IPT genes in the development of clubroot disease.

Table of contents

List of figures

List of tables

List of supplementary figures

List of supplementary tables

Due to the size of the Supplementary Data tables, all files have been provided as separate excel workbook in archive folder. All file names are listed below:

Table S 1: Summary of *P.brassicae* transporter families and their substrates

Table S 2: Summary of *P.brassicae* TCDB transporter families

Table S 3: *P.brassicae* transporters and their putative substrates

Table S 4: Unusual *P.brassicae* transporters

Table S 5: List of organisms and protein sequence IDs used in phylogenetic analysis

Table S 6: Primers used in the study

Abbreviations and Acronyms

10X	10X genomics
$^{14}CO_2$	Radioactive carbon 14 isotope
AA	Arachidonic acid
ABC	ATP binding cassette transporter superfamily
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BNYVV	Beet necrotic yellow vein virus
BoSTP	Brassica oleracea sugar transporter
CAZymes	Carbohydrate-active enzymes
cDNA	Complementary DNA derived from RNA
CKs	Cytokinins
CKS1	Cytokinin Synthesis 1CoA Coenzyme A
Col-0	Columbia-0
CR	Clubroot resistance
cZ	Cis-zeatin
dH2O	Deionised water
DMAPP	Dimethylallyl pyrophosphate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dpi	Days post inoculation
DZ	Dihydrozeatin
EC	Enzyme commission
ECD	European Clubroot Differential set
ET	Ethylene
FADH2	Flavin adenine dinucleotide
FBRNT	Freebase, riboside and nucleotide CKs
FPKM	Fragments per kilobase of transcripts per million mapped reads
G1-to-S	Cell cycle stage G1 phase (gap1) in which the cell grows, and the S phase (synthesis), during which DNA is replicated
G2-M	Cell cycle transition checkpoint between S phase and the M (mitosis) phase
gDNA	Genomic DNA
GFP	Green fluorescent protein

GPCR	G Protein Coupled Receptors
IGV	Integrated genome viewer
iP	Isopentenyl adenine
iPA	Isopentenyl adenosine
IPT	Isopentenyl transferase
ipt1,3,5,7	Arabidopsis thaliana quadruple mutant of isopentenyl transferases
IUBMB	International Union of Biochemistry and Molecular Biology
JA	Jasmonic Acid
LB	Luria-Bertani broth
Mb	Megabase
MCSs	Membrane contact sites
MeSA	Methyl salicylate
Min	minutes
mNAT	Mouse histidine and glutamine amino acid transporter
mRNA	messenger RNA
MSF	Major facilitator superfamily
MYCOCOSM	The fungal genomics resource
NADH	Nicotinamide adenine dinucleotide
NCBI	National Centre for Biotechnology Information
ONT	Oxford nanopore technology
PbIPTs	Plasmodiophora brassicae isopentenyl transferases
PCR	Polymerase Chain Reaction
PCV	Peanut clump virus
PFAM	Protein family database
pH	Potential of Hydrogen
PI	Propidium iodide
PMSF	Phenylmethylsulfonyl Fluoride
PMTV	Potato mop-top virus
PRO-EU	Clubroot transporters that has parts of prokaryotic and eukaryotic transporters
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
RNAseq	RNA sequencing
SA	Salicylic Acid
SBD	Substrate binding domain

SBP	Substrate binding protein
scRNAseq	Single cell RNA sequencing
SD	Synthetic dropout media
Srt1	Sucrose transporter
STP	Sugar transporter
SUCs/SUTs	Sucrose transporters
SWEET	Sugars Will Eventually be Exported Transporters
TAGs	Triacylglycerol
TC	Transport classification
TCA cycle	Tricarboxylic Acid (TCA) cycle
TCDB	Transporter classification database
TCID	Transport classification identity code
tRFs	tRNA-derived RNA fragments
tRNA	Transfer RNA
tRNA-IPT	tRNA-Isopentenyl transferase
tZ	Trans-zeatin
UK	United Kingdom
YNB	Yeast nitrogen base without amino acids
YPD	Yeast extract peptone dextrose

CHAPTER 1: GENERAL INTRODUCTION

1.1. Plasmodiophora brassicae and clubroot disease

Plasmodiophora brassicae is an intracellular obligate biotrophic pathogen that causes clubroot disease, one of the most damaging diseases within the family Brassicaceae and which can also infect the model plant *Arabidopsis thaliana*. The resting spore stage is the only stage that takes place outside the host in soil and is the only non-parasitic stage in the pathogen's lifecycle. The disease exists in more than 60 countries and is a major challenge in areas where brassica crops are grown on a large scale (Dixon, 2009). Clubroot-infected plants are dwarfed compared to healthy plants and their roots grow abnormally large into galls. The mature galls remain in soil at harvest where they rot and disintegrate thereby releasing spores from the plant tissue and remain infectious in the soil for 18 years or more rendering the fields unsuitable for brassica cultivation (Wallenhammar, 1996).

1.2. Taxonomic classification of *P.brassicae*

For a long time *P. brassicae* was classified as a fungus because of the chitinous cell wall of resting spores (Javed *et al.*, 2023). With the availability of molecular data as more organisms are sequenced, *P. brassicae* is now classified as a protist in the phytomyxea (Bass *et al.*, 2009; Burki *et al.*, 2010; Neuhauser *et al.*, 2014). The current classification places *P. brassicae* in the eukaryotic supergroup of Rhizaria, Phylum: Cercozoa; Class: Phytomyxea; Order: Plasmodiophorales; Family: Plasmodiophoraceae; Genus: *Plasmodiophora*; Species: *Plasmodiophora brassicae* (Bass *et al.*, 2009; Burki *et al.*, 2010; Javed *et al.*, 2023). The class phytomyxea consists of obligate biotrophic pathogens of plants and stramenopile (Hittorf *et al.*, 2020; Javed *et al.*, 2023). Apart from *Plasmodiophora brassicae*, another major economic plant pathogen in phytomyxea is *Spongospora subterranean* which causes potato powdery scab disease. Phytomyxea are also vectors of plant viruses such as watercress yellow spot Virus, beet necrotic yellow vein virus (BNYVV), Potato mop-top virus (PMTV), wheat mosaic virus and the peanut clump virus (PCV) (Hittorf *et al.*, 2020).

1.3. The life cycle of *P. brassicae*

P. brassicae has a complex life style which involves three main stages: a dormant resting spore stage in the soil, root hair infection stage and cortical infection stage (Kageyama & Asano, 2009).

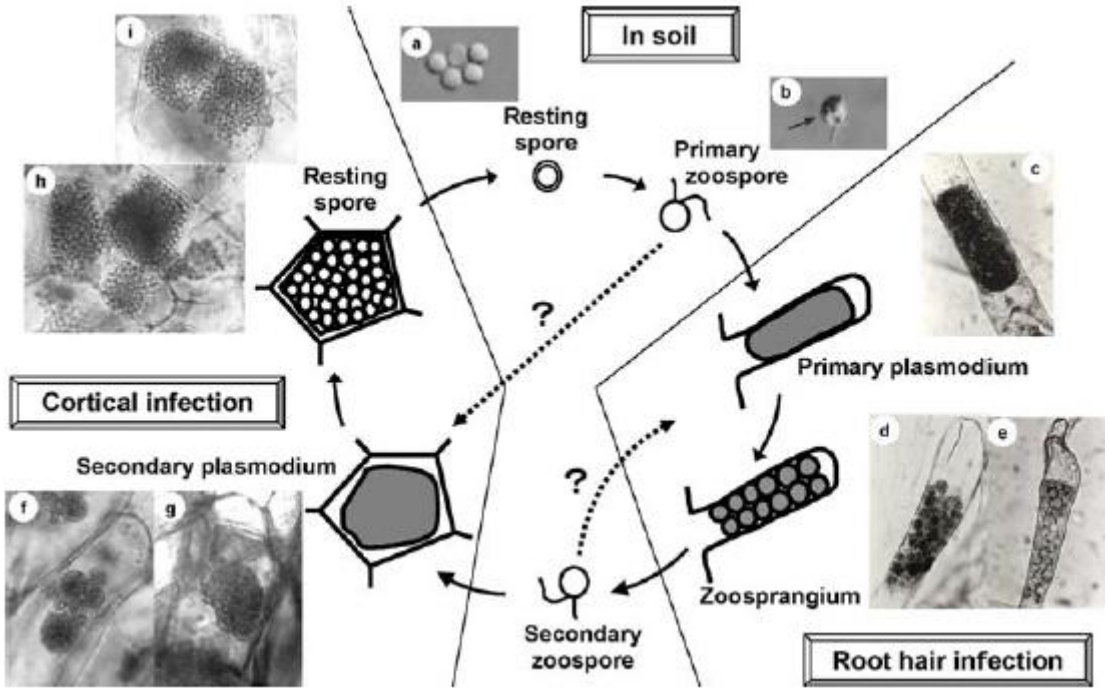


Figure 1: The life cycle of Plasmodiophora brassicae.
The sketch pictures a-i shows the pathogen transition steps from free-living resting spores in the soil to parasitic primary and secondary plasmodia stages in the host. a) Resting spore, b) primary zoospore, c) primary plasmodia in the root hair d) Zoosporangial cluster in root hair, e) Empty zoosporangium, f,g) Secondary plasmodia in cortical cells, h, i)Resting spores in cortical cells. Dotted lines and question marks indicates that there is a possibility of primary zoospores directly infecting the root cortex and the secondary zoospores re-infecting the root hairs but there is not enough evidence to support these hypotheses. Taken from (Kageyama & Asano, 2009)

Resting spores in the soil germinate into biflagellate primary zoospores which swim in the soil water film to the host. The primary zoospore starts the root hair infection stage by penetrating the root hair cell walls. Within the root hair, the primary zoospore undergoes several nuclear divisions to form primary plasmodia. The primary plasmodia then form zoosporangia with each zoosporangia containing four to sixteen zoospores (Kageyama & Asano, 2009; Hwang *et al.*, 2012). The root hair infection stage is symptomless and does not cause significant yield and quality loss (Howard *et al.*, 2010). While the traditional view holds that zoospores released into the soil from zoosporangia initiate cortical infection, recent evidence suggests that primary zoospores can directly infect cortical cells, despite the indistinguishable appearance of primary

and secondary zoospores (Feng *et al.*, 2013b). The zoospores germinate into secondary zoospores which penetrate the epidermal cells of the main root and eventually colonise the root cortical cells. It is not known if *P.brassicae* has an apoplastic stage and how secondary plasmodia move within the host cells remains unclear. Some root colonising plant pathogens use the host vascular tissues to colonise distant cells from the site of initial infection (Malinowski *et al.*, 2024) but *P.brassicae* might not use this mechanism to spread in distal host root cells as *P.brassicae* secondary plasmodia has not been isolated from the xylem or phloem sap of infected plants. *P.brassicae* has been proposed to move within the infected cells by host cell cytoplasmic streaming (Asano & Kageyama, 2006). This claim is supported by the presence of secondary zoospores in the cortical cells which had moved from the epidermal cells with 3 days after inoculation (Feng *et al.*, 2013b; Liu *et al.*, 2020). Asano and Kageyama speculated that the increase in the number of *P. brassicae*-infected cells was due to the division of these infected cells. They proposed that if the pathogen material was localized to a single site within the cell, there would be a low probability of both newly divided daughter cells inheriting the pathogen. However, if the pathogen material was evenly distributed in the host cell cytoplasm, the chances of both daughter cells containing the pathogen would be high, thereby increasing the number of infected host cells. Based on these observations, the infected cells should be clustered together but microscopic observations by Malinowski et al (2012) revealed that *P.brassicae* plasmodia are randomly distributed surrounded with patches of xylem and phloem tissues. Another grey area in the lifecycle of *P.brassicae* is its mode of reproduction. It is not clear if *P.brassicae* undergoes karyogamy but binucleated zoospores have been found (Hwang *et al.*, 2012). Once in the cortex, the secondary zoospores grow into secondary plasmodia. As the secondary plasmodia grows, the infected cells start to swell, followed by development of galls in the roots causing the roots to develop club-shaped malformations, hence the name of the disease. Secondary plasmodia develop and form millions of resting spores in the host roots. The infected roots rot and disintegrate releasing the resting spore into the soil completing the disease cycle.

1.4. Clubroot disease management practices

1.4.1. Cultural control

Prevention. Preventing a field from being infected with *P.brassicae* remains the best management option. Clubroot is a soil borne disease and once the field is infested with clubroot, management practices will minimize the disease incidence, severity and the economic loss of the disease but will not eliminate the pathogen. Clubroot is mostly transmitted from infested fields to clubroot-free fields mainly by the movement of humans and machinery. This mode of transmission is supported by the observation that areas with clubroot infection are usually found within 300m from the field entrance although isolated clubroot infested spots have been found in the middle of fields (Howard *et al.*, 2010). Avoiding sharing of farm equipment and machinery where possible and proper sanitation of any farm equipment, machinery and personal clothing such as footwear can prevent transmission of clubroot to health fields. Use of planting materials free from *P.brassicae* resting spores can also help to prevent clubroot spread. The isolated clubroot infested spots in the middle of the fields can be attributed to the use of *P.brassicae* contaminated material as *P.brassicae* DNA has been detected on seeds from clubroot-infested fields (Cao *et al.*, 2007). Where irrigation is practiced, it is important to use clean water, free from *P.brassicae* resting spores, as the spores can move into runoff water from infested fields and thus contaminating the water source(Howard *et al.*, 2010).

Field selection, soil amendments and planting dates. Although *P.brassicae* spores can germinate on all types of soil, clubroot disease incidence and severity varies among different soil types (Wallenhammar, 1996). Field conditions such as soil water content, temperature and pH are essential for clubroot disease development. Soil moisture provides a medium through which biflagellate primary zoospores swim through to potential host plants. Clubroot disease incidence was reported to increase with increase in rainfall (Gossen *et al.*, 2012). Clubroot disease severity has been reported to increase with temperature with severe disease symptoms developing at 25°C (Sharma *et al.*, 2011; Gossen *et al.*, 2012). Soil pH is also important for clubroot disease development and soil amendments that change soil pH can be used to reduce the disease pressure. Clubroot disease severity is high in acidic soil (Wallenhammar, 1996) but spore density and disease severity is decreased in alkaline soils (Fox *et al.*, 2022; Hennig *et al.*, 2022). However, the effectiveness of liming varies between seasons and liming material (Gossen *et al.*, 2014; McGrann *et al.*, 2016; Hennig *et al.*, 2022). Careful selection of fields with good drainage for brassica cultivation, application of appropriate liming material and planting dates could all minimise the loss from clubroot disease. It should be noted that liming

could affect the choice of crops to rotate with brassica and it is expensive practice considering the size of fields where brassica crops are grown.

Cultivar selection. Major clubroot resistance (CR) genes and quantitative trait loci (QTLs) conferring clubroot resistance have been identified across diverse Brassica species. Importantly, this resistance is often pathotype-specific, meaning it is effective against particular strains of *P. brassicae* that possess differing virulence(Hwang *et al.*, 2012; Botero-Ramirez *et al.*, 2024). Different Brassica species house distinct sets of these resistance genes. For example, majority of CR genes, including Crr1a (Hatakeyama *et al.*, 2013), Crr3 (Saito et al., 2006), CRb (Zhang et al., 2014), Rcr1 (Chu et al., 2014), Rcr3 and Rcr9wa (Karim et al., 2020), are primarily observed in B. rapa, while clubroot resistance QTLs genes are mainly found in other brassicas (Voorrips et al., 1997; Rocherieux et al., 2004; Yu et al., 2017). Breeding for clubroot resistance in brassicas is challenging due in part because of the emerging pathotypes quickly break resistance in clubroot resistant cultivars within two to four years after the release of the resistant cultivars (Hollman *et al.*, 2021; Botero-Ramirez *et al.*, 2024) and limitations of gene introgressions within brassicas (Rahman, 2013). Despite these challenges, clubroot resistant cultivars have been on the market from early 2000s with 75 clubroot resistant cultivars in Canada alone (Botero-Ramirez *et al.*, 2024). Cultivation of susceptible cultivars increases the spore density in infected fields (Hwang *et al.*, 2011). The increase in spore density in fields infected by clubroot is associated with large galls formed on roots of susceptible canola cultivars that are sometimes 250 fold bigger than galls formed on resistant cultivars (Hwang *et al.*, 2015). Growing resistant cultivars however, reduces the disease incidence and severity by two fold compared to susceptible cultivars (Hwang *et al.*, 2011). Growing clubroot resistant cultivars adapted to local pathotypes where available is recommended and remains the most successful clubroot management practice (Howard *et al.*, 2010; Javed *et al.*, 2023; Botero-Ramirez *et al.*, 2024).

Crop rotation and weed control. Crop rotations are designed to reduce the spore inoculum in the soil (Howard *et al.*, 2010). The *P.brassicae* spore inoculum is increased by continuous growing of resistant or susceptible cultivars (Hwang *et al.*, 2015) and keeping brassica weeds on the field as it believed that *P.brassicae* infects all brassicas (Javed *et al.*, 2023). *P.brassicae* spores can live in soil and remain infectious for up to 17 years (Wallenhammar, 1996) but their half-life is estimated to be 3.6years (Ernst *et al.*, 2019) meaning that 50% of the spores are not viable after 3.6 years although a 90% decrease in spore density has been reported after two years of non-host cultivation (Ernst *et al.*, 2019). Different rotation sequences with a

combination of non-host crops, resistant and susceptible cultivars and fallow have produced promising results in reducing *P.brassicae* spore density, disease incidence and severity (Hwang *et al.*, 2015; Peng *et al.*, 2015; Hwang *et al.*, 2019). Growing canola after a three-year non-host rotation increased yield by 3600% (Hwang *et al.*, 2019) thus a more than two year non-host rotation is recommended as a standard.

Disease monitoring. Continuous monitoring of clubroot disease in the fields from sowing to harvest could help farmers identify hot spots and take actions that minimise the risk of transmitting the disease to the rest of the field. Plants suspected to be infected with *P.brassicae* should be examined for both above and below ground clubroot symptoms as other soilborne diseases could cause the same above ground symptoms as clubroot (Howard *et al.*, 2010).

Cultivar mixtures. Continuous cultivation of resistant cultivars in clubroot-infested fields increases selection pressure on the pathogen resulting in the breakdown of resistance genes and emergence of highly virulent pathotypes (Strelkov *et al.*, 2016; Hollman *et al.*, 2021; Javed *et al.*, 2023). Cultivar mixtures, a disease and pest management practice that is used in cereal crops (Borg *et al.*, 2018; Duan, XY *et al.*, 2022) has been proposed as a new clubroot management practice (Botero-Ramirez *et al.*, 2024). The idea of using cultivar mixtures in the clubroot management leverages the potential of preserving clubroot population structure by promoting the coexistence of virulent and avirulent pathotypes thereby minimising the population shift towards highly virulent resistance-breaking pathotypes (Botero-Ramirez *et al.*, 2024).

1.4.2. Biological control

Biocontrol microorganisms. Biological control of *P.brassicae* with microorganism is an attractive clubroot disease management practice as it is environmentally friendly. Biocontrol of *P.brassicae* with microorganisms supressed the disease on canola under controlled environment but the disease suppression was lost when the inoculum density was high in soil (Peng *et al.*, 2014). Several studies have reported decrease in clubroot disease incidence and severity when endophytic fungi strains such as *Trichoderma* spp. ReTk1 and ReTv2 (Hasan *et al.*, 2023), *Trichoderma harzianum* (Gulzar et al., 2024), *Acremonium alternatum* (Auer *et al.*, 2024), *Didymella macrostoma P2* (Cheng *et al.*, 2024) and bacterial strains *Paenibacillus chitinolyticus* and *Sophora flavescens* (Gulzar et al., 2024; Rudsari et al., 2024) were used as clubroot biocontrol agents. However, most of the studies were done in greenhouse conditions. Despite the fact that efficacy of biocontrol microorganisms vary between greenhouse conditions and field conditions (Peng *et al.*, 2014) and also between cultivars (Auer *et al.*, 2024), biocontrol

agents against *P.brassicae* are promising and may play an important role in the clubroot disease management especially in fields with low inoculum density.

Trap crops. Trap crops are crops grown to induce the germination of resting spores without being infected hence reducing the spore inoculum in the soil (Javed *et al.*, 2023). *Allium porrum* , *Secale cereale* and *Lolium perenne* has been reported as bait crops for *P.brassicae* (Zamani-Noor et al., 2022; Javed et al., 2023). A screen of 86 plant species against clubroot showed that some brassica species such as *B. orientalis, C.* squanatus and *R. sativus* as well nonhost plants including *P.sativum*, *P. rhoeas* and *A. myosuroides* which were not known to be infected by *P.brassicae* are infected but do not grow root galls and thus they can be used as bait crops to trigger the germination of *P.brassicae* resting spores (Zamani-Noor *et al.*, 2022). Plants which are infected by *P.brassicae* that forms small, or no root galls are potential trap crops and can reduce the *P.brassicae* spore inoculum even if the pathogen completes its life cycle as they form small or no galls and the number of spores in the gall is proportional to the gall size.

1.4.3. Chemical control

Chemical control of soilborne diseases is always challenging as the disease-causing microorganisms may not be in reach of chemicals due to soil environment. Nonetheless, a number of fungicides are able to reduce the severity of clubroot disease when applied as soil treatments by spot drenching, banding, broadcasting, irrigation or as seed treatment (Botero-Ramirez *et al.*, 2024). Soil fumigation has been shown to reduce clubroot severity (Hwang *et al.*, 2018; Zuzak *et al.*, 2022). Chemical fungicides can negatively impact other soil organisms, human health, and the environment. Additionally, fumigants can be expensive to apply, especially for large fields. These factors make chemical control of clubroot an impractical option and should be considered only as a last resort.

1.5. Physiological impacts of *P.brassicae* infection on host development and metabolism

1.5.1. Impact on regulation of host cell cycle.

The *P.brassicae* infected roots of brassica plants and, in case of *Arabidopsis thaliana,* roots and hypocotyls undergo abnormal cell enlargement and uncontrolled cell division leading to the development of galls. This characteristic growth points to the involvement of plant hormones during infection and disease development. Auxin and cytokinin, the major plant hormones that control most aspects of plant developmental processes are thought to play a major role in clubroot disease symptom development (Ludwig-Müller *et al.*, 2009; Schuller *et al.*, 2014).

Cytokinins (CKs) are naturally occurring isoprenoid and aromatic compounds produced by plants but have also been found to be synthesised by other organisms (Akhtar *et al.*, 2019). T*rans*-zeatin (*t*Z), isopentenyl adenine (iP), *cis*-zeatin (*c*Z) and dihydrozeatin *(DZ)* are the four naturally occurring CKs that are common in plants, but their amount varies between plant species (Sakakibara, H., 2006). The role of cytokinin in clubroot disease development has interested researchers for several decades (Dekhuijzen and Overeem 1971; Müller and Hilgenberg 1986; Siemens et al. 2006; Devos et al. 2006; Malinowski et al. 2016). Cell division activity is an important factor in gall formation during clubroot disease development. *P.brassicae* may regulate host cell cycle regulation by producing cytokinin which is important in the regulation of G1-to-S transition in plants (Inzé & De Veylder, 2006). Devos et al. (2006) postulated that at the very first stages of *P.brassicae* infection, plasmodial-produced cytokinins trigger a local re-initiation of cell division in the root cortex establishing a de novo meristematic area that acts as a sink for host derived indole-3-acetic acid, carbohydrates, nitrogen, and energy to maintain the pathogen and to trigger gall development. However, Malinowski *et al* (2012) showed that this was incorrect, and that gall formation occurs because of reprogramming of existing meristems i.e. the vascular cambium that leads to secondary growth in uninfected plants. This view has been further supported by the findings that *P.brassicae* promotes host tissue proliferation by delaying G2-M cell cycle transition checkpoint (Olszak *et al.*, 2019). Whether the reprograming of the host cell cycle is entirely due to the changes in host CKs homeostasis is debatable. Devos *et al* (2006) reported that *P.brassicae* infected Arabidopsis accumulated isopentenyl adenosine (iPA), an active cytokinin which was twofold higher than in control plants at four days post inoculation (dpi). However, the experimental evidence is poor. There is no evidence for an early increase in CKs as the host cytokinin biosynthesis and cytokinin oxidase/dehydrogenase genes were repressed as early as 7dpi (Siemens et al. 2006; Malinowski et al. 2016). Although *P.brassicae* contains CK biosynthetic genes (Rolfe et al. 2016) and could synthesise small amounts of CKs (Müller and Hilgenberg 1986; Malinowski et al. 2016), their impact on the Arabidopsis host remains unclear as *P.brassicae* infection of atipt1;3;5;7 mutants that are severely restricted in CK biosynthesis could not rescue this phenotype (Malinowski et al. 2016) and thus the role of CKs in gall formation remains unclear. However, *P.brassicae* produced CKs may be important in other Brassicas.

1.5.2. Impact of *P.brassicae* infection on host carbohydrate partitioning

During the cortical invasion infection stage, *P.brassicae* interferes with host developmental processes which leads to the formation of galls in the roots (Kageyama & Asano, 2009). The

galls are therefore a biotrophic interaction zone, providing *P. brassicae* a location for shelter and for nutrient acquisition (Rolfe *et al.*, 2016; Sosso *et al.*, 2019). At this stage the pathogen is metabolically active and able to utilize host energy sources - carbon, nitrogen, amino acids and other nutrients (Ludwig-Müller *et al.*, 2009).

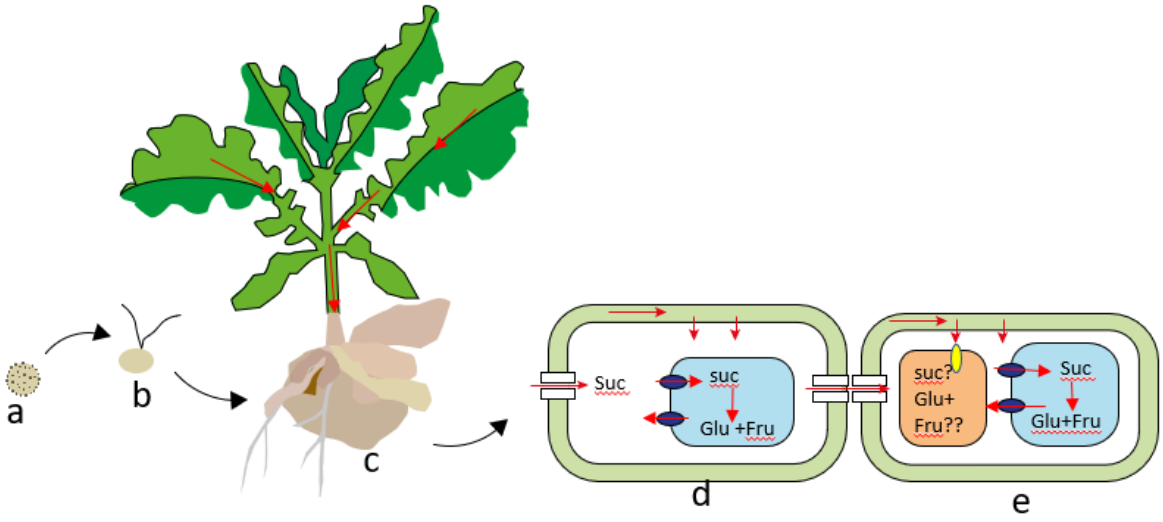


Figure 2: Diagram showing infection of host plant by P.brassicae and the partitioning of photosynthates and nutrients in infected roots and P.brassicae .
P.brassicae resting spores in soil (a) germinate into biflagellate primary zoospore (b). The zoospore infects root hairs and eventually enters the cortex causing gall formation(c). The red arrows show the movement of photosynthates from the leaves to the roots. After phloem unloading, the nutrients move from uninfected plant cells (d) via the apoplast (the light-green coloured space) or plasmodesmata (space between the two white rectangles) to the P.brassicae infected plant cells (e). P.brassicae plasmodia then take up nutrients from the plant cell cytoplasm via its nutrient transporters. the sky blue rectangle represents plant cell vacuole, the orange rectangle represents P.brassicae plasmodia and the blue and yellow spheres represents plant and P.brassicae sugar transporters.

Roots and hypocotyls of *P.brassicae* infected Arabidopsis plants were found to accumulate high levels of glucose, fructose and starch (Brodmann et al. 2002). A sugar translocation study (Keen and Williams 1969) revealed a 4-fold increase of glucose in the hypocotyls of infected cabbage. Fructose concentrations also increased about 2-fold in the diseased hypocotyls respectively. Pulse feeding of radioactive $^{14}CO_2$ revealed that radioactive sucrose was the most abundant sugar in the galls. These results are in corroboration with Evans and Scholes (Evans and Scholes 1995) who showed that *P.brassicae* infected Arabidopsis thaliana leaves had less sucrose compared to uninfected plants during daylight implying that sucrose was rapidly metabolized or exported from the leaf. Starch also accumulates in the hypocotyls and galls following *P.brassicae* infection (Evans and Scholes 1995; Brodmann et al. 2002; Walerowski et al. 2018).

Biotrophic pathogens use different strategies to translocate photosynthates preferentially to infection site. One of the common strategies is recruitment of host sugar transporters to the infection site (Chen *et al.*, 2010). *P.brassicae* has been shown to enhance sink strengths of the galls by recruiting host transporters and enzymes involved in carbohydrate transport and metabolism. Transcriptome analysis of *Brassica oleracea* var. capitate (Zhang et al. 2019) showed that monosaccharide sugar transporters BoSTP4b and BoSTP12 genes were upregulated in infected plants. Consistent to this, *P.brassicae* infection of Arabidopsis and *Brassica rapa* affected the distribution of hexoses and sucrose resulting in the formation of a strong physiological sink in the roots with an upregulation of the expression of STP8, STP13 SWEET11 and SWEET12 facilitating local distribution of sugars toward the pathogen (Walerowski et al. 2018; Li et al. 2018). Sucrose synthase, an enzyme that breaks down sucrose, and extracellular invertase were upregulated in infected *Arabidopsis thaliana* (Siemens *et al.*, 2006; Siemens *et al.*, 2011). *P.brassicae* not only recruits host sugar transporters in the galls, it also disrupt vascular differentiation promoting phloem formation at the expense of xylem formation (Malinowski *et al.*, 2012).

P.brassicae may also use the pathogen-produced cytokinin to regulate source-sink relationship in the infected plants. Studies on how CKs regulate source-sink relationships have been mainly carried on leaves by treating the leaves with CKs, inoculations with CK producing pathogens and expression of CKs biosynthesis genes. Mothes and Engelbrecht, (1961) reported the movement and accumulation of amino acids in the parts of the leaves that were treated with CK. Inoculation of rice leaves with *Magnaporthe oryzae* expressing tRNA-IPT named Cytokinin Synthesis 1 (CKS1) gene increased the concentration of glucose and fructose in and around the inoculation site whereas inoculation with mutants deleted for this gene did not affect sugar distribution (Chanclud et al. 2016). Axillary bud localized overexpression of cytokinin biosynthesis gene Isopentenyl transferase (IPT) in tobacco resulted in the accumulation of starch thus causing local sink enhancement (Guivarc'h et al., 2002). Cytokinins may also modulate redistribution of photosynthates by the induction of extracellular invertases. Induction of extracellular invertase was observed in axillary buds with high CKs and high starch content (Guivarc'h et al., 2002) and caused delayed leaf senescence in tobacco leaves expressing IPT gene under the control of senescence induced promoter (Balibrea-Lara et al., 2004). High levels of cytokinins in a particular tissue makes it a potential metabolic sink (Mothes & Engelbrecht, 1961) thus localised plasmodial-produced cytokinins would make the infected plant cells a new metabolic sink by enhancing the movement of metabolites to the galls thereby affecting the

normal partitioning of photosynthates in the infected plant. Despite being a major pathogen of cruciferous plants, many aspects of *P. brassicae*-host interactions remains obscure (Feng et al. 2012). There is much to learn about the nutrition of this important biotrophic pathogen in all stages of its lifecycle. Understanding of, and elucidation of *P.brassicae* carbon metabolism mechanisms could contribute to development of novel strategies in resistance breeding or clubroot management.

1.6. *P.brassicae* carbon metabolism

Living organisms use different metabolic pathways to generate energy in the form of ATP which supports cellular functions and survival and for the synthesis of intermediate compounds needed for cellular growth and proliferation (Al-Khami *et al.*, 2017). Tissues have different abilities to utilise different metabolites thus the metabolic pathways are selected based on metabolic fuels available (Judge & Dodd, 2020). The metabolic pathways that may be activated in spores and the secondary plasmodia (Figure 1) are different due to different metabolic fuels present in these stages of the pathogen. The resting spores are metabolically inactive but they must contain energy reserves that fuel metabolic activities that take place between the emergence of biflagellate zoospores and zoospore encysting on the root hairs of the host plant (Dixon 2014). *P.brassicae* converts carbohydrates for energy production and biosynthesis intermediates mainly via glycolysis, glyoxylate, pyruvate metabolism, pentose phosphate pathway and the TCA cycle (Schwelm *et al.*, 2015). Carbon metabolism in the germinating spores that produce the primary zoospores is mainly via the glyoxylate pathway (Schwelm *et al.*, 2015) (Figure 3).

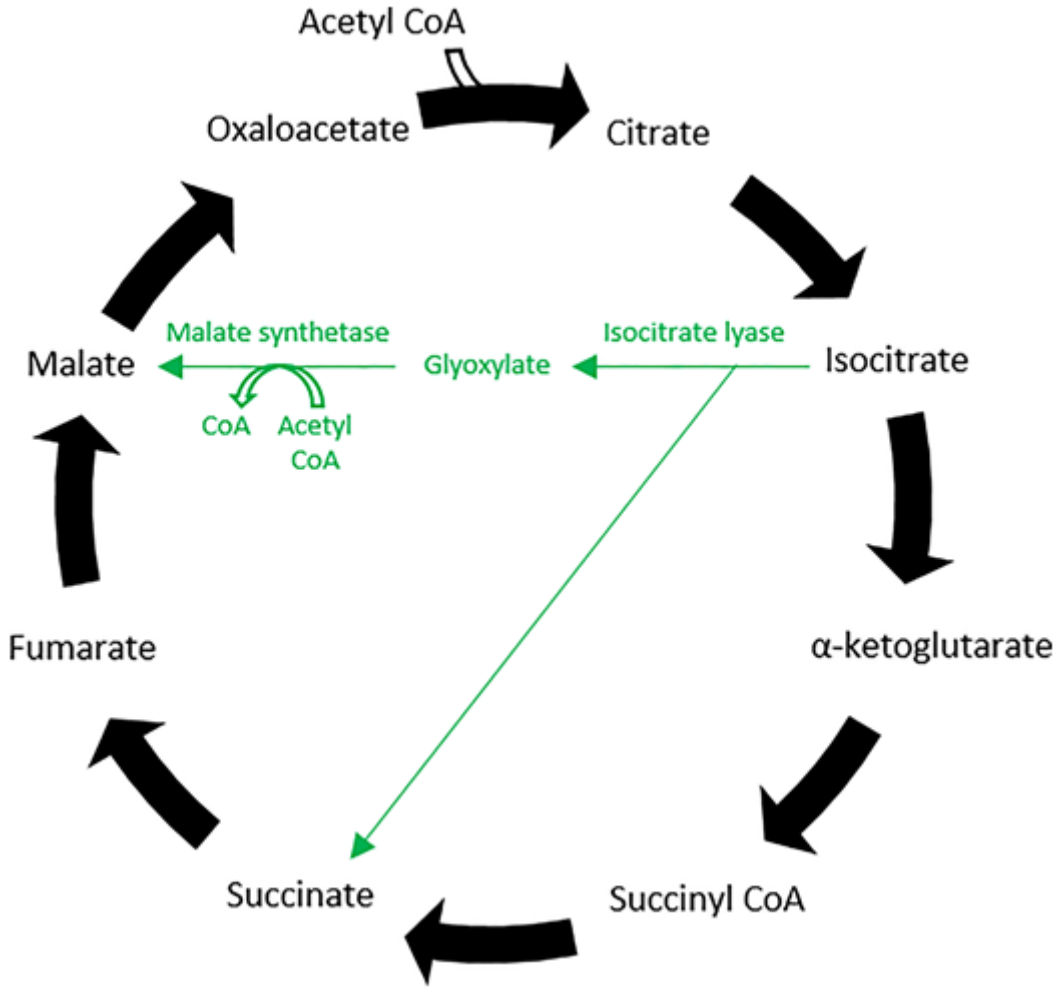


Figure 3: The glyoxylate shunt (green) coexisting amongst the TCA cycle (black). The glyoxylate shunt converts fatty acids into carbohydrates by bypassing decarboxylation steps of the TCA cycle. Taken from (Judge & Dodd, 2020).

It has been reported that *P.brassicae* spores and secondary plasmodia contains large amounts of trehalose and lipid droplets (Brodmann et al. 2002; Bi et al. 2016). *P. brassicae* lipid droplets are highly enriched in Triacylglycerol (TAGs), the most important compounds of lipids for energy storage in mammals (Bi et al. 2016). Fatty acids can be used to produce energy via ß-oxidation. Fatty acid β-oxidation is a cyclic process of breaking down a long-chain acyl-CoA molecule to acetyl-CoA molecules. During this process, the stepwise shortening of acyl-CoA produces one molecule of NADH and FADH2 for every two carbon units released and acetyl-CoA. The acetyl-CoA produced can then be used in the glyoxylate cycle or in the TCA cycle to further generate more energy thus the energy yield from fatty acid oxidation is as much as double than the energy produced from glucose (Houten *et al.*, 2016). Given that the germinating spores require high amounts of energy to power the revived cellular metabolic processes, swimming, and root hair infection, it is plausible to think that the pathogen will utilise the stored

TAGs to meet these energy demands. Bi *et al.*, (2016) isolated proteins associated with TAGs in *P.brassicae* and found that these proteins were enriched in enzymes involved in TAG biosynthesis and metabolism. *P.brassicae* can also break down trehalose to produce glucose which can be used for energy production via other TCA and glyoxylate pathways (Figure 3). Trehalase, which breaks down trehalose to glucose, was primarily expressed in the spores. Breakdown of trehalose to provide carbon and energy for the revived enzymes and initial development have been reported in the germinating spores of *Streptomyces griseus* (McBride and Ensign 1987) and *Mycobacterium smegmatis* (Shleeva et al. 2017).

Following root hair infection, the primary zoospore will develop into primary plasmodia (Figure1). The primary plasmodia also contain lipid droplets which can be metabolised for energy production. At this stage, the *P.brassicae* is in contact with the host so it may take up some host nutrients from the root hairs. The supply of nutrients in root hair is limited thus the primary plasmodia may meet its energy requirements primarily from the stored lipid droplets. Once inside the root cortical cells, *P.brassicae* can utilise host sugar nutrients. *P.brassicae* was found to express hexose sugar transporters and carbohydrate metabolism genes of the glycolysis cycle during the secondary plasmodial stage (Schwelm *et al.*, 2015; Kong *et al.*, 2022) indicating the pathogen dependence on host carbohydrates. Hexose sugars would be phosphorylated to hexose-P and then enter general metabolism. Glucokinase (EC 2.7.1.2) and fructokinase (EC 2.7.1.4) have been identified in the genome (Rolfe *et al.*, 2016). However, if sucrose is taken up, it have to be metabolised quickly to drive the sucrose uptake but potential sucrolytic enzymes have not yet been identified in *P.brassicae* (Rolfe *et al.*, 2016).

1.7. Intracellular pathogens nutrient acquisition mechanisms

Nutrients can be defined as indispensable elements required for growth by all living organisms. One common feature regarding nutrient metabolism in biotrophic pathogens is the loss of genes encoding enzymes involved in primary and secondary metabolism, lack of some transporter families which could imply redundance in transporter repertoire and expansion of some transporter families (Baxter *et al.*, 2010; Spanu *et al.*, 2010; Duplessis *et al.*, 2011). The host cell environment is critical for the survival of biotrophic pathogens. One of the factors affecting cellular environment is the pH gradient between different cellular compartments which ranges from 5.5 in the apoplast and vacuole (Bosch & Franklin-Tong, 2024) to 8.4 in the peroxisomes (Shen *et al.*, 2013). Changes in cellular pH possesses the potential to exert a substantial influence upon virtually all cellular processes, encompassing metabolic pathways, membrane potential, cellular proliferation, and the translocation of substances across the plasma

membranes (Putnam, 2012). Intracellular pathogens therefore must overcome any challenges posed by the host intracellular environment which could reduce access and transport of nutrients in the pathogens. As a consequence, intracellular pathogens deploy different nutrient acquisition strategies such as transport via membrane contact sites, endocytosis and expression of a wide array of plasma membrane embedded nutrient transporters to compete with their host cells for cytosol nutrients (Dumoux & Hayward, 2016; Garvetto *et al.*, 2023).

1.7.1. Transport through Membrane contact sites

Membrane contact sites (MCSs) are regions where two organelles come in proximity with a separating distance of no more than 30nm and without membrane fusion (Vormittag *et al.*, 2023). MCSs are essential for the interaction and communication between organelles within the cell. MCSs are involved in the non-vesicular trafficking of lipids, proteins and calcium ion between organelles (Dumoux & Hayward, 2016; Medeiros *et al.*, 2021).

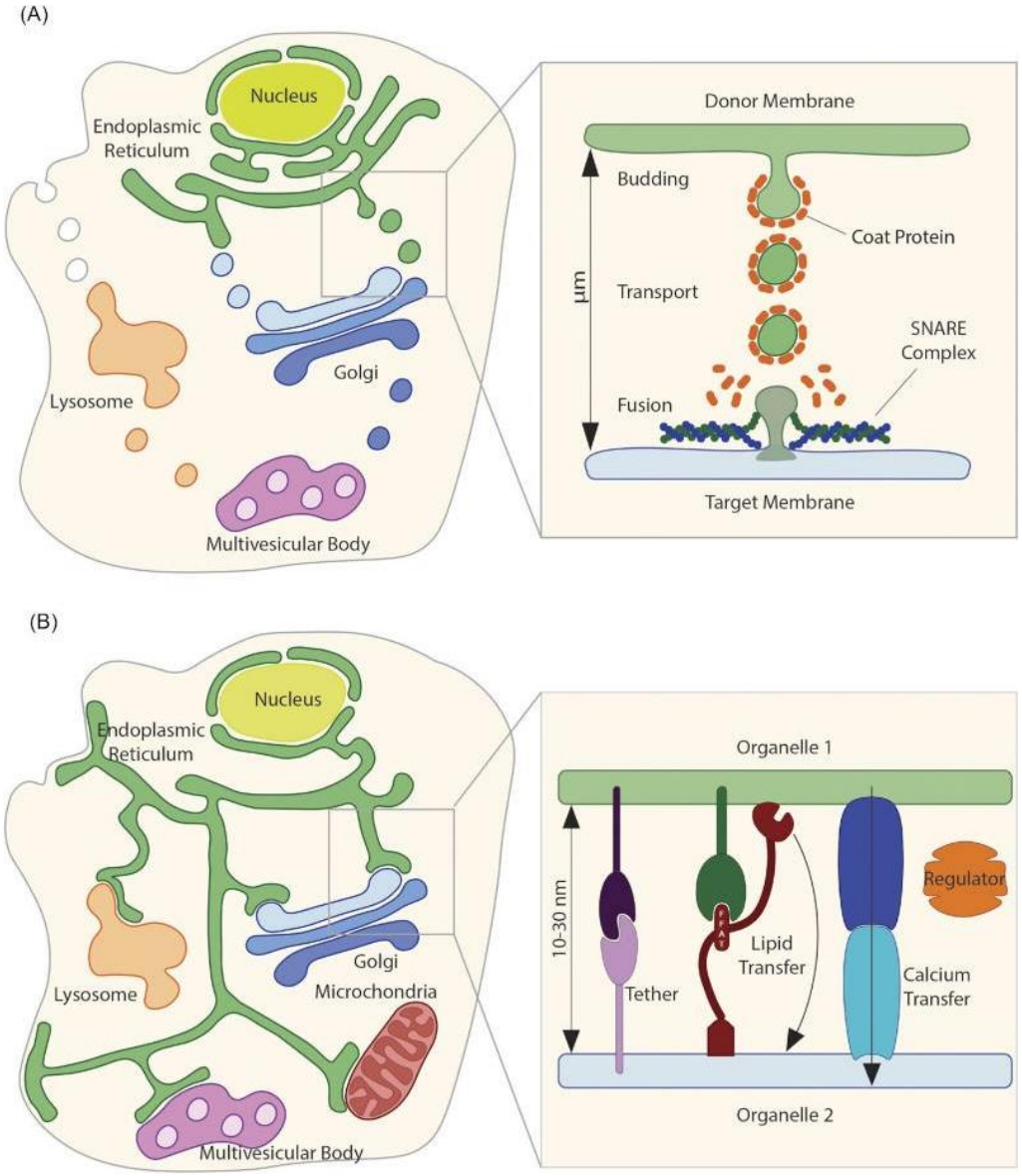


Figure 4: Schematic representation of transfer of metabolites at MCSs.
A) transient formation of MCS. The two interacting organelles transfer the lipids and other solutes from the doner organelles in vesicles which move and fuse with the target membrane in a SNARE-mediated process. B) static MCS sites. The membranes of interacting organelles are held in place with attachment proteins attached to their membranes. The transfer of solutes from the doner to the target membrane is facilitated by the binding (tethering) proteins. Taken from (Vormittag et al., 2023).

Nutrient acquisition through MCSs has been found in intracellular bacterial and protist pathogens of humans and animals (Dumoux & Hayward, 2016; Medeiros *et al.*, 2021; Santos & Nozaki, 2021). Such mode of nutrient acquisition has not so far been reported to exist in intracellular plant pathogens. *P.brassicae* resides in the host cytosol which makes it an additional organelle and thus it may also form these MCSs with host endoplasmic reticulum and mitochondria.

1.7.2. Nutrient acquisition by endocytosis

Endocytosis is a mechanism in which eukaryotic cells take up extracellular macromolecules into their cytoplasm (de Souza, 2024). The uptake of extracellular solid particles is phagocytosis whereas the uptake of extracellular fluids is pinocytosis (Garvetto *et al.*, 2023).

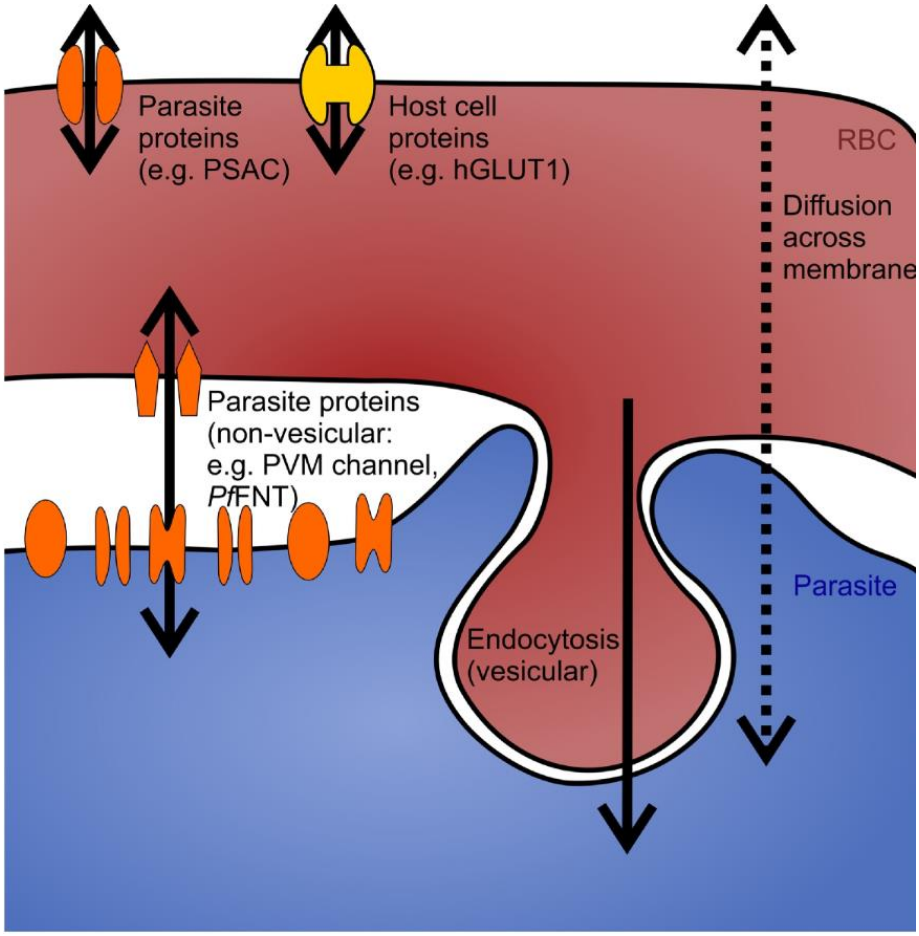


Figure 5: schematic representation of P.falciparum nutrient acquisition pathways. Nutrients are transported across the red blood cell (RBC)plasma membrane by a mix of parasite and host transporters. The nutrients move across the parasotophorous vacuole membrane (PVM) and parasite plasma membrane (PPM) through parasite established pathways such as endocytosis or non-vesicular. Nutrients may also diffuse across the plasma membranes. Taken from (Henshall & Spielmann, 2023).

Most eukaryotes including the parasitic protists depend on the endocytic pathway for the uptake of macromolecules which are broken down in specialised endosomal-lysosomal system to provide metabolic fuels for different cellular processes (de Souza, 2024). Phagocytosis is speculated to be the main mode of nutrition in free living Rhizarians (Cavalier-Smith *et al.*, 2018). Phagocytosis of host cytoplasm and organelles has been reported in other phytomyxea and have been recently found to exist in *P.brassicae* (Garvetto *et al.*, 2023).

1.7.3. Transport by membrane transporter proteins

Transporters are membrane embedded proteins that function in solute trafficking across the cell membrane and membranes of cellular organelles. Transporters that have a similar protein structure and function belong to the same family and the families are further grouped into super families.

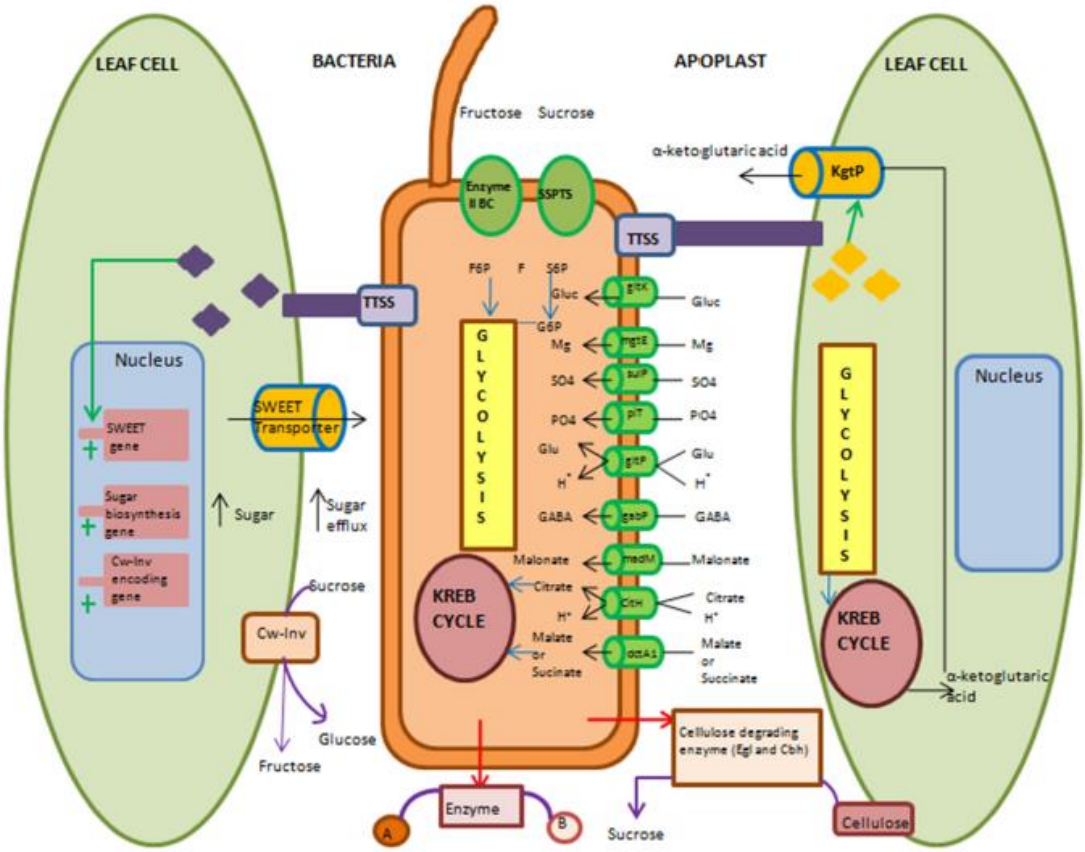


Figure 6: Overview of molecular and biochemical events used by biotrophic or hemi-biotrophic bacteria to acquire nutrients during apoplast colonisation.
Pathogens secrete effector proteins (i.e Type three secretion system(TTSS) in this figure) that target host sugar transporters such as SWEET (sugar will eventually be exported transporter) (yellow cylinders) and reprogram host sugar partitioning there by increasing sugar supply in the infected sites. Pathogens then express a wide array of enzymes and transporters (green cylinders and green circles) to break down and assimilate host derived nutrients. S6P, Sucrose-6-phosphate; F6P, Fructose-6-phosphate; G6P, glucose-6-phosphate; SSPTS, Sucrose-specific phosphotransferase system; gltK, Glucose ATP-binding cassette transporters (ABC) transporter permease; mgtE, Magnesium transporter; fecC, Iron ABC transporter permease; sulP, sulphatepermease; piT, Inorganic phosphate transporter; gltP, Proton/glutamate symport; gabP, gamma-amino butyric acid (GABA) permease; madM, malonate transporter; citH, Citrate/proton symport; dctA1, Dicarboxylate transporter; kgtP, Ketoglutaric acid transporter; egl, Endoglucanase; cbhA, Cellobiohydrolase. Taken from (Fatima & Senthil-Kumar, 2015).

Pathogens expresses different nutrient transporters during their interaction with host that facilitate the trafficking of host-derived nutrients across the pathogen plasma membrane (Garnica *et al.*, 2013; Fatima & Senthil-Kumar, 2015). *P.brassicae* has been reported to express sugar transporters, amino acid transporters and thiamine transporters during clubroot disease development (Rolfe *et al.*, 2016; Kong *et al.*, 2022). There are fundamentally two types of sugar transporters – permeases, where sugar flows down a concentration gradient and active transporters where a proton gradient is used to drive sugar uptake against a concentration gradient (Chen et al. 2010; Sauer et al. 1990) the former are SWEETs, the latter Sucrose transporters (SUCs/SUTs). The number of nutrient transporters encoded by *P.brassicae* genome, and their mode of transport is not known. This information is critical to understand how the pathogens obtain nutrients from the host.

1.7.3.1. Types of membrane transporters

Transporters facilitate the movement of water and solutes such as inorganic ions, sugar, amino acids, and other molecules to travel through cell membranes. Transporters are grouped into three classes depending on their mode of transport. These are channels, carriers and pumps.

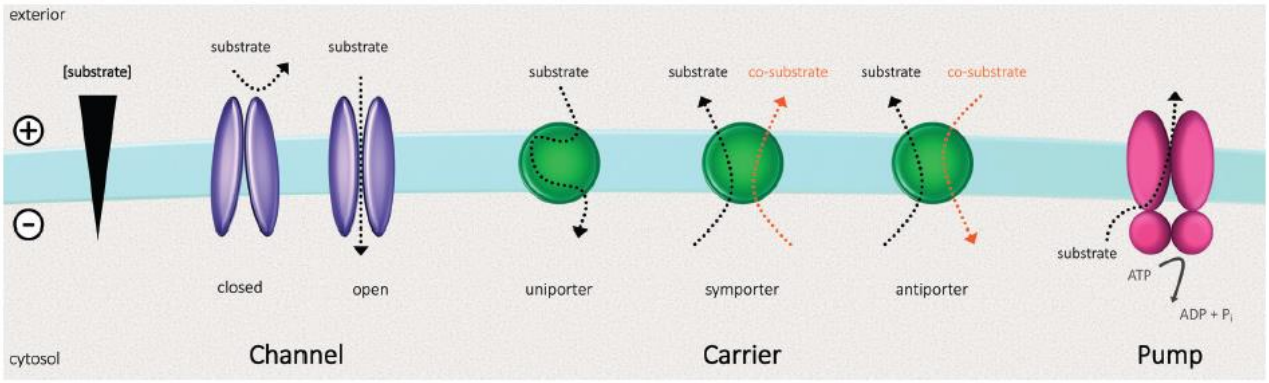


Figure 7: Schematic diagram of modes of substrate transport across the plasma membrane. The diagram shows a plasma membrane in (blue) separating the cell cytosol and its environment (exterior). The plus and minus signs indicate a membrane potential, and the black triangle indicates an inward electrochemical gradient. Channels (purple) move substrate down its electrochemical gradient. Channels open or close in response to stimuli such as changes in membrane voltage. Carriers (green) bind substrates and they undergo several conformational changes to deliver the substrate to the opposite side of the membrane. Pumps (pink) use energy from primary source (e.g ATP) to move substrates against their electrochemical gradient. Taken from (Martin, 2020).

Channels (Figure 7, purple) are made of large proteins which transverse the plasma membrane forming a pore through which ions or molecules of a specific size and charge cross the plasma membrane down their electrochemical gradient (Martin, 2020; Picci *et al.*, 2022). Channels are categorised based on the ion selectivity and the type of gating mechanism and they are involved

in a wide range of physiological and pathological processes (Picci *et al.*, 2022). Carriers bind and transport substrates across the membrane by undergoing reversible conformational changes. A carrier can be uniporter transferring a substrate down its electrochemical gradient, symporter which couples the transport of substrate and co-substrate in the same direction or antiporter which transports substrate and co-substrate in different directions. Symporters and antiporters use energy stored in the electrochemical gradient of co-substrates (Figure 7 ,orange) to move the substrate against its concentration gradient (Stillwell, 2016; Martin, 2020). Pumps (Figure 7,pink) use energy from the primary source (ATP) to pump a solute against its electrochemical gradient.

Transporters can also be classified based on whether the transport type involves the use of energy. This groups transporters as passive and active transporters. Passive transport can be either by simple diffusion where the substrate dissolves into and diffuse the plasma membrane or facilitated diffusion where the substrate movement down its concentration gradient is aided by a carrier to cross the plasma membrane. Active transport is divided into primary and secondary active transporters. Primary active transporters use the primary source of energy (e.g ATP) while secondary active transporters use energy stored in electrochemical gradient of ions to transport substrates against their concentration gradients (Stillwell, 2016; Martin, 2020).

Due to a lack of uniform, well-defined communication concepts of transmembrane transporters that are universally understood by international scientific community, a new transporter classification system "the Transporter classification (TC) system was created (Saier, 2000). TC-system assigns transmembrane transporters into TC-class, subclass, superfamilies, families and subfamilies based on function, sequence homology and substrate specificity. TC-system is regarded as the best way to characterise membrane transporters and has been approved and recommended by the nomenclature committee of the International Union of Biochemistry and Molecular Biology (IUBMB)(Saier, 2000).

1.8. Transporter identification methods

1.8.1. Function-driven screens

This method is based on the cloning and expression of putative transporters in heterologous expression hosts. Transport activity is attributed to the putative transporter if it can functionally complement the mutation in the heterologous expression host (Larsen *et al.*, 2017). The advantage of this approach is that both the potential substrate and the function of the putative transporter is experimentally obtained. The downside though is it requires knowledge about the

putative substrate, availability of suitable heterologous expression hosts, biological sensors and it is a low throughput and laborious. Nevertheless, functional-driven screens have been instrumental in identifying membrane transporters in plants, animals, and plant pathogens. Sucrose transporters that facilitate sucrose import into the sieve elements and companion cells of the phloem (Sauer & Stolz, 1994) and the SWEET transporters that facilitate sugar movement down the concentration gradient (Chen *et al.*, 2012) were identified by function-driven screens. A similar approach was employed by (Gu *et al.*, 2000) to identify the mouse histidine and glutamine amino acid transporter mNAT. By heterologous expressing cDNA isolated from mouse kidney into *Xenopus laevis* oocytes, they observed that these oocytes accumulated significantly higher amounts of radioactive labelled amino acids compared to control oocytes. Heterologous expression system has also been successfully utilised to characterise the hexose sugar transporter of broad bean pathogen *Uromyces fabae* (Voegele *et al.*, 2001). This method can be harnessed to identify and characterise *P.brassicae* nutrient transporters.

1.8.2. Phenotype-driven screens

Phenotype-driven transporter identification is achieved by feeding toxic analogues of substrates or mutating putative target transporters. It requires prior knowledge about the physiological role of the transporter. These assays are conducted in the native transporter environment or hosts (Larsen *et al.*, 2017). Limitations to this method are prior knowledge about the physiological role of the transporter, knowledge of and availability of toxic analogues, ability to physically separate the sink and source tissues, availability of fluorescently or radioactive labelled substrates where physical separation of tissues is not possible, and works on the assumption that labelled substrates will be recognised and transported by the transporters (Larsen *et al.*, 2017). Phenotype driven transporter identification led to the identification of a first sucrose transporter of plant pathogen *Ustilago maydis* (Wahl et al., 2010) . Wahl et al created individual hexose transporter or hexose transporter-like mutants and assayed their development by inoculating them on young corn seedling. They observed that two of the mutations influenced the virulence of the pathogen. One of two mutation was characterised by heterologous expression in yeast and was found to encode a high affinity sucrose transporter. This approach has been used to identify nutrient transporters of human pathogenic protists (Martin, 2020) but is practically difficult to use for transporter identification in *P.brassicae* . The primary obstacle lies in the extreme difficulty of genetically transforming *P.brassicae*. To date, only a single publication has reported successful genetic transformation of this pathogen (Feng *et al.*, 2013a).

This limited success renders targeted mutagenesis of putative transporters an insurmountable challenge for investigating nutrient transport in *P.brassicae*.

1.8.3. In silico-based approaches

Computational methods are computer-based approaches used to identify putative transporters based on guilt by association concept. Expression based, gene clustering and homology-based approaches are used for identification of candidate genes (Larsen *et al.*, 2017). This approach leverages high-throughput-omics technologies that generates big biological data (David *et al.*, 2019). Homology-based approach is the most widely used in silico method for identification of candidate transporters due to the availability of data in public databases such as NCBI (https://www.ncbi.nlm.nih.gov/), InterProScan (https://www.ebi.ac.uk/interpro/) and MYCOCOSM (https://mycocosm.jgi.doe.gov/mycocosm/home). There are also databases that are specific for transporter proteins such as ligand-gated ion channel database (Donizelli *et al.*, 2006), human transporter database (Ye *et al.*, 2014), the transporter classification database, TCDB (Saier *et al.*, 2021) and transporter annotation pipelines (Li *et al.*, 2009; Elbourne, LD *et al.*, 2017). Homology-based approaches use information extracted from experimentally characterised transporters that is computationally inferred to unknown transporters based on primary sequence identity, predicted secondary structural homology and topology (Saier, 2000; Elbourne, LD *et al.*, 2017). It allows the identification of membrane transporters relatively quickly compared to other membrane transporter proteomic approaches. It also allows the identification of transporters in non-model species whose mutant collections or full-length cDNA libraries are not available (Larsen *et al.*, 2017). However, transporters identified using homology-based approaches requires experimental confirmation due to challenges with annotating genes and proteins and mismatch between genomic and functional data on protein which may lead to misleading information about the role of the putative transporter (David *et al.*, 2019). In silico based approach has been used to identify and characterise hexose sugar transporter PsHTX1 of wheat stripe rust (Chang *et al.*, 2020), sugar transporters of Rice False Smut Pathogen, *Villosiclava virens* (Qin *et al.*, 2024) and *P.brassicae* (Kong *et al.*, 2022).

1.9. Problem statement and justification of the study

Despite being a major pathogen of cruciferous plants, many aspects of *P. brassicae*-host interactions remain obscure (Feng et al. 2012). There is much to learn about the nutrition of this important biotrophic pathogen in all stages of its lifecycle. Although there have been efforts to examine how *P.brassicae* infection interferes with the host carbohydrate distribution (Keen and Williams 1969; Evans and Scholes 1995; Brodmann et al. 2002; Walerowski et al. 2018; Li et al. 2018), little is known about the chemical nature of nutrient compounds and the pathogen transporters involved in their uptake from its hosts. Lack of information on *P.brassicae* nutrient transporters can be attributed to the lack of *P.brassicae* transformation systems, inability to grow in axenic cultures and its intracellular lifestyle. This is compounded by challenges associated with membrane transporter proteomics such as low protein abundance of transporter proteins, their hydrophobic nature, solubility and purification challenges and their requirement for a membrane system for transport assays (Quick & Javitch, 2007; Hediger *et al.*, 2013; Larsen *et al.*, 2017). In silico homology-based approaches offers an opportunity to identify *P.brassicae* transporters by leveraging the *P.brassicae* -host interaction transcriptome and genomic data deposited in public databases. Nonetheless, application of computational methods to study *P.brassicae* biology is still a challenge due to the lack of well annotated *P.brassicae* refence genome.

The role of cytokinin in clubroot disease development remains obscure. Early studies suggested that *P.brassicae* may produce cytokinin during disease development (Dekhuijzen and Overeem 1971; Müller and Hilgenberg 1986) and these observations have been supported by the identification of cytokinin biosynthesis genes in *P.brassicae* genome (Rolfe *et al.*, 2016). Different types and levels of cytokinins in *P.brassicae* infected plants have been reported (Muller & Hilgenberg, 1986; Devos *et al.*, 2006; Malinowski *et al.*, 2016). It is therefore not known whether the *P.brassicae* IPT genes are either ADP/ATP isopentenyl transferases (ADP/ATP IPTs) and synthesise isopentenyl adenosine (iPA) and *trans*-zeatin cytokinin or tRNA IPTs and synthesise cis-zeatin. Experimental characterisation of *P.brassicae* IPT genes is important because the activity of different cytokinin varies among plant species (Yonekura-Sakakibara *et al.*, 2004; Sakakibara, H., 2006). Pertry et al. (2009) found that formation of differentiated leafy galls in Arabidopsis infected by *Rhodococcus fascians* resulted from continuous stimulation of tissue proliferation caused by local increase in cZ and a synergistic activity of cZ and other CKs secreted by the pathogen. Infection of maize by dikaryon *U. maydis* leads to the decrease of total CKs due to a reduction in CK glucosides that are the predominant

type of CKs in healthy maize plant but results in the increases in specific freebase, riboside and nucleotide (FBRNT) CKs which in turn increases its virulence (Morrison et al. 2015). Despite availability of a large body of experimental work on the role of cytokinin in *P.brassicae* disease development, the type of cytokinins produced by *P.brassicae* and their role in pathogen development or their influence on host development and/or defence responses has never been demonstrated. Knowing when the Pb-IPT genes are expressed and the type cytokinin they make would advance our knowledge in deciphering the role of cytokinin in *P.brassicae* development. Understanding of and elucidation of *P.brassicae* nutrient acquisition mechanisms and *P.brassicae* cytokinin metabolism could contribute to development of novel strategies in resistance breeding or clubroot management.

1.9.1. Overall objectives

The main objectives of this study project were to contribute to the understanding of the biology of this important biotrophic pathogen by identifying all transporters present in the genome and their role in nutrient acquisition from the host, to improve the *P.brassicae* gene models and characterisation of *P.brassicae* cytokinin biosynthesis genes (PbIPTs).

1.9.2. Specific objectives

1. To correct *P. brassicae* models for genes of interest
2. To create a *P.brassicae* gene model database that allows public collaboration for manual curation of existing *P.brassicae* gene models
3. To create a *P.brassicae* transporter gene catalogue and assign them to transporter families using TC-system nomenclature.
4. To identify *P.brassicae* nutrient transporters, their substrates and their families.
5. To identify putative sugar transporters, determine their substrate specificity, subcellular localisation and expression at different stages of the pathogen lifecycle
6. To characterise cytokinin biosynthesis genes in *P. brassicae*, determining which forms of CKs the pathogen can produce and their role in controlling host and pathogen development, and host susceptibility and defence.

1.9.3. Hypotheses

1. *P. brassicae* expresses different classes of nutrient transporters at different stages of the pathogen lifecycle.

2. *P. brassicae* sugar transporters mediate hexose or sucrose import and are localized at the *P. brassicae* plasma membrane for carbon allocation of the pathogen.
3. PbIPTs are either ADP/ATP isopentenyl transferases (ADP/ATP IPTs) and synthesise isopentenyl adenosine (iPA) and *trans*-zeatin cytokinin or tRNA-IPTs and synthesise *cis* zeatin.
4. *P.brassicae* produces CK which influences host development and/or defence responses

CHAPTER 2: MANUAL CURATION OF PLASMODIOPHORA BRASSICAE GENE MODELS

Abstract

The basic biology of *P.brassicae* remains elusive due to challenges such as the lack of transformation systems, the inability to grow in culture, and its intracellular lifestyle. In silico, homology-based approaches offer a valuable opportunity to explore *P.brassicae* biology by leveraging the *P.brassicae* -host interaction transcriptome and genomic data available in public databases. However, applying computational methods to study *P.brassicae* biology remains difficult, primarily due to the absence of a well-annotated *P.brassicae* reference genome. Since the initial publications of *P.brassicae* genomes, 52 *P.brassicae* genomes have been released, with the most recent featuring a complete telomere-to-telomere assembly. Genome annotation programs often rely on the same tools and algorithms, and these inherent limitations, combined with the unique nature of the *P.brassicae* genome, lead to flawed gene models and similar errors across different annotations. This chapter focuses on the experimental validation of *P.brassicae* gene models using transcriptome data from short-read RNA and single-molecule, full-length/long-read RNA sequencing technologies. Using bioinformatics tools, I compare the accuracy of gene models and identify errors between PbPT3 and Pb314v2 *P.brassicae* genome annotations. The analysis reveals that a significant number of gene models from both annotations are incorrect. I demonstrate that full-length/long-read mRNA sequencing is crucial for resolving gene model errors, particularly at the start and end positions of genes. The findings of this study were used to validate the presence of unusual *P.brassicae* transporters, as discussed in Chapter 3, and to design correct gene constructs for functional characterisation of *P.brassicae* isopentenyl transferase genes in chapter 4. Additionally, the results inspired the creation of a community resource aimed at facilitating the manual curation of *P.brassicae* gene models and the development of a well-annotated *P.brassicae* reference genome. This resource will, in future, provide benefit to the wider clubroot research community.

2.1. Introduction

The basic biology and mechanistic basis of parasitism in *P.brassicae* is not well understood owing to its intracellular lifestyle, lack of transformation systems and a strict requirement of a living host. Whilst genomic information allows researchers to study the biology of these experimentally challenging important pathogens (Rolfe *et al.*, 2016; Stjelja *et al.*, 2019; Bryant *et al.*, 2023), accurate annotation is important to facilitate research efforts. Gene structure annotations are usually predicted from sequenced genomes by computational methods. These gene models are then used to predict and assign biological function to the predicted genes. Although gene-prediction computer programs are powerful and continuously improving, they are still prone to errors (McDonnell *et al.*, 2018), particularly in organisms that are in poorly represented taxonomic groups. Errors in gene model predictions therefore results in incomplete or misleading information about the putative roles of proteins (David *et al.*, 2019). Computational gene model errors include missing of gene models, false gene prediction and merged or split gene models (Williams *et al.*, 2011; McDonnell *et al.*, 2018).

The advances in next generation sequencing and the fall in sequencing costs has allowed clubroot researchers to sequence genomes of strains of interest. Two early *P.brassicae* sequencing efforts (Schwelm et al. 2015; Rolfe et al. 2016) described the *P.brassicae* genome and defined initial gene models. Schwelm et al. (2015) assembled 24.1 Mb genomic DNA into 165 scaffolds with 9,437 genes. Rolfe et al. (2016) also sequenced and assembled the *P.brassicae* genome size of 24.2Mb into 107 scaffolds and 10,851 genes. In a later sequencing project, Stjelja et al. (2019) assembled 25.1 Mb (includes the mitochondrial genome) into 20 nuclear contigs probably equivalent to 20 chromosomes and 9,231 genes (nuclear genome). Since these original publications, 52 *P.brassicae* genomes have been published with the most recent finally complete genome with telomere-to-telomere assembly (Javed *et al.*, 2024). The new genome is 25.3Mb assembly comprising of 20 chromosomes and 10,521 protein-coding gene models.

The current *P.brassicae* gene models have problems which make the use of *P.brassicae* genomic information challenging. For instance, Rolfe et al. (2016) in their annotation identified unusual nutrient transporters that consists of a prokaryotic substrate binding domain fused to a 7 alpha-helical transmembrane domain. The sequences of these unusual transporters are present in the *P.brassicae* e3 genome (Schwelm et al. 2015) and either annotated as two different genes or not annotated at all. A well curated and annotated *P.brassicae* genome is needed given that clubroot researchers are using *P.brassicae* genomic information to study the role of specific

P.brassicae proteins during *P.brassicae* -host interactions (Ludwig-Müller *et al.*, 2015; Kong *et al.*, 2022; Feng *et al.*, 2024).

Accurate and reliable reference genomes are created by manual curation of gene models because the best gene prediction computer programs can only predict approximately 80% of complete gene structures correctly (Williams *et al.*, 2011; McDonnell *et al.*, 2018). In *C.elegans,* it was reported that computer gene prediction programs struggle to predict correctly gene structures for genes with poorly conserved orthologs, long introns, many exons, short exons, or weak translation initiation signals (Williams *et al.*, 2011). However, manual curation uses transcript data as a direct source of evidence but also uses indirect sources of evidence such as protein alignments, homology to orthologs and paralogs and conserved protein domains which are useful to annotate genes with very few or no transcripts (Williams et al., 2011; McDonnell et al., 2018).

P.brassicae genome contains two isopentenyl transferase (PbIPT) genes. The gene model structure of these genes is not the same in the different *P.brassicae* genome annotations which makes it difficult to know the correct gene to use in functional studies. The objective of this study was to resolve gene model structures of *P.brassicae* IPT genes by using experimental evidence generated by full length/ long read mRNA sequencing technologies before their functional characterisation. The data produced in full length/mRNA sequencing was also used to inspect the quality and accuracy of other *P.brassicae* gene models. To this end, a web-based community-driven annotation resource (Lee *et al.*, 2013) that is hosted by the University of Sheffield has been created. This resource would facilitate manual curation of *P.brassicae* gene models by incorporating both full length/long reads mRNA and short reads RNAseq studies multiple research groups thus addressing the inherent limitations of automated gene prediction methods, which coupled with the compact nature of *P.brassicae* genome with short intergenic regions and overlapping genes may struggle to correctly predict gene structures.

2.2. Materials and methods

2.2.1. Biological materials.

Arabidopsis thaliana Col-0 plants were used for all full-length single-molecule mRNA sequencing experiments, while Chinese cabbage (Wong Bok) plants were used to maintain the *P.brassicae* inoculum and for single-cell sequencing experiments. For Arabidopsis, 5 cm diameter pots were filled with M3 compost, and seeds were stratified by placing them in water and refrigerating at 4°C for 3-5 days before being sown directly into the pots using a pipette. For Chinese cabbage, 10 cm x 10 cm x 20 cm pots were filled with M3 compost and seeds were sown directly. All pots were then placed in a growth room with an irradiance of 100 µmol m^{-2} sec^{-1}, a photoperiod of 10 hours light and 14 hours dark, and day/night temperatures of 20°C/18°C. After germination, seedlings were thinned to one plant per pot.

A *P. brassicae* isolate defined as 16/2/12 by European Clubroot Differential (ECD) set (Devos *et al.*, 2005) was used in this study. *P.brassicae* spore preparation was performed following (Mithen & Magrath, 1992). Galls of infected Chinese Cabbage Wong Bok were homogenized with dH_2O, filtered through 3 layers of muslin, and centrifuged (3 min, 4000 x g, 4°C). Soil, starch and plant tissue were removed from the pellet with a spatula, and the remaining spores re-suspended in dH_2O. The spores are in the tan layer at the top of the pellet. The centrifugation step was repeated to remove additional unwanted soil, starch and plant tissue. The inoculum was stored at 4°C until required with a maximum storage time of 4 weeks. For each experiment a new inoculum was prepared. For the determination of spore concentration, a staining solution (Donald *et al.*, 2002) containing 0.5 µL calcofluor white (100 µg/mL) was mixed with 10 µL of a 100-fold diluted spore stock (10 µL spore stock + 990 µL dH_2O), and 10 µL loaded onto a haemocytometer. Spore density was determined via fluorescence microscopy (BX51 Olympus) using excitation at 360 nm and detecting emission at 430 nm. Wall layers of active spores showed intense blue fluorescence as calcofluor white binds chitin. The plants were inoculated 14 days after germination with 200 µL of 1 x 10^7 mL^{-1} *P.brassicae* spores suspended in water.

2.2.2. Mapping gene models from different annotations to reference genome

Gene models from three early *P.brassicae* genome annotations PbPT3 (Rolfe et al., 2016), Pbe3_h15 (Schwelm et al., 2015) and Pb314v2 (Stjelja et al., 2019) were used to compare gene models produced by different research groups. The Pb314v2 genome (accession: GCA_900303365.2) (Stjelja et al., 2019) was selected as a reference genome for analyses as it was the most complete, and scaffolds most likely match complete chromosomes. Note however,

that a new *P.brassicae* genome with telomere-to-telomere assembly (Javed *et al.*, 2024) which is believed to be the most complete genome assembly was published after the analyses had been done. Gene models from Pbe3_h15 and PbPT3 genome annotations were converted to the same gff format of Pb314v2 gene models using AGAT (https://github.com/NBISweden/AGAT). The resulting gff files were aligned to the reference genome using a lift over program flo (Pracana *et al.*, 2017).

Long read RNA sequencing data (described in section 2.2.3.2) and 210 RNAseq datasets from 14 publicly available *P.brassicae* RNAseq studies (Table1) were mapped to the reference genome using MINIMAP2 (Li, 2018) and STAR (Dobin *et al.*, 2013) respectively.

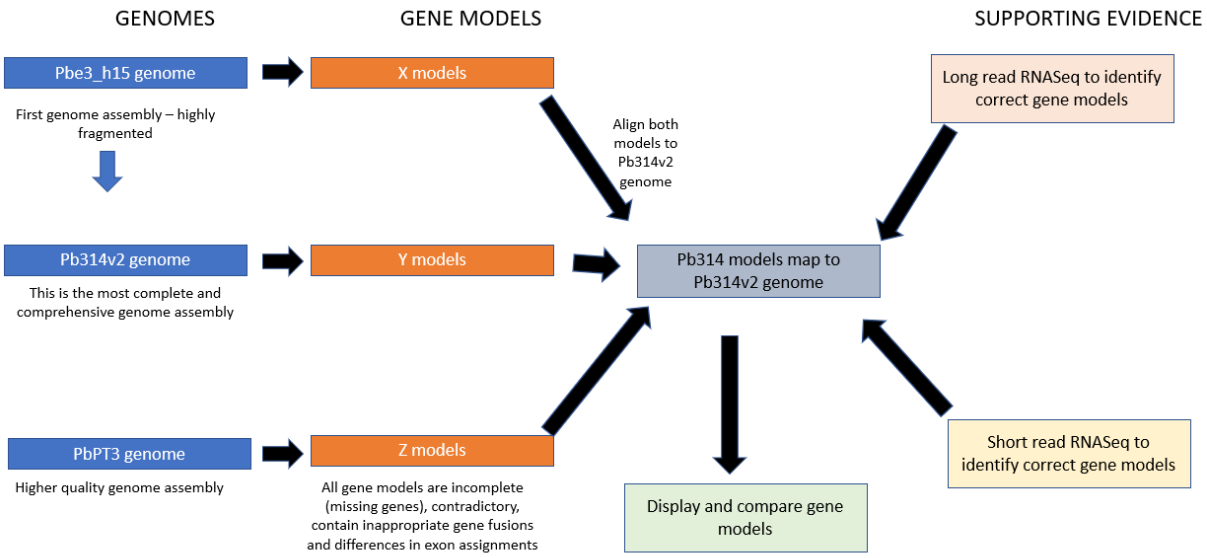


Figure 8: Flowchart showing steps used to assess the accuracy of P.brassicae gene models. Visualisation and comparison of gene models were done using Integrated genome viewer and custom R code. Gene model comparisons

Using the R packages Gviz and rtracklayer, the reference *P.brassicae* genome was read as a fasta file. The re-mapped gene models (as gtf files) were read into R and a database was created and saved in SQL format. This was done in turn for each gene model annotation. To compare gene models from different genome annotations, custom R code was then employed to generate PNG files visualizing *P.brassicae* IPT gene models and gene models of selected genes. These visualizations displayed intron/exon boundaries, overlaps with genes from other annotations, and supporting evidence from long-read sequencing and RNAseq studies.

Whole genome gene model comparisons were done using AGAT and details of gene models errors were revealed using gffcompare program (Pertea & Pertea, 2020).

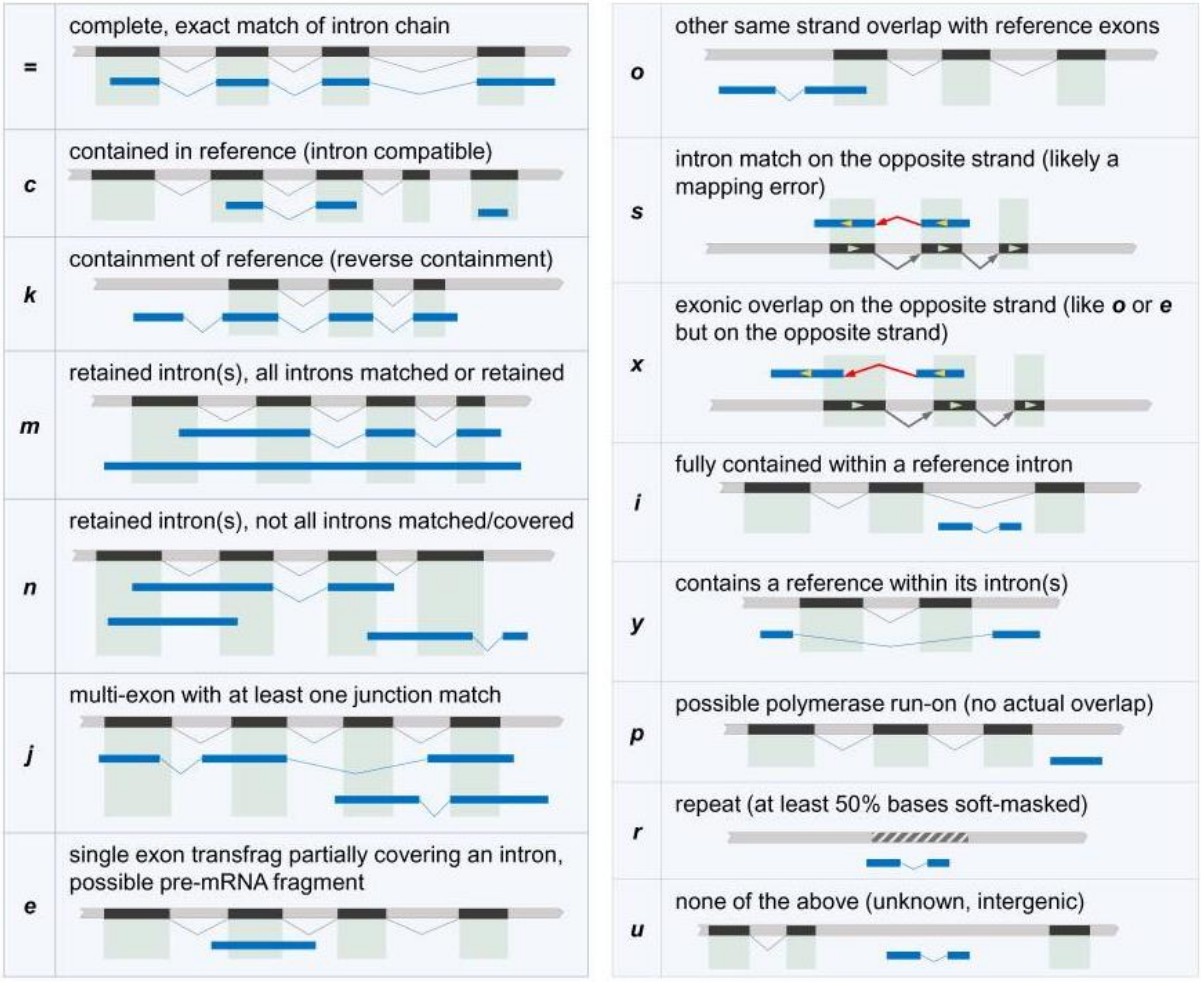


Figure 9: Common gene model errors.
Description of gene model error categories that are found in genome annotation comparisons using gffcompare program. Reference gene models are shown in black, gene models to be classified are shown in blue, and hashed regions represent repeated regions in the genome (Pertea & Pertea, 2020)

To further investigate errors in gene models, gene models associated with error codes in figure 9 were inspected using the (IGV), as described in McDonnell et al., (2018). Briefly, the Pb314v2 genome was loaded into IGV along with gff files from the three genome annotations and bam files from the long-read sequencing and RNAseq datasets. The gff and bam files were displayed in parallel tracks below the reference genome, enabling comprehensive visualisation of all information pertaining to a specific genomic region. PNG files were then saved from IGV for genomic regions exhibiting gene model errors.

2.2.3. Gene model supporting evidence

In-house long read sequencing data and RNAseq datasets from publicly available *P.brassicae* RNAseq studies were mapped to the reference genome and the resulting bam files were used as transcription evidence for gene models.

2.2.3.1. Publicly available RNAseq datasets

Short-read sequences from 210 samples across 14 publicly available RNAseq projects (Table 1) were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/). Control samples were removed before the analysis. Reads from infected samples were analysed for the presence of sequencing primers. The primers were trimmed from the reads by using trimmomatic (Bolger *et al.*, 2014) removing 15 bases from the read ends. The reads that had a minimum quality score of 15 were mapped to the *P.brassicae* genome using the STAR aligner (Dobin et al., 2013) with default settings except for intron size which was set to max 5Kb. Because the samples are normally a mixture of plant and pathogen transcriptomes, low *P.brassicae* reads were expected especially at the early time points.

Table 1: Bioprojects that contained RNAseq datasets used in *P.brassicae* gene model analysis.

SRA Accession	Host	Pathotype	Time points (dpi)	Input reads after trimming	Unique mapped	% of reads mapped to *P.brassicae* genome
PRJNA564005	Brassica napus	pathotype_4	0.5, 1, 2.5,4	859,309,635	6,758	< 0.01
PRJEB26435	Brassica oleracea	unknown	field	162,192,525	19,059,543	11.75
PRJEB36458	Brassica napus	eH	28, 36, 48	1,162,083,990	270,615,025	23.29
PRJNA322393	Brassica rapa	Pathotype_4	30	39,509,482	3,546,811	8.98
PRJNA298858	Brassica rapa	Pathotype_4	0, 0.5, 3, 4	159,530,688	1,019	< 0.01
PRJNA309985	Arabidopsis	Pb3	17	130,037,915	1,401,712	1.08
PRJNA453960	Brassica oleracea	unknown	7, 28	281,934,396	10,707,311	3.8
PRJNA528807	Brassica campestris	pathotype_race_7	42	144,835,123	52,771,218	36.44
PRJNA345072	Brassica oleracea	Pathotype_4	0, 7,14	335,848,328	18,182,510	5.41
PRJNA398451	Brassica oleracea	ECD16/4/0	3, 7, 20	275,746,680	14,776	0.01
PRJNA597078	Brassica napus	pathotype_5X	7, 14, 21	666,583,614	5,057,518	0.76
PRJNA641167	Brassica napus	F3_14_pathotype_3A	7, 14, 21	684,809,583	25,384,276	3.71
PRJEB12261	Arabidopsis	16_02_2012	16, 26	313,626,061	39,394,090	24.70
PRJNA331021	Brassica napus	ZJ-1	R.spore, germ.spores, cortical infection	365,627,627	278,299,119	76.12

2.2.3.2. Full-length single molecule mRNA sequencing

Total RNA was isolated from Arabidopsis roots and hypocotyls 26 days post inoculation with *P.brassicae* using TRIzol reagent (Sigma-Aldrick, UK) following manufacturer's instructions with some modifications. 200 µL (24:1) of chloroform isoamyl-alcohol were added to the aqueous phase and incubated on ice for three minutes. The mixture was then centrifuged at 12,000 x g for 20 minutes at 4°C. The RNA-containing aqueous phase was transferred to a new tube, mixed with 1/10 volume of 3M sodium acetate (pH 5.3) and 2.5x volumes of isopropanol, and precipitated at -20°C for at least 30 minutes. The RNA was resuspended in 70 µL of RNAse

free water and the tube was incubated at 65°C for five minutes to improve RNA solubility. The RNA was aliquoted by putting 10 µL of RNA into six 0.2 mL RNAse-free PCR tubes. RNA was quantified using Nano-Drop and Qubit equipment (Thermo Fisher Scientific). RNA integrity was checked by running the 5 µL of RNA on 1% agarose gel for 100 min at constant 90 V. The total RNA was then sent for full-length single-molecule mRNA sequencing at NEOF-NERC environmental omics facility, University of Sheffield. Sequencing libraries for PacBio sequencing technology were successfully prepared but sequencing failed for unknown reasons. Freshly prepared RNA samples were then sent to the facility for sequencing using Oxford Nanopore Technology (ONT) sequencing technology. The ONT sequencing worked but the sequencing was repeated two times because the quality and quantity of full length/ long reads obtained at each run was relatively low. The reads were mapped to *P.brassicae* reference genome using MINIMAP2 aligner (Li, 2018).

2.3. Results

2.3.1. Evaluation of initial *P.brassicae* gene models

Since the first publication of *P.brassicae* draft genome (Schwelm et al., 2015), 51 more *P.brassicae* genomes have been published, although gene models have not been created for some. Where gene models have been created, the number of genes annotated in each annotation is significantly different. To find the causes of variations in predicted number of genes in different *P.brassicae* genome annotations, gene models from the *P.brassicae* genome annotations PbPT3 (Rolfe et al., 2016), and Pb314v2 (Stjelja et al., 2019) were compared.

Table 2: Comparison of two *P.brassicae* genome annotations.

PbPT3	Pb314v2	Number of cases
0	0	17
0	1	824
1	0	2110
1	1	7186
1	2	242
1	3	18
1	4	1
1	6	1
1	11	1
2	0	34
2	1	501
2	2	23
2	3	3
3	0	8
3	1	51
3	2	2
4	1	11
4	2	3
5	1	3
5	2	1
7	1	1
9	2	1
9	4	1
26	7	1

Zero means no gene was annotated in that genomic region. One means a gene was annotated in the same genomic region. Zero to many means multiple genes were annotated in the same

region in one genome annotation but not in the other and the reverse is true. One to many and many to many means on gene or multiple genes overlap

PbPT3 genome has 10851 genes and Pb314v2 has 9231 genes predicted. 7186 genes or genomic regions in PbPT3 had overlaps in the same genomic region of Pb314v2 and 17 genomic regions did not have a predicted gene in both genome annotations. PbPT3 genome annotations had 2110 genes or genomic regions predicted to encode a gene that were not present in the refence genome. Additionally, 42 genomic regions, each predicted to contain two or three genes, lacked corresponding gene predictions in the reference genome. To get more detailed information on the problems present in gene models, Pb314v2 genome annotation was compared to PbPT3 genome as a reference. The analysis showed that 8182 loci in Pb314v2 genome annotation had overlapping exons in the reference while 632 were novel loci. Of the 8814 super union loci (combined loci of query and reference annotations), 2793 had matching intron chains, 4015 had matching transcripts and 4001 matching loci.

Table 3: Summary of results of Pb314v2 genome annotation compared to PbPT3 genome annotation

Accuracy estimation	Sensitivity (%)	Precision (%)
Base level	94.5	85.9
Exon level	73.5	59.4
Intron level	86.1	70.9
Intron chain level	39	18.2
Transcript level	44.7	18.9
Locus level	45.1	43.5

Missed exons: 950/41565 (2.3%)

Novel exons: 8520/51443 (16.6%)

Missed introns: 1664/32564 (5.1%)

Novel introns: 8167/39562 (20.6%)

Missed loci: 0/8880 (0.0%)

Novel loci: 632/9230 (6.8%)

This analysis revealed that although the accuracy level was high at the base level (94.5%), the accuracy level was low at the locus level and even lower at the intron chain level (Table 3). The low sensitivity and precision values at the intron chain level indicated that many genes in the two annotations have gene model structures wrongly annotated. This was further highlighted by a high number of exons and introns that were annotated as novel. The analysis also grouped

Pb314v2 gene models in 13 categories based on transcript (Gene model) relationships class codes (Figure 9).

RNAseq data from 14 publicly available *P.brassicae* RNAseq (Table 1) projects were mapped to the Pb314v2 *P.brassicae* genome as a reference. The mapped transcripts were then used as evidence to support the accuracy of the existing gene model structures. As shown in table 2, manual inspection of PbPT3 and Pb314v2 gene models using RNAseq transcript data confirmed cases where expressed genes were not predicted by the annotation programs used in either *P.brassicae* genome annotation project (Figure 10)

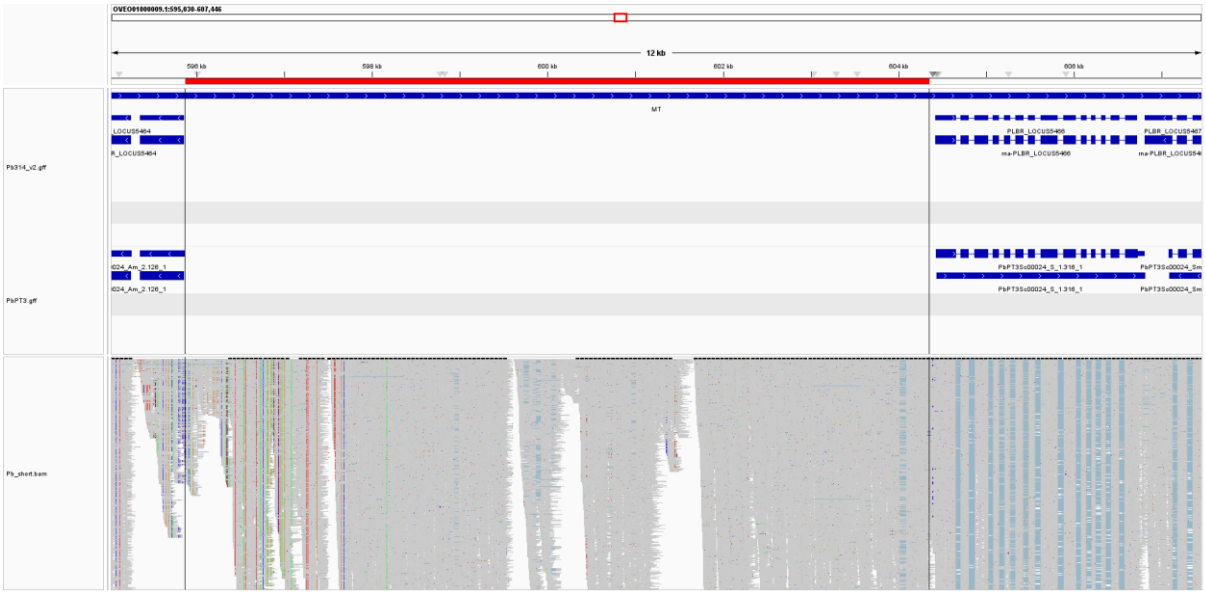


Figure 10: Missing gene models.
The red line indicates OVEO01000009.1:595,030-607,446 of scaffold 9 of the reference genome showing missing gene models in all annotations with transcript evidence.

Gene models assigned transcript relationship class code of "=", were expected to contain two groups of gene models where in group 1 gene models were expected to be correct because both query and reference annotations have exact matches for exon length and intron chains (Figure 11a). For group 2, gene models were expected to be incorrect because the gene models were assigned to this class because the internal intron coordinates matched between reference and query annotations but external exon lengths either matched or differed based on their start and end coordinates (Figure 11b).

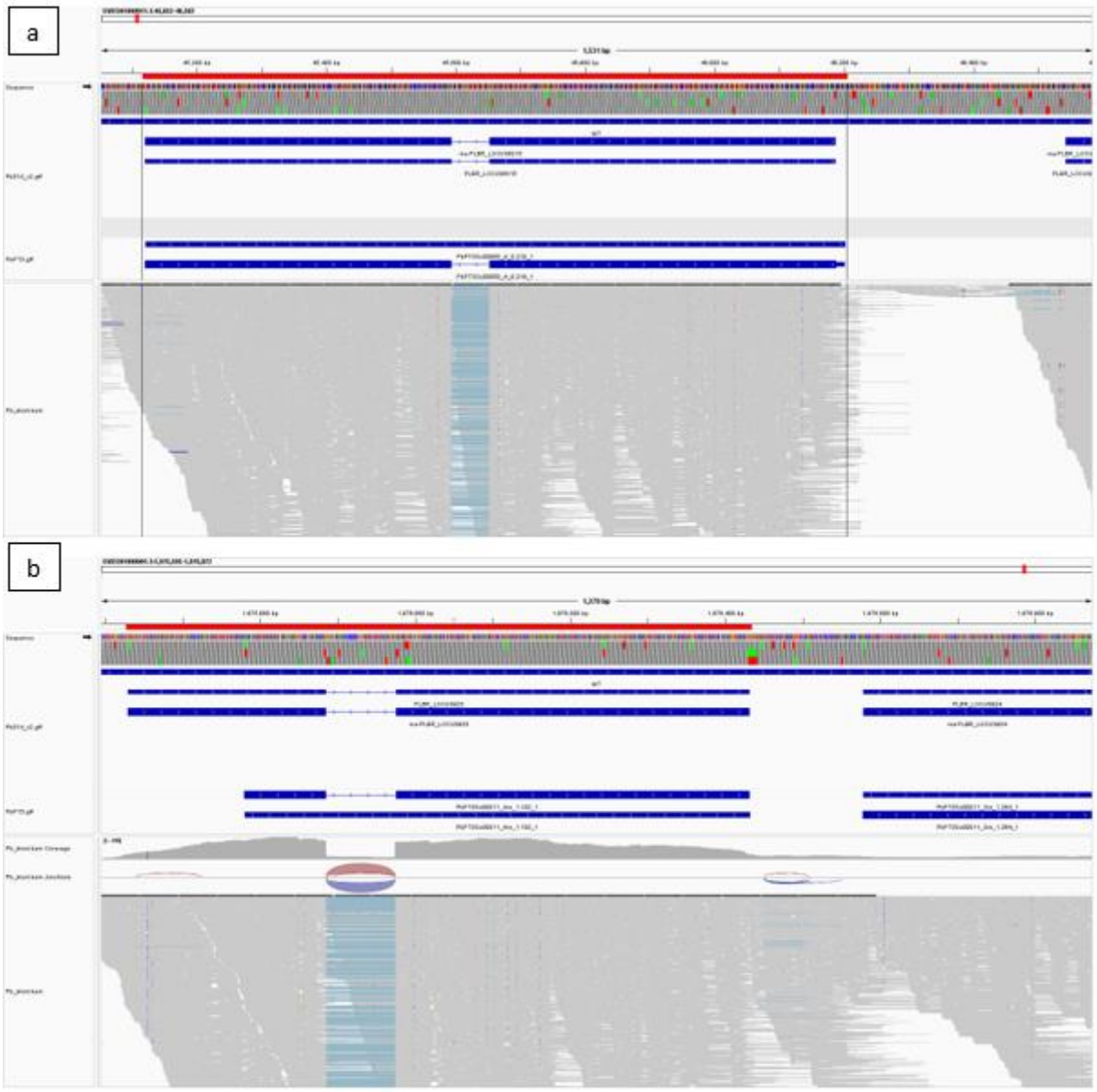


Figure 11: Manual inspection of gene models with a transcript relationship class code of "=". The red line indicates a predicated gene model with a match of exons and introns chains between reference and query annotations with regard to their start and end coordinates (a) and intron chains match, but external exon lengths differ due to variations in start and end coordinates (b).

Surprisingly, manual inspection identified two scenarios where incorrect gene models assigned class code of "=" could be regarded as correct. In both cases, the exons and introns for both query and reference annotations matched perfectly, but manual inspection identified this gene model to be incorrect because both annotations missed an exon that was supported by expression evidence (Figure 12a) or contained an intron which was not supported by expression data (Figure 12b).

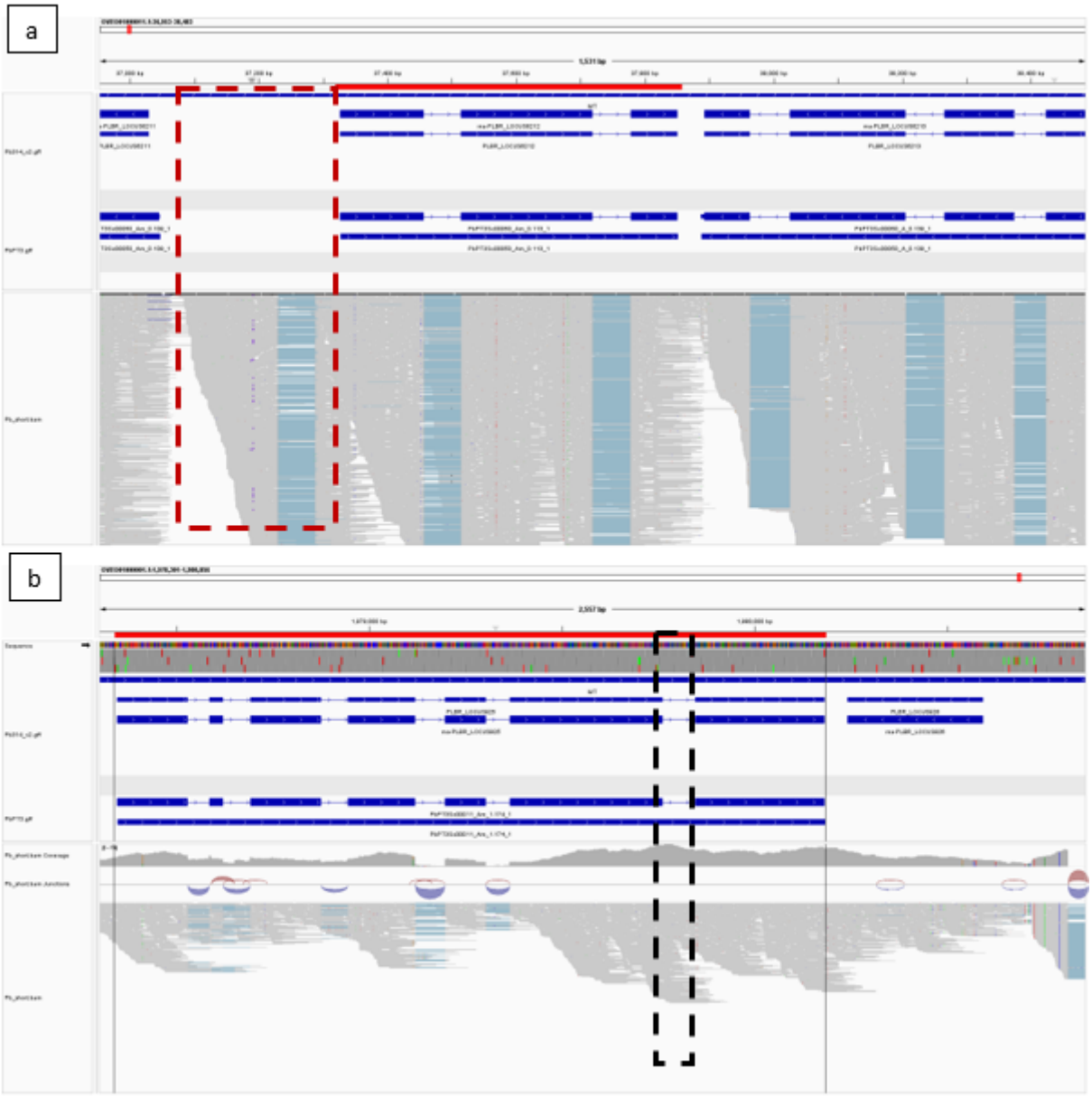


Figure 12: Incorrect gene models with a transcript relationship class code of "=".
a) shows incorrect gene model due to a missed exon which is indicated by a red rectangle and b) shows incorrected gene model with unsupported intron shown by a black rectangle.

2.3.2. Full length/ long read RNAseq data

Total RNA from *P.brassicae* -infected *Arabidopsis thaliana* was sequenced in two sequencing runs on Oxford Nanopore Technology (ONT) sequencing platform producing 2.02 and 19.81 million reads respectively. Data from the first sequencing run was not used due to very low number of reads that was generated. Reads with quality score below 9 were removed and the remaining reads were used for analysis. The read length varied from 0.11kb to 2.11kb (Figure 13) indicating that there were high number of full length/long reads. Reads were mapped to the *P.brassicae* genome, with a proportion of the reads successfully aligned (Table 4). Given the

time point post inoculation selected for the collection of sequencing samples, it was expected that 20 to 30% of the total reads will map the *P.brassicae* genome. The reason for low *P.brassicae* reads is unclear. Nevertheless, the data was useful in resolving some of the gene model errors that could not be resolved by the short read RNAseq data.

Table 4: Summary of full length/ long read mRNA sequencing

Total reads	Total reads after quality filtering	Reads mapped to *P.brassicae* genome
19.81 M	16.3 M	3,347,774

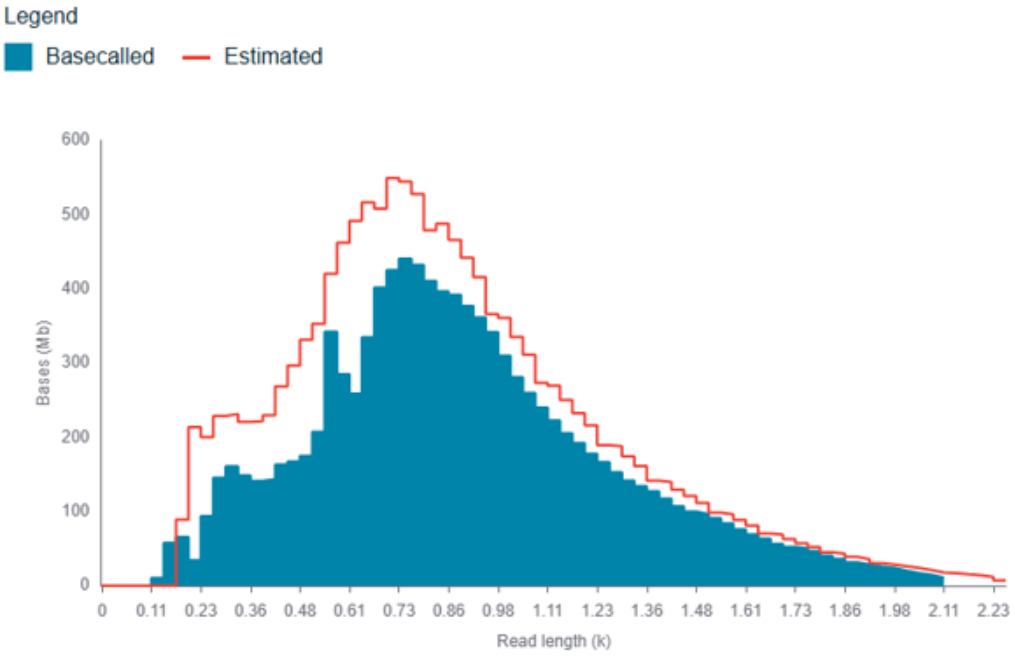


Figure 13: Distribution of read length.
The longest 1% of strands were classified as outliers and excluded to allow focus on the main body of data

2.3.3. Improving gene models by using full length/long read mRNA sequencing

Comparison of PbPT3 and Pb314v2 genome annotations showed that PbIPT1gene model encoded by PLBR_LOCUS7342 in Pb314v2 was similar to PbPT3Sc00040_Am_6.202_1 in PbPT3 annotation despite PLBR_LOCUS7342 having a long 5' end (Figure14). PLBR_LOCUS7342 gene model also had two in frame translation initiation start codons and was supported by short read RNAseq data (Black rectangle). *P.brassicae* IPT2 encoded PLBR_LOCUS5459 in Pb314v2 was identified as different from PbPT3Sc00024_Sm_2.181_1 in PbPT3. PLBR_LOCUS5459 had a long 3'end and was also supported by the RNAseq data (Figure S1). Because the sequences of *P.brassicae* IPT genes in PbPT3 annotation were contained in gene models of Pb314v2, it was not possible to identify the start and end positions

based on short reads RNAseq data. To address this challenge, full length/long read mRNA sequencing was done on *P.brassicae* infected Arabidopsis thaliana using oxford nanopore sequencing technology (ONT). These full length/long reads (Red broken rectangle) were used to correct PbIPTs gene models which could not be curated based on short read RNAseq data.

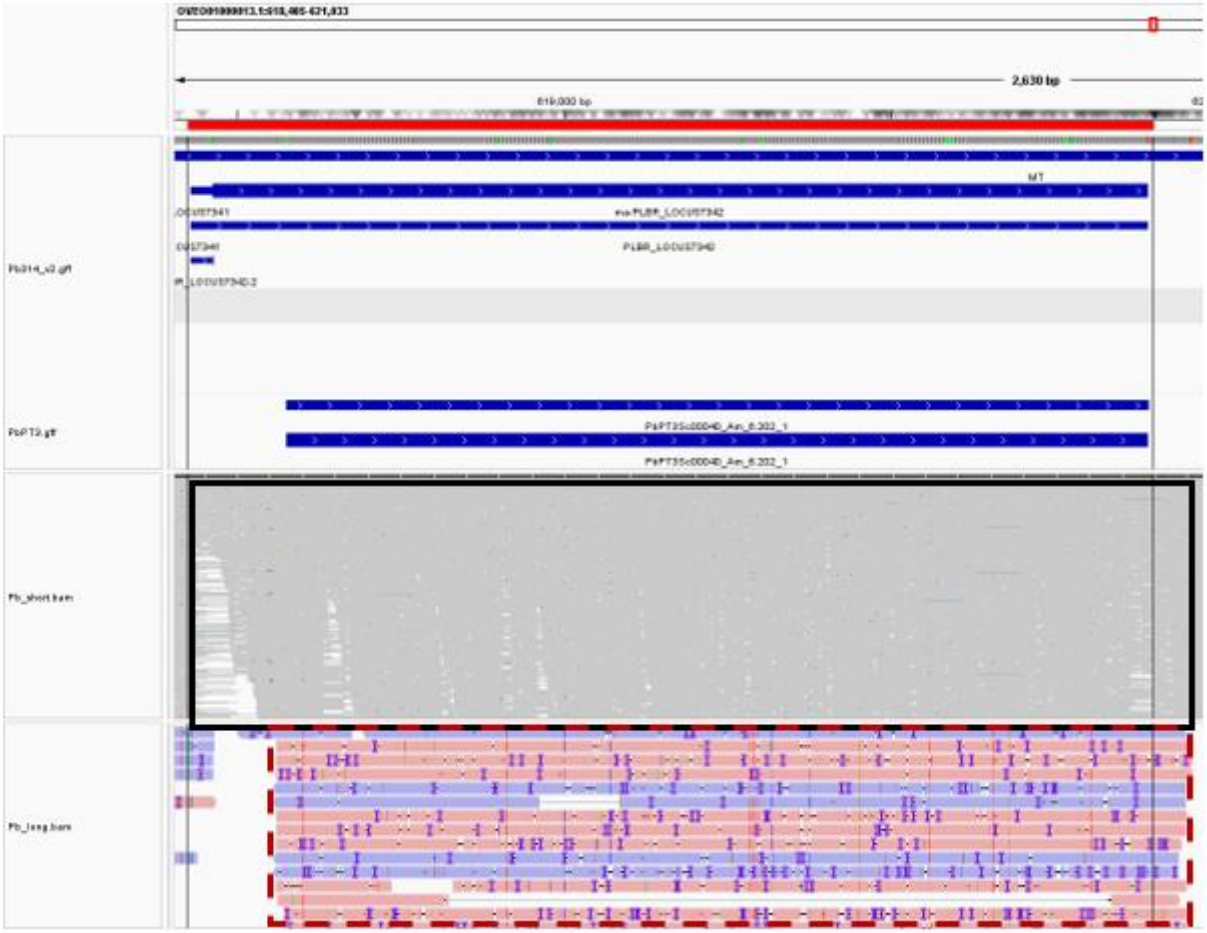


Figure 14: Manual curation of PbIPT1 gene models.
Black rectangle shows the alignment of short reads from RNAseq sequencing supporting the overhang at 5' end of PLBR_LOCUS7342. Red rectangle shows the alignment of full length/long reads from long read sequencing in support of PbIPT1 gene model PbPT3Sc00024_Sm_2.181_1 in PbPT3

The use of short-read RNAseq data failed to clearly support correct gene models in one annotation and incorrect gene models from another annotation in genomic regions that had uniform distribution of reads across the region. Long reads however clearly indicated the start and end of transcripts (Figure 15) which supported correct gene model structures if they were present from either annotation.

Figure 15: Correction of gene models using full length/ long read mRNA sequencing. The upper panel shows the Pb314v2 annotation which proposes 2 gene models in this region. The PbPT3 annotation shows a single model with unsupported introns. Short reads map to this region. Long read sequencing supports the Pb314v2 annotation. The middle panel indicates short read alignment, and the bottom panel show long read alignment.

The compact nature of *P.brassicae* genome means that several genes overlap which makes gene prediction programs struggle to predict gene models correctly in these regions. For example, PLBR_LOCUS5469 and PLBR_LOCUS5470 were annotated as two genes but the short reads alignment showed that these gene models should have been joined (Figure 16a). However, when full length/ long reads became available, the alignment of the reads revealed that this region encodes three different overlapping genes (Figure 16 b-d) while there were no long reads to support a proposed joined gene model (a).

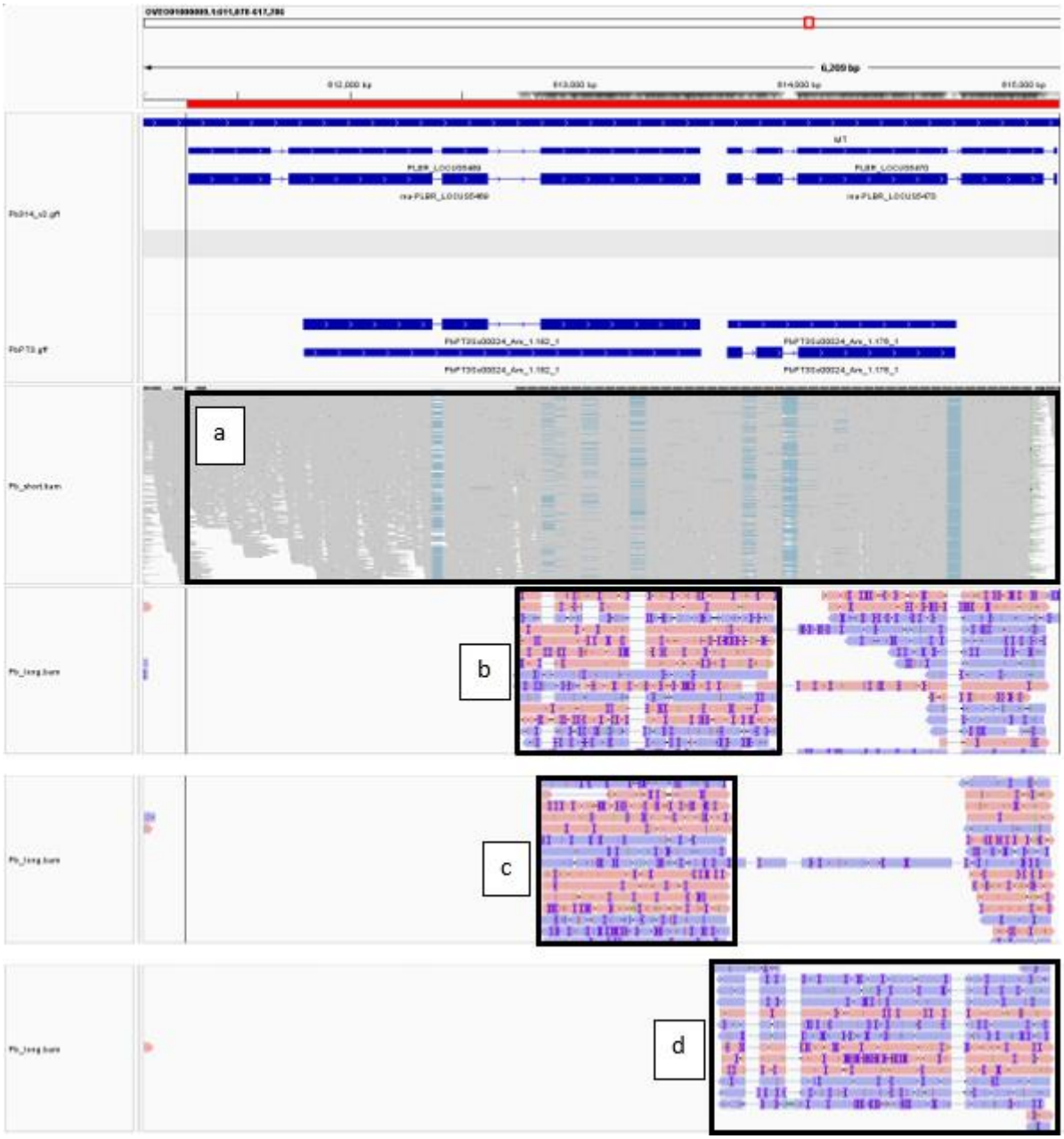


Figure 16: Overlapping genes.
(a) shows a proposed joined model of PLBR_LOCUS5469 and PLBR_LOCUS5470 (b) indicated a gene that combines the last exon of PLBR_LOCUS5469 and the first two exons of PLBR_LOCUS5470. (c) show exon four of PLBR_LOCUS5469 which should be annotated as a gene without intron and (d) shows PLBR_LOCUS5470 gene model with additional exon that is not supported by both short and long reads alignment. Gene models a-c are missing in the annotation of Pb314v2.

Full length/long read sequencing reads also were more informative in the identification of gene isoforms (Figure 17) which could not be solved by short read RNAseq expression data.

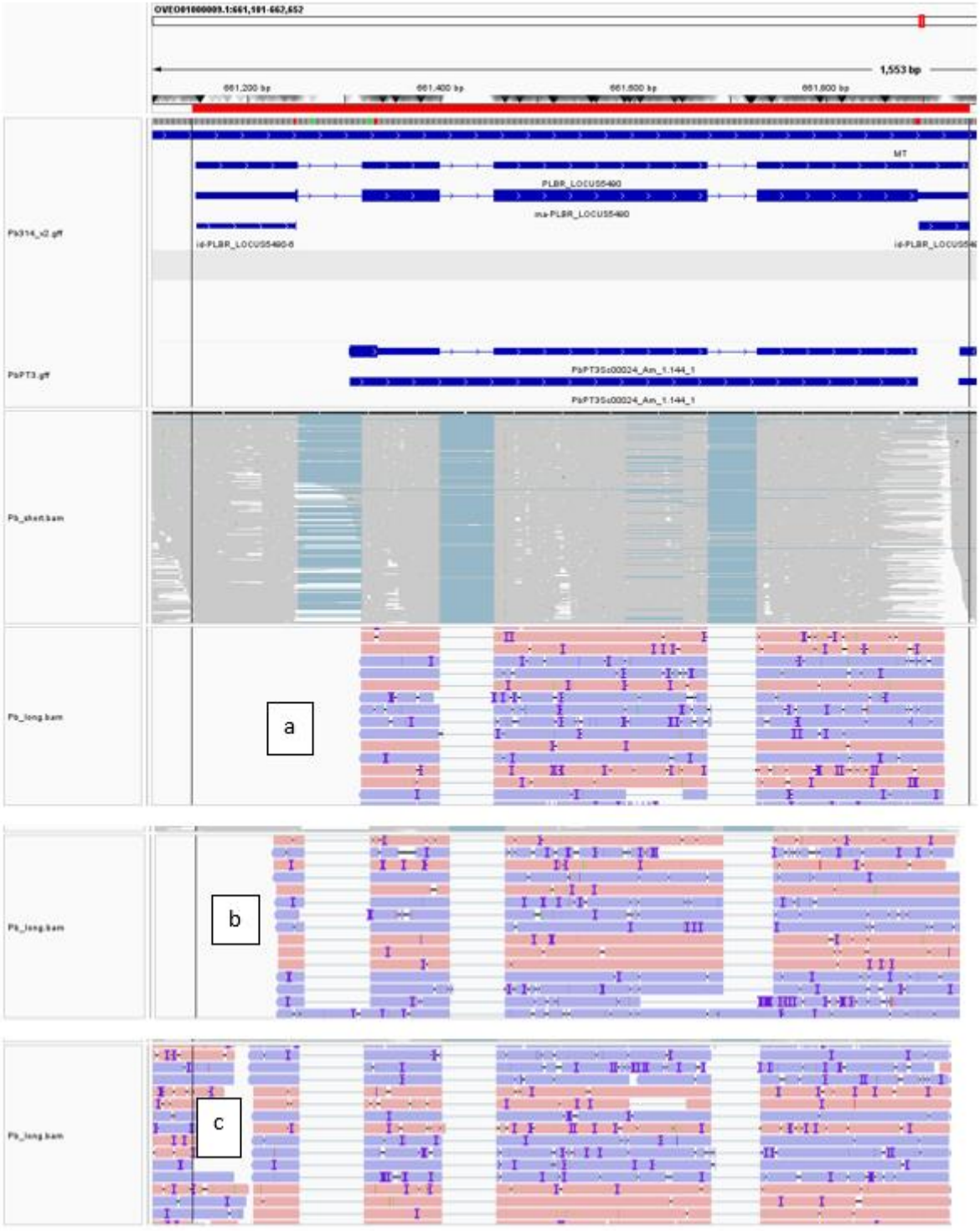


Figure 17: Different gene isoforms of PLBR_LOCUS5490.

2.4. Discussion

Most of the current *P.brassicae* gene models have been created using gene prediction computer programs. Although these programs are powerful and are continuously improved, they have inherent limitations and the best programs are able to predict 80% of gene structures correctly thus requiring manual curation (Williams et al., 2011). One obvious issue for *P.brassicae* genome annotation projects is the big differences in the number of predicted genes (Table 2). These differences are associated with gene prediction programs that fail to predict gene models or wrongly combine or split gene models. Due to the compact nature of *P.brassicae* genome which causes genes to overlap coupled with the insufficient supporting evidence used to create the initial gene models, the high number of missing and incorrect gene models is likely to occur. The accuracy of gene prediction heavily depends on the amount of supporting evidence used. For example, the re-annotation of the Arabidopsis genome to create Araport11, using transcript evidence from 113 RNAseq datasets, resulted in significant improvements compared to the TAIR10 annotation, which in addition to other supporting evidence relied on two RNAseq datasets (Cheng et al., 2017). This improvement was particularly noticeable in the low-confidence gene models of TAIR10, suggesting that these models lacked sufficient supporting evidence during the initial TAIR10 annotation (Cheng et al., 2017). Furthermore, gene prediction programs struggle to correctly predict gene structures of genes with weak translation initiation signals, short exons, weak splice sites, many exons, long introns or poorly conserved orthologs (Williams et al., 2011). A mere 45.1% of the Pb341v2 reference genome's gene models (Table 3) matched their corresponding PbPT3 gene models at the locus level. However, agreement at the transcript or locus level does not guarantee accuracy. Manual inspection, using short and long read expression data, revealed inaccuracies in some gene models that were initially deemed correct in both annotations (Figure 12). This suggests the actual percentage of correctly predicted gene models may be even lower than the observed 45.1%. High number of incorrect gene models in the initial gene model annotations have been reported in *C.elegans* where 54.4 % of the initial 19099 predicted genes have been manually corrected (Williams et al., 2011).

While RNAseq data used in this study helped correct some gene model errors (Figures 11b and 12), it was insufficient for others. Short read length of RNAseq datasets was reported to be limited in providing unambiguous gene boundaries for overlapping genes (Cheng et al., 2017). Long read sequencing in this study successfully resolved issues with the PbIPT1 and PbIPT2 gene models in the Pb314v2 genome annotation that were difficult or impossible to solve using

short-read alignment. Specifically, the long reads provided sufficient coverage to correct the 5' end of PbIPT1 and the 3' end of PbIPT2, which short-read alignment had failed to address (Figure 14 and Figure S1). This correction was crucial for ensuring the accurate functional characterization of the PbIPT1 and PbIPT2 proteins. However, while PacBio sequencing of *P.brassicae* -infected samples failed despite a high-quality sequencing library, ONT sequencing, though successful, yielded significantly fewer reads (Table 4) than the RNAseq data generated from samples collected at the same time point using Illumina technology (Table 1). Nevertheless, these reads improved gene model curation by confirming the accuracy of gene models in regions with uniformly distributed short reads (Figure 15) and enabled the identification of overlapping genes and splicing isoforms (Figures 16 and 17). Alternative splicing, a key driver of eukaryotic protein complexity, occurs in varying degrees across organisms. While approximately 40% of Arabidopsis (Cheng et al., 2017), genes encode multiple splicing isoforms, this figure rises to 81% in humans (Gerstein et al., 2014). The current *P.brassicae* gene models lack alternative splicing gene models, and this may have contributed to high number of incorrect gene models as failure to account for alternative splicing results in false gene models created by combining different gene isoforms (Cheng et al., 2017).

Correct exon-intron gene structure is crucial for maintaining high-quality genome annotations and this is achieved by manually inspecting gene models (McDonnell et al., 2018). The gene model errors identified in this study highlight the need for manual curation of *P.brassicae* gene models. This process will not only improve the accuracy of existing models but also likely reveal additional genes within the *P.brassicae* genome. For example, manual curation of initial *C. elegans* gene models significantly increased the gene count from 19,099 to 24,891, demonstrating the potential impact of this approach. Manual curation however is labour and resource intensive and is mainly done for model organisms. Since annotations from model organisms are limited in annotating the specific biology of other organisms and given the number of genomes being sequenced each year, the centralised approach of annotating model organisms cannot be scaled without significant funding (Menda *et al.*, 2008). Research community-based annotation methods have been developed to enhance the scale and quality of gene model curation and to increase the number of annotated organisms (Menda *et al.*, 2008). Research community annotation databases for specific organisms (Menda *et al.*, 2008; Bryant *et al.*, 2023) and for eukaryotic genomes (Humann *et al.*, 2019) have been created. To facilitate concerted efforts in creating a high quality and reliable *P.brassicae* reference genome through

manual curation of the current gene models, a web-based (Lee *et al.*, 2013) community driven *P.brassicae* annotation resource has been created.

2.5. Supplementary information

2.5.1. Supplementary figures

Figure S 1: Manual curation of PbIPT2 gene models.
Black rectangle shows the alignment of short reads from RNAseq sequencing supporting the overhang at 3' end of PLBR_LOCUS5459. Red rectangle shows the alignment of full length/long reads from long read sequencing in support of PbIPT2 gene model PLBR_LOCUS5459 in Pb314v2

CHAPTER3: GENOME-WIDE IDENTIFICATION OF PLASMODIOPHORA BRASSICAE TRANSPORTERS AND THEIR ROLE IN NUTRITION.

Abstract

Biotrophic pathogens create a nutrient sink at the infection site by reprogramming the host's metabolism to divert nutrients toward these areas, ensuring the host is weakened but not killed. Carbon acquisition is critical for the development and proliferation of pathogens, as it provides energy and building blocks for cellular components and signalling. Carbohydrates are the primary source of energy and carbon for many pathogens, while biotrophs also acquire essential nutrients as simple compounds, such as inorganic phosphates, or more complex ones, ranging from hexoses and sucrose to complex vitamins. *P.brassicae* redirects host carbohydrates in infected roots by stimulating phloem proliferation and recruiting host sugar transporters in the developing galls. However, surprisingly little is known about how *P.brassicae* acquires carbon and other nutrients, and which transporters are involved in this process. This chapter focuses on identifying the nutrient transporter families in the *P.brassicae* genome, creating a comprehensive catalogue of *P.brassicae* transporters. To validate the bioinformatics methods, gene expression analysis was conducted using publicly available RNAseq data, and the heterologous expression of the *P.brassicae* SWEET transporter in a yeast sucrose transporter mutant was performed. Comparing clubroot transport systems with those of two other obligate biotrophic pathogens revealed that *P.brassicae* possesses more transporters and unique transport systems not found in these other pathogens. A detailed analysis of clubroot transporters uncovered unusual transporters within its genome. This work highlights the diverse transporters expressed by *P.brassicae* during disease development, further emphasizing the crucial role of nutrient transporters in the progression of clubroot disease.

3.1. Introduction

Biotrophic pathogens are found on, or in, living plants where they obtain nutrients without rapidly killing the host (Chen *et al.*, 2010). Biotrophic pathogens create a nutrient sink at the infection site by reprograming the host metabolism to distribute nutrients preferentially to infection sites, so that the host is disadvantaged but is not killed (Doidy *et al.*, 2012; Sosso *et al.*, 2019). Carbon acquisition is crucial for the development and proliferation of pathogens, providing energy and building blocks of cellular components and signalling (Goulet & Saville, 2017). Carbohydrates are the main source of energy and carbon for many pathogens. Biotrophs also acquire some essential nutrients as simple compounds e.g. inorganic phosphates, but others may be more complicated (ranging from simple hexose/sucrose to highly complex vitamins). As a consequence, biotrophs express numerous transporter genes that enable diverse nutrients to be acquired from the host (Chen *et al.*, 2010; Wahl *et al.*, 2010).

Plasmodiophora brassicae is an intracellular obligate biotrophic pathogen that causes clubroot disease, one of the most damaging diseases within the family Brassicaceae, including the model plant, *Arabidopsis thaliana* (Siemens *et al.*, 2002; Dixon, 2009). Infected roots and hypocotyls undergo abnormal cell enlargement and uncontrolled cell division leading to the development of galls that provide space for *P. brassicae* growth and nutrient acquisition. Soluble sugars, such as sucrose and hexoses, are probably the major form in which carbon is supplied although the precise form in which they are acquired by *P.brassicae* is unclear (Keen & Williams, 1969; Brodmann *et al.*, 2002; Li *et al.*, 2018). *P. brassicae* redirects the host carbohydrates in the infected roots by stimulating phloem proliferation and recruiting host sugar transporters in the developing galls (Li *et al.*, 2018; Walerowski *et al.*, 2018; Zhang *et al.*, 2019). Transcriptome analysis of *P.brassicae* -infected plants has revealed that host sugar transporters, such as sugar permeases or hexose transporters, are expressed in the roots and may be involved in the unloading of sugars in the infection site (Li *et al.*, 2018; Walerowski *et al.*, 2018; Zhang *et al.*, 2019). This facilitates the flow of nutrients to the galls and subsequent metabolism of these nutrients drives the sink strength of the galls. Brodmann et al. (2002) found that roots and hypocotyls of infected Arabidopsis plants accumulated high levels of glucose, fructose and starch. They also accumulated large amounts of trehalose, a disaccharide sugar synthesised in small amount by Arabidopsis, but large amounts by *P. brassicae.*

Most of plant pathogen carbohydrate transporters that have been identified so far facilitate the uptake of hexoses (Voegele *et al.*, 2001; Schuler *et al.*, 2015; Chang *et al.*, 2020; Yuan *et al.*, 2021). Extracellular hydrolysis of sucrose to hexoses by pathogens can trigger plant defence

responses, such as the expression of pathogenesis-related genes (Schaarschmidt *et al.*, 2007; Opitz *et al.*, 2021). Wahl et al. (2010) identified a sucrose transporter (Srt1) in smut fungus *Ustilago maydis* that was localised on the plasma membrane of the haustoria. Srt1 has high sucrose affinity enabling the pathogen to directly uptake sucrose from the host, circumventing these defence responses that are elicited by the extracellular breakdown of sucrose into hexose sugars. Putative hexose and sucrose transporters have also been identified in *P.brassicae* but are yet to be confirmed experimentally (Rolfe *et al.*, 2016). Recently, 4 putative sugar transporters in the Major facilitator superfamily (MSF) and 13 from the SWEET family were found in *P.brassicae* (Kong *et al.*, 2022). When expressed in EBYVW400 yeast strain that is unable to transport glucose and fructose, ten of the putative hexose transporters restored the growth of the strain on fructose or glucose indicating that these *P.brassicae* transporters can transport hexose sugars. As most of its lifecycle within the plant is intracellular, it would be difficult to generate a proton gradient (the cytoplasm is pH 7) so SWEETs would be favourable sugar transporters. Moreover, proton gradient may not be needed as metabolism of hexoses or conversion of sucrose to trehalose would drive the sink strength of plasmodia.

The *P.brassicae* genome also encodes unusual putative transporters which have a prokaryotic-like substrate binding domain (SBD) that belong to the ATP binding cassette (ABC) transporter superfamily fused to a Tm7 transmembrane helix domain. In recognition of their prokaryotic-eukaryotic structure, we have designated these PRO-EU transporters. The PRO-EU transporters encode domains predicted to transport either sugars and sugar alcohols, or phosphates (Rolfe *et al.*, 2016). Initially, these transporters were thought to be unique to *P.brassicae* , but homologs have also been found in anaerobic gut fungi (Seppala *et al.*, 2016), although the fungal homologs exhibit extensive extracellular scaffolds and carbohydrate-active enzymes (CAZymes) which are lacking in *P.brassicae* . In the later stages of infection the enlarged galls and high metabolic activity generate anaerobic conditions in galls (Gravot *et al.*, 2016) and it is tempting to speculate that these environmental similarities might be associated with PRO-EUs, although this remains unproven. Additionally, the genomes of anaerobic gut fungi contain SWEET transporters (Seppälä et al. 2016) allowing the fungi novel co-consumption of glucose and xylose (Podolsky *et al.*, 2021).

P. brassicae has lost the ability to synthesise many essential macromolecules (e.g. vitamins, amino acids) (Schwelm *et al.*, 2015; Rolfe *et al.*, 2016) as they are biochemically expensive to make and acquiring them, ‘ready-made’ is energetically more favourable (Rolfe et al. 2016). Lack or presence of impaired biosynthetic pathways of some important nutrients and expansion

of some specific transporter gene families is a characteristic lifestyle of obligate biotrophic pathogens (Guerillot *et al.*, 2023). Genomic studies in rust fungi biotrophs have revealed they are unable to directly assimilate inorganic nitrates and sulphur thus acquiring them in complex forms (Duplessis *et al.*, 2011). They have been reported to have transporters that acquire nitrogen in the form of ammonia, amino acids and oligopeptides (Struck, 2015). In rust fungi, expression of amino acid transporters has been reported to occur mainly in the haustoria (Hahn *et al.*, 1997; Struck *et al.*, 2002; Struck *et al.*, 2004). Biotrophs require essential nutrients that are needed by different cellular biological processes such as synthesis of nucleic acids, phospholipids, enzyme cofactors among others. For example, *Puccinia striiformis* f.sp. *tritici* (*Pst*) obtains these essential nutrients from the host by expressing inorganic phosphate transporters and vitamin B transporters in the haustoria. The pathogen also expressed a glycerol transporter in the germinating spores, a compound that could be incorporated in the gluconeogenic pathway (Garnica *et al.*, 2013). Like these rust fungi biotrophs, *P.brassicae* might have transporters that enable it to get these nutrients which it is unable to synthesise.

The role of host nutrient transporters in delivering nutrients to host cells in the galls is documented. However, for *P.brassicae* to be able to use host nutrients, it must import these across its plasma membrane, presumably requiring the expression of its own transporters (Figure 2). Surprisingly, little is known about the carbon source uptake and the transporters involved in *P. brassicae* nutrient acquisition from its host plants, in contrast to other important plant biotrophic pathogens where there is a wealth of knowledge regarding chemical nature of nutrient and the mode of acquisition. This is partly due to its strict obligate biotrophic lifestyle meaning it cannot be grown in culture, the lack of transformation systems and its unusual intracellular lifestyle compared with most (hemi-)biotrophs which are extracellular. The availability of the complete *P.brassicae* genome sequence presents an opportunity to computationally identify putative nutrient transporters, despite challenges posed by its unique biology. However, the compact nature of the *P.brassicae* genome, characterised by small intergenic regions and overlapping genes, can lead to inaccurate gene model predictions, diminishing the reliability of bioinformatic analyses. Nevertheless, *in silico* homology-based approaches remain the most effective method for identifying and cataloguing *P. brassicae* transporters. This is because results obtained from *in silico* analyses can be rigorously validated through independent bioinformatic methods and subsequent experimental verification

This study utilised the *P.brassicae* genome Pb314V2 as a reference to conduct bioinformatic analyses aimed at identifying putative *P.brassicae* transporters, thereby enhancing understanding of this unique plant pathogen's physiology. The primary objective was to identify and characterise nutrient transporters crucial for *P.brassica'*s ability to acquire and utilise host-derived nutrients during its interaction with the host plant. Sugar transporters were of particular interest due to their pivotal role in enabling the pathogen to utilise host carbohydrates as a carbon and energy source. The spatial and temporal expression patterns of putative sugar transporters were subsequently analysed during clubroot disease development. One putative sugar transporter was further characterised through heterologous expression in yeast.

3.2. Materials and methods

3.2.1. Identification of putative transporters and their substrates

The Pb314v2 was selected as a reference genome for analyses as it was the most complete, and scaffolds most likely match complete chromosomes. However, a new *P.brassicae* genome with telomere-to-telomere assembly has been published (Javed *et al.*, 2024). Protein sequences from the Pb314v2 gene models were extracted using AGAT program (https://github.com/NBISweden/AGAT).

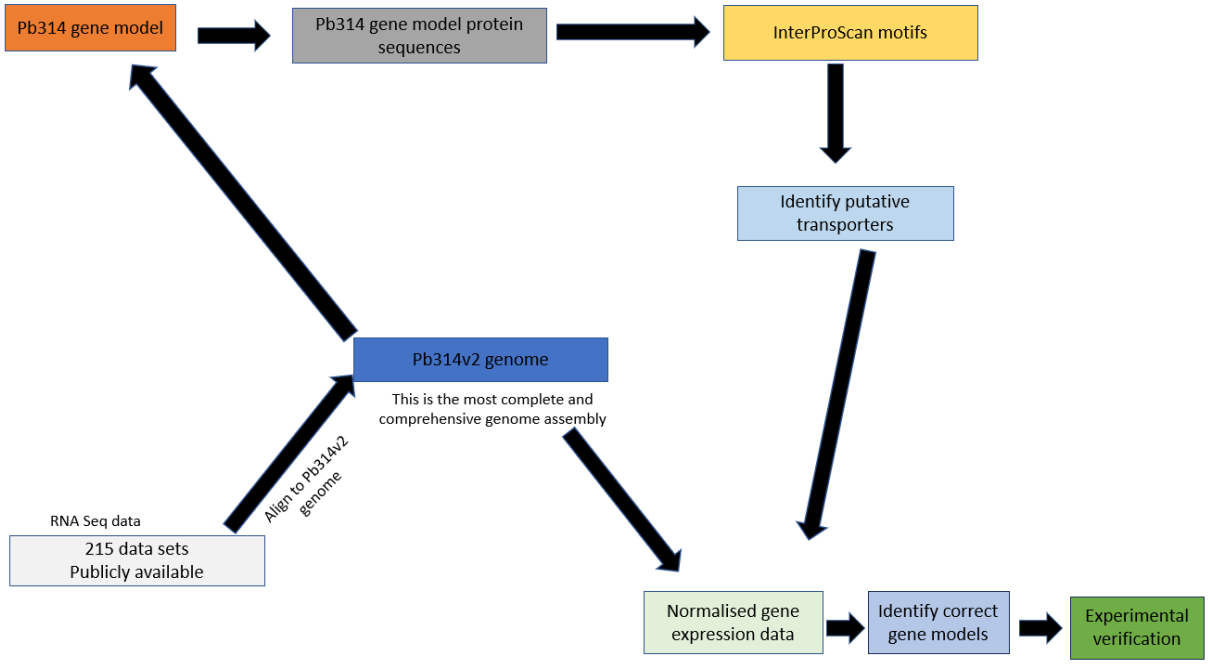


Figure 18: Flowchart diagram detailing steps taken to identify putative transporters from P.brassicae gene models

Protein domains (PFAM domains) present in *P.brassicae* genome were identified by analysing gene models using InterProScan (Blum *et al.*, 2021). To identify putative *P.brassicae* transporters, *P.brassicae* protein domains were subsequently compared to PFAM domains associated with cellular transport, as curated in the transporter classification database (TCDB) (Saier *et al.*, 2021). The TCDB PFAM domains utilised in this study are accessible under the 'TC-DOM' tab, a tab-delineated table mapping Pfam entries to TC systems. Putative transporters were assigned to superfamilies, families and subfamilies in accordance with the TC-system nomenclature using the transport classification identity code (TCID) associated with the PFAM domains (Saier *et al.*, 2021). TCDB analysis of *P.brassicae* transporters yielded multiple TCIDs per transporter. Since a single PFAM domain within a gene could be linked to more than one TCID, the output was summarised by selecting the first TCID encountered for

each unique gene ID and PFAM domain combination. Given that proteins can have multiple domains, indicating diverse potential functions, genes that had more than one PFAM domain were counted only once to prevent inflating the number of putative transporters. Putative substrates for each transporter were searched from the TCDB by using the TCID assigned to the transporters.

3.2.2. Transporter expression analysis

3.2.2.1. Single cell RNA sequencing

The asynchronous development of *P.brassicae* results in infected plant roots harbouring the pathogen at diverse developmental stages during the course of the disease cycle (Feng *et al.*, 2013b; Liu *et al.*, 2020). To gain understanding of pathogen gene expression in different pathogen developmental stages within the infected roots, single cell transcriptomic approach was used. Protoplasts were isolated from both healthy and *P.brassicae* -infected Chinese cabbage (Wong Bok) plants. Protoplast isolation was performed following the Arabidopsis protoplast isolation protocol (Shahan et al., 2022) with modifications. Roots from healthy and *P.brassicae*-infected Chinese cabbage were harvested 23 days post-inoculation. The roots and galls were thoroughly cleaned with distilled water and then sliced into small pieces using a surgical blade in base buffer lacking enzymes. The base buffer was prepared by adding 21.86g of mannitol, 1.17g of MESS, 0.447g of KCl, and 0.442g $CaCl_2$ in 300 mL of deionised water. The base buffer pH was adjusted to 5.7 using KOH, followed by the addition of 0.1% (w/v) bovine serum albumin and 0.000194% (v/v) β-mercaptoethanol. To extract protoplasts, sliced roots and galls were transferred to a petri dish containing 10 mL of protoplasting buffer (base buffer supplemented with 1% cellulase R-10 and 0.25% Macerozyme R-10, Duchefa Biochemie B.V., The Netherlands). The petri dishes were then sealed with parafilm, covered with aluminium foil, and incubated at room temperature on a rotary shaker at 100 rpm for 90 minutes. Following digestion, the cell suspension was filtered sequentially through 70 µm and 30 µm cell strainers mounted on 50 mL Falcon tubes. The filtrate was then centrifuged at 250 x g for 5 minutes at 4°C in a swinging bucket centrifuge using minimum acceleration and deceleration. The resulting pellet was washed twice by resuspension in 5 mL of base buffer followed by centrifugation at 100 x g for 3 minutes. The pellet was resuspended in 1 mL of base buffer. To assess cell viability, concentration, and the presence of *P.brassicae* plasmodia , 10 µL of cell suspension was stained with 1 µL of Nile Red and loaded onto a Neubauer haemocytometer. Cell counting and the detection of *P.brassicae* plasmodia were performed using a fluorescence microscope (BX51 Olympus) with excitation at 360 nm and emission

detection at 430 nm. Subsequently, to determine the number of viable cells, 10 µL of cell suspension was stained with 1 µL of trypan blue and analysed using fluorescence microscope (BX51 Olympus) with excitation at 488 nm and emission detection from 500 nm. The protoplast suspension was then taken to NEOF-NERC environmental omics facility, university of Sheffield for single cell sequencing using 10X Genomics technology.

3.2.2.2. Publicly available RNAseq Datasets

Short read sequences from 14 publicly available RNAseq projects (Table 1) were mapped to the *P.brassicae* genome using STAR (Dobin *et al.*, 2013). Normalised gene expression and differential gene expression were calculated using R package DESQ2 (Anders & Huber, 2010). Control samples and genes with thirty or less reads across the samples were removed before normalisation of the counts. The normalised read counts expressed as fragments per kilobase of transcripts per million mapped reads (FPKM) was used to analyse gene expression profiles. A transporter was considered as expressed if the sum of counts across all the samples was greater or equal to 10 FPKM. Expression of genes with lower FPKM (<=2FPKM) read counts were manually inspected using the Integrative Genome Viewer (IGV). Transporters that were differentially expressed in resting spores, germinating spores and cortical infection stages were analysed using data from PRJNA331021 RNAseq project. Project PRJNA331021 was selected because it had the most reads mapped to *P.brassicae* genome in different pathogen development stages. Transporter spatial and temporal expression profiles were visualised by using Pheatmap R package (Kolde, 2018). Transporter expression profiles were scaled by row.

3.2.3. Identification of unusual and rare transporters in the *P.brassicae* genome

The unusual transporters with prokaryotic-eukaryotic structure designated as PRO-EU were identified by filtering *P.brassicae* PFAM domains present in bacterial solute binding proteins (SBP) PFAM domains (Ortega *et al.*, 2022).

Table 5: Bacterial solute binding domains used to identify *P.brassicae* PRO-EU transporters

PFAM	SBP FAMILY	PFAM	SBP FAMILY
PF00497	SBP_bac_3	PF03180	Lipoprotein_9
PF00496	SBP_bac_5	PF12849	PBP_like_2
PF13416	SBP_bac_8	PF01297	ZnuA
PF03401	TctC	PF02608	Bmp
PF13407	Peripla_BP_4	PF16868	NMT1_3
PF01497	Peripla_BP_2	PF04392	ABC_sub_bind
PF13458	Peripla_BP_6	PF04348	LppC
PF09084	NMT1	PF00532	Peripla_BP_1
PF01547	SBP_bac_1	PF05048	NosD
PF03480	DctP	PF12727	PBP_like
PF04069	OpuAC	PF01094	ANF_receptor

PFAM domains associated with putative *P.brassicae* transporters were individually analysed using InterProScan, specifically utilising the 'Browse by Member databases' feature to access and query the PFAM database. The InterPro database provides an overview of information for each PFAM domain, including taxonomic data. This is displayed as a sunburst diagram (Figure 19) and colour coded by kingdom. Organisms sharing the same protein domain as *P.brassicae* were selected using the taxonomy tab, weighting segments by sequence number, while other settings remained default. *P.brassicae* genes were referred to as: a) 'PRO-EU' either if 95% of the sunburst is prokaryotes and 5% is eukaryotic organisms (Figure 19a, PF00005) or more than 98 % of the sunburst is prokaryotes and less than 2% is no kingdom (Figure 19b, PF10531) and also had one of the 22 bacteria SBP domains (Table 5). b) 'not common' in eukaryotes either if 95% of the sunburst is prokaryotes and 5% is eukaryotic organisms (Figure 19a, PF00005) or more than 98 % of the sunburst is prokaryotes and less than 2% is no kingdom (Figure 19b, PF10531) and lacked bacterial SBP domain. These transporters were considered as 'not common' in eukaryotes because they were found in few organisms of different eukaryotic kingdoms. c) 'restricted' to few organisms if only 95% of the sunburst is occupied at most by two eukaryotic kingdoms and 5% is prokaryotes (Figure 19c, PF03083).


A
Legends
bacteria
viruses
archaea
eukaryota
Other
Weight Segments by
Number of sequences
Font Size
14
Sunburst Depth
5 rings
2
8
Selected Taxon
B
Legends
bacteria
viruses
archaea
eukaryota
Other
Weight Segments by
Number of sequences
Font Size
14
Sunburst Depth
5 rings
2
8
Selected Taxon

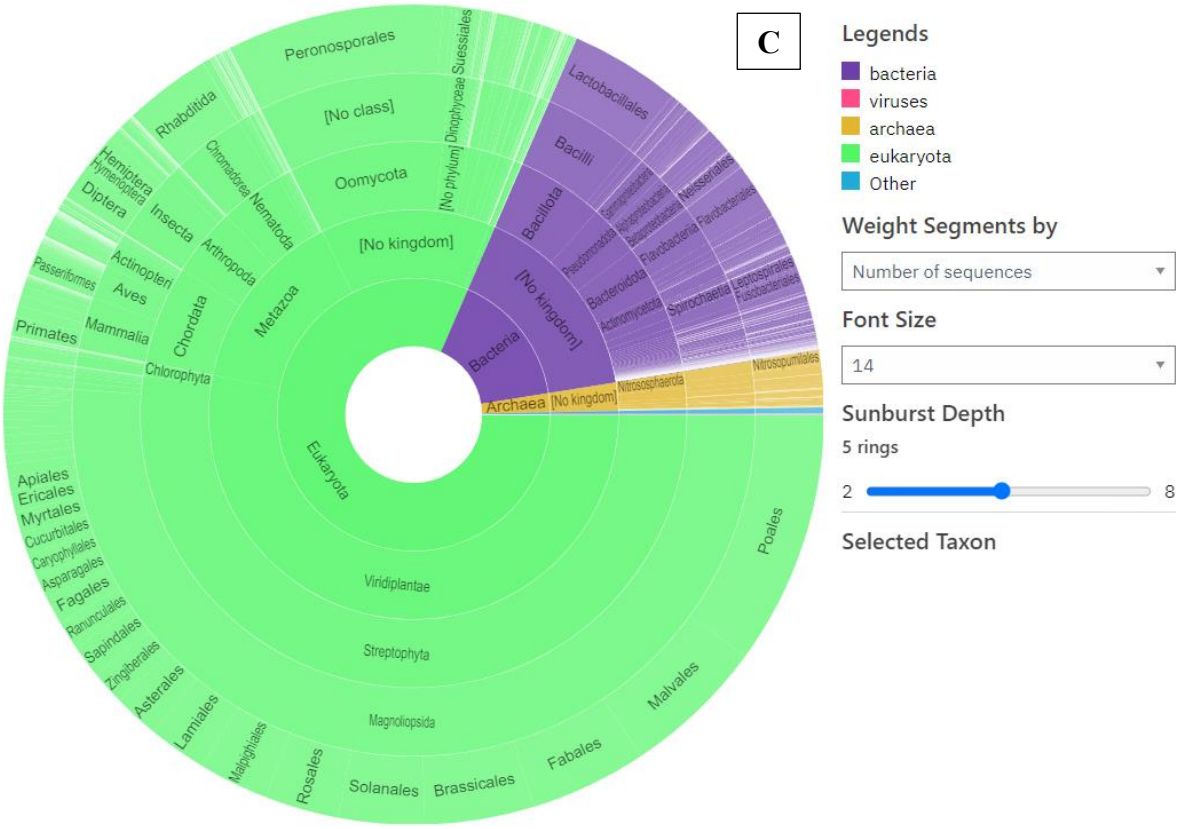


Figure 19: Categorisation of unusual and rare P. brassicae transporters.
Transporters were categorized based on their distribution within the sunburst plot and the presence or absence of bacterial substrate-binding proteins (SBP). Transporters were grouped as PRO-EU (A) if At least 95% of the sunburst plot represents prokaryotes and up to 5% are eukaryotic organisms and contained a bacterial SBP; Not Common in Eukaryotes (B) if Over 98% of the sunburst plot consists of prokaryotes, with less than 2% having no assigned kingdom and lacked a bacterial SBP and as Restricted to Few Organisms (C) if At least 95% of the sunburst plot is occupied by a maximum of two eukaryotic kingdoms, with up to 5% being prokaryotes.

3.2.4. Refinement of unusual and rare transporter gene models

For refinement and improving the accuracy of gene models, BAM files created by mapping of long reads/full length RNA reads and publicly available short RNAseq reads (Figure 8) together with gene models of unusual and rare transporters were displayed for comparison using custom code written in R. A gene model was considered as correct if it was supported by alignment of either short reads, long/full length RNA reads or both.

3.2.5. Transport systems present in *P.brassicae* and other obligate biotrophic plant pathogens

Transporters associated with the fungal pathogens *Puccinia triticina-pt76* (Duan, H *et al.*, 2022) and *Ustilago maydis 521* (Kämper *et al.*, 2006) with TCDB annotations were identified and obtained from (https://mycocosm.jgi.doe.gov/). This involved navigating to the website and selecting the corresponding taxonomic groups: Pucciniomycotina for *P. triticina-pt76* and Ustilaginomycotina for *U. maydis 521*. Next, the 'transporter' tab was selected, allowing access to transporters of organisms in Pucciniomycotina and Ustilaginomycotina. Finally, the transporters of *P. triticina-pt76* or *U. maydis 521* were downloaded by choosing the desired organism from the available transporter table. The transport systems of *P. triticina-pt76* or *U. maydis 521* were then compared to those of *P.brassicae* to assess if the *P.brassicae* transport system was complete in relation to other obligate biotrophic plant pathogens.

3.2.6. Functional characterization of PbSWEET transporter by heterologous expression in yeast

P.brassicae SWEET transporter (PLBR_LOCUS1241) was characterised using growth complementation assay in an invertase mutant yeast strain CSY4000 (Li *et al.*, 2020). CSY4000 strain cannot grow on hexose and sucrose because of concurrent knockout of endogenous hexose transporters and SUC2, which encodes an invertase. The genotype of CSY4000 is MATa leu2-3,112 ura3-52 trp1-289 his3Δ1 MAL2-8c SUC2::LEU2 Δhxt1-17 gal2Δ::loxP stl1Δ::loxP agt1Δ::loxP mph2Δ::loxP mph3Δ::loxP . PbSWEET construct (Appendix 1) was synthesised and codon-optimized for expression in *E. coli* by Thermo Fisher Scientific Inc. To assess if PbSWEET transports sucrose, the synthetic PbSWEET gene construct was cloned in pDR196 yeast expression vector (Meyer *et al.*, 2006). Vector free CSY4000 and CSY4000 with empty vector were used as negative controls.

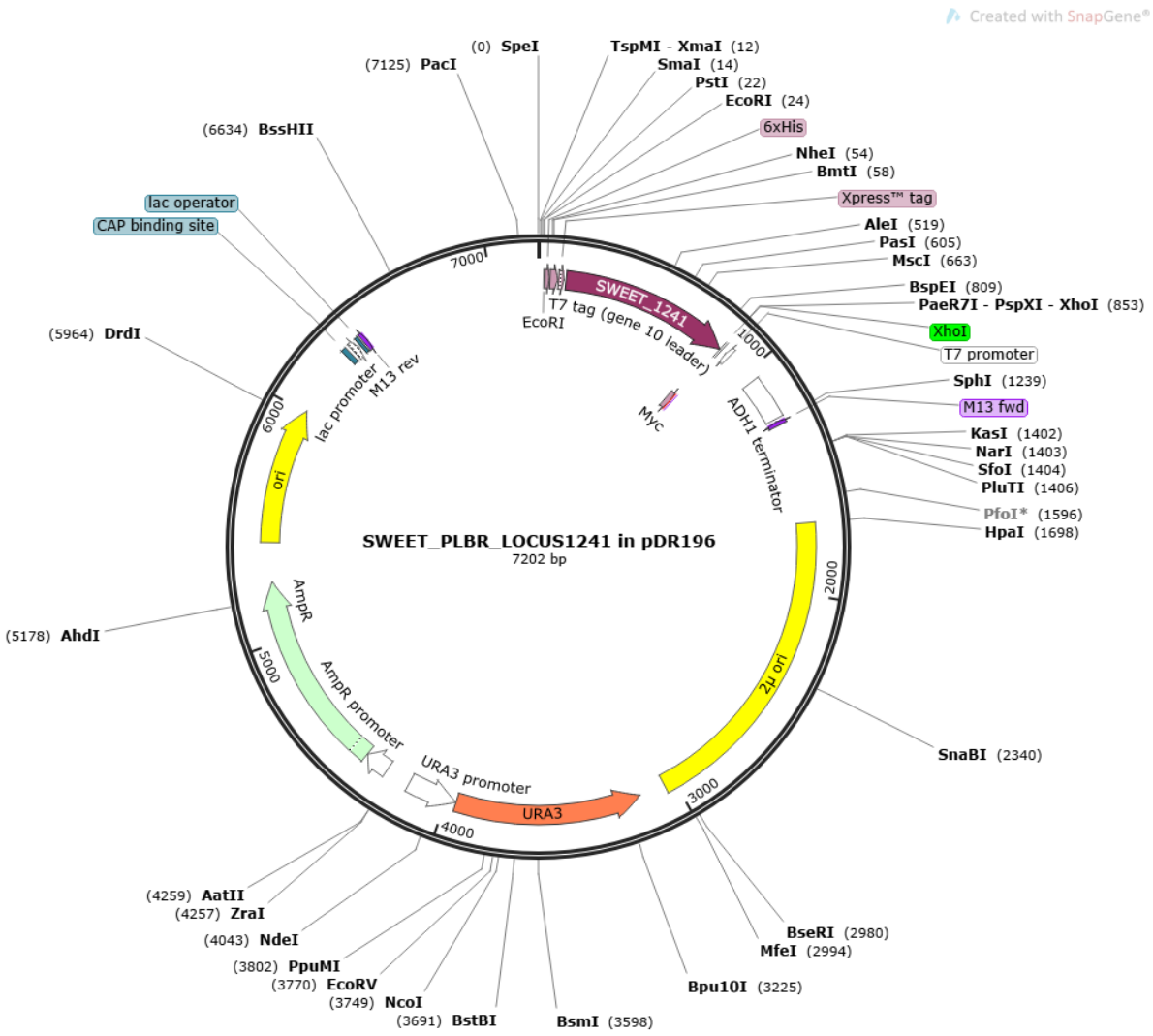


Figure 20: Vector map showing the cloning of P.brassicae SWEET gene in the yeast expression vector pDR196

3.3. Results

3.3.1. Identification of genes encoding putative transporters and transport types in *P.brassicae*

To understand how *P.brassicae* manipulates its host development and nutrition, the *P.brassicae* genome was searched for genes encoding transport proteins. In total, 2172 genes (Table S1) were predicted to encode transporters. These transporters were identified by filtering PFAM domains associated with transport that were found in *P.brassicae* gene models. Because a protein can have multiple domains indicating diverse potential functions, transporters that had more than one PFAM protein domain involved in transport were counted once. Out of the initial 2172 putative transporters identified, 327 were considered duplicated and hence counted once giving a final set of 1845 unique transporters in *P.brassicae* equivalent to 18% of total *P.brassicae* genes reported in the Pb314 reference genome (Stjelja et al., 2019). To determine the transport systems present in *P.brassicae* (Figure 21), the transporter classification database (TCDB) (Saier et al., 2021) was utilised. There are seven categories (TC classes 1-5,8,9), the last two (TC classes 8 and 9) being less clearly defined. All the 1845 putative *P.brassicae* transporters were classified as belonging to one of these classes (Figure 21) with 74 % (1360) in the well-defined classes 1-5.

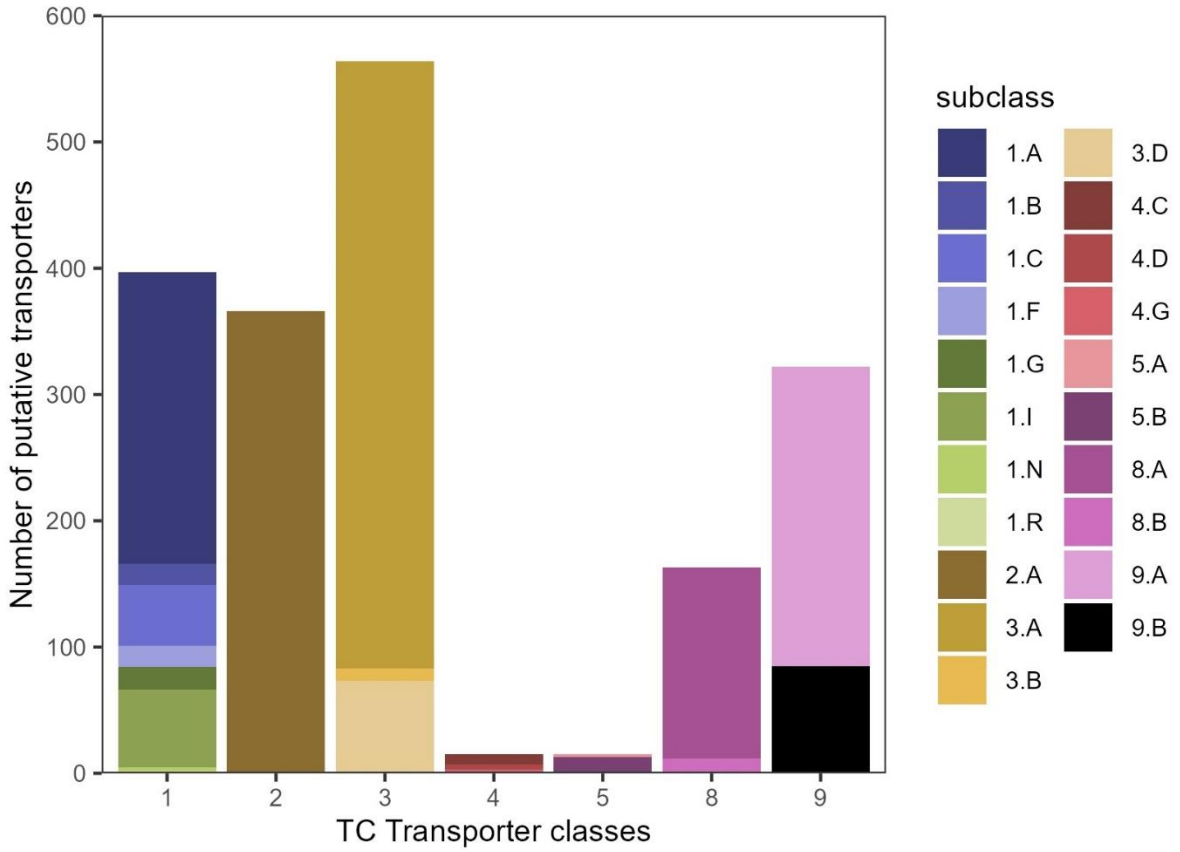


Figure 21: Transport systems present in P.brassicae genome.
There are seven categories (TC classes 1-5,8,9) recognised in TCDB. 1) channels/pores, 2) electrochemical potential-driven transporters known as secondary carriers, 3) primary active transporters, 4) group translocators, 5) transmembrane electron carriers, 8) auxiliary transport proteins and 9) putative or incompletely characterized transport systems contain proteins that might be involved in transport but are less understood.

Next, the transporters were assigned to TCDB families. *P.brassicae* transporters were grouped into 193 transporter families (Table 6, Table S2).

Table 6: Top 50 *P.brassicae* transporter families with the highest number of transporters

class	subclass	family	number	class	subclass	family	number
3	3.A	3.A.20	89	9	9.A	9.A.8	16
9	9.A	9.A.15	87	1	1.C	1.C.57	15
3	3.A	3.A.1	71	8	8.A	8.A.114	15
3	3.A	3.A.18	65	1	1.G	1.G.3	14
1	1.I	1.I.1	62	9	9.A	9.A.1	14
8	8.A	8.A.104	61	9	9.B	9.B.104	14
1	1.A	1.A.1	56	1	1.A	1.A.23	13
2	2.A	2.A.133	55	1	1.A	1.A.76	13
3	3.D	3.D.1	48	1	1.B	1.B.12	13
2	2.A	2.A.1	46	2	2.A	2.A.5	13
9	9.A	9.A.60	45	3	3.A	3.A.16	13
3	3.A	3.A.22	38	3	3.A	3.A.23	13
2	2.A	2.A.29	35	1	1.A	1.A.12	12
9	9.A	9.A.14	35	2	2.A	2.A.49	12
2	2.A	2.A.7	34	8	8.B	8.B.9	12
3	3.A	3.A.8	34	1	1.A	1.A.5	11
3	3.A	3.A.5	33	2	2.A	2.A.48	11
3	3.A	3.A.3	30	1	1.A	1.A.10	10
3	3.A	3.A.11	29	1	1.A	1.A.30	10
1	1.A	1.A.4	28	1	1.A	1.A.53	10
2	2.A	2.A.6	28	2	2.A	2.A.21	10
8	8.A	8.A.30	23	2	2.A	2.A.4	10
3	3.A	3.A.2	18	3	3.B	3.B.1	10
9	9.A	9.A.50	18	3	3.D	3.D.4	10
1	1.F	1.F.1	17	1	1.A	1.A.3	9

The top 50 transporter families, ranked by the number of transporters within each family, are predominantly represented by members of classes 1-3 (Figure 21). 10 families of the top 50 families were in the subclass 1.A which consists of transporters that facilitate the passive movement of solutes through a pore or channel embedded in the cell membrane, without requiring energy or the aid of a carrier molecule. However, in the pore/channels category, family 1.I.1, the nuclear pore complex (NPC) family had the greatest number of transporters. *P.brassicae* transporters in subclass 2.A that contains transporters which function as secondary carrier facilitators belong to 45 families (Table S2). Most transporters belonged to The RDD Na+(Li+)(K+)/H+ Antiporter (RDD) Family (2.A.133) which is essential in response to saline-alkaline stress. The major facilitator superfamily (MFS, 2.A.1) which catalyse uniport, solute: cation (H+, but seldom Na+) symport and/or solute: H+ or solute: solute antiport was the second largest family with 46 transporters. Primary active transporters accounted for 30% (565) of all the transporters (Figure 21). Majority of these transporters are primary active transporters that derive their transport energy from hydrolysis of a phosphate bond of inorganic pyrophosphate,

ATP or another nucleotide triphosphate (Table S2) with 11 families in this subgroup making it to the top 50 families (Table 6). 61 *P.brassicae* transporter families representing 26% of all *P.brassicae* transporters were in the poorly defined class 8 and 9 which contains proteins that are believed to be involved in transport but there is not sufficient evidence to assign them in the well-defined classes (Table S2). The 5'-AMP-activated protein kinase (AMPK) Family (8.A.104) which plays an essential role in regulation of cellular metabolism and the autophagy-related phagophore-formation transporter (APT) family, which is involved in the formation of phagophore an organelle responsible for cellular degradation and nutrient storage had the most number of transporters.

3.3.2. *P.brassicae* transport system in relation to other obligate biotrophic pathogens

While the TCDB lacks explicit information on transport subclasses exclusive to prokaryotes or eukaryotes, a comparative analysis with *Puccinia triticina-pt76* and *Ustilago maydis*, two prominent obligate biotrophic plant pathogens, was conducted to assess the completeness of *P.brassicae* transport systems (Figure 22). The results showed that *P.brassicae* has all the transport systems found in the two pathogens except subclass for 3.E, 4.B and 9.C. *P.brassicae* however, possessed seven unique transport systems absent in the other two pathogens. With 1845 transporter proteins, *P.brassicae* possesses significantly more transporters than *Puccinia triticina-pt76* (1484) and *Ustilago maydis* (567). While the presence of unique transport systems could be attributed to *P.brassicae* 's unique life cycle, a more striking observation is the disproportionately high percentage of transporter proteins within its genome (18%) compared to 4.65% in *Puccinia triticina-pt76* and 8.36% in *Ustilago maydis*. This disproportionate distribution was unexpected given *P.brassicae* genome (25.1 Mb) is slightly larger than *Ustilago maydis* (19.66 Mb) and significantly smaller compared to *Puccinia triticina-pt76* (253.66 Mb).

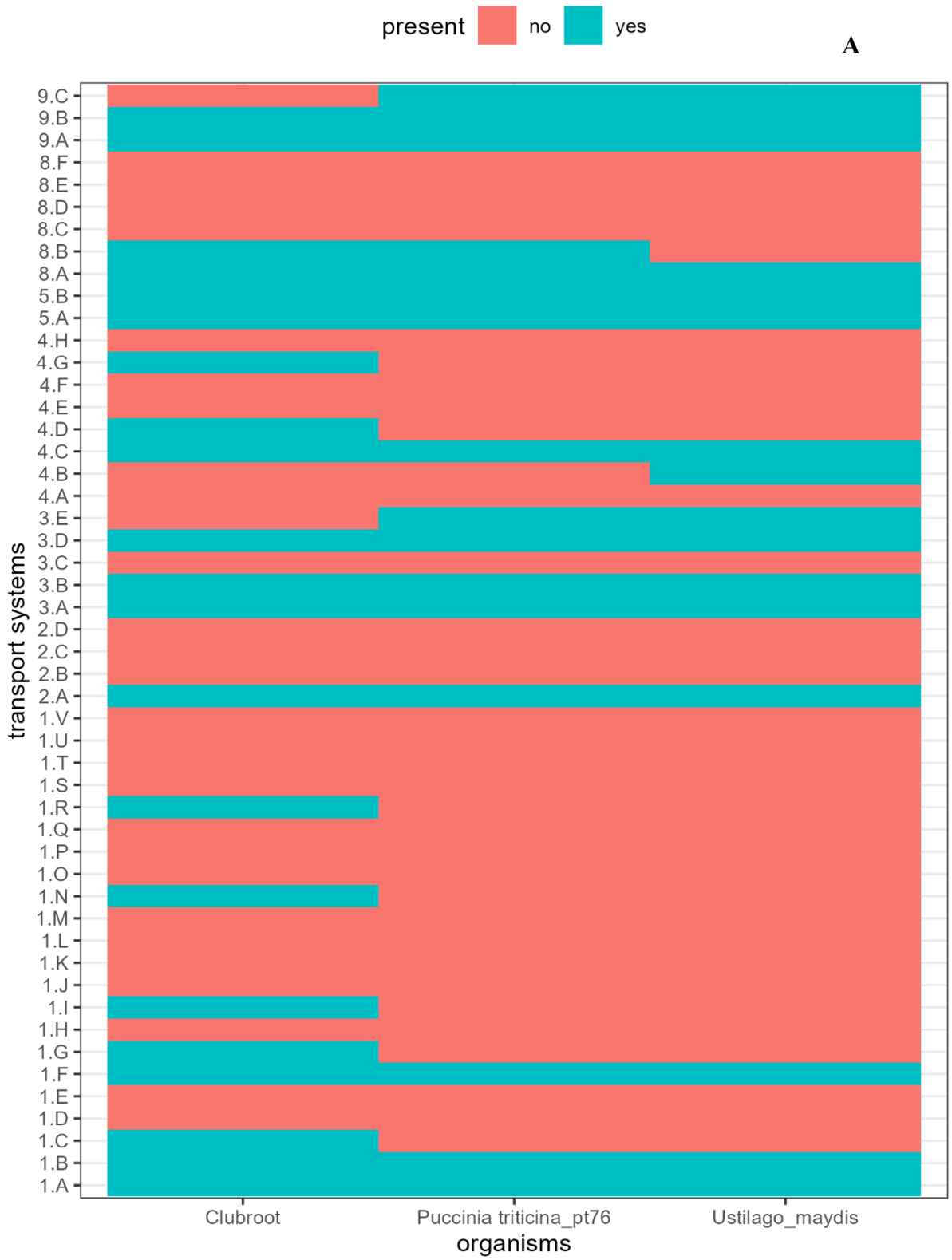

present
no
yes
A
transport systems
9.C
9.B
9.A
8.F
8.E
8.D
8.C
8.B
8.A
5.B
5.A
4.H
4.G
4.F
4.E
4.D
4.C
4.B
4.A
3.E
3.D
3.C
3.B
3.A
2.D
2.C
2.B
2.A
1.V
1.U
1.T
1.S
1.R
1.Q
1.P
1.O
1.N
1.M
1.L
1.K
1.J
1.I
1.H
1.G
1.F
1.E
1.D
1.C
1.B
1.A
Clubroot
Puccinia triticina_pt76
Ustilago_maydis
organisms

Figure 22: Number of transporters and transport systems present in P.brassicae, Puccinia triticina-pt76, and Ustilago maydis.
A) shows the transport systems that are present or missing in each of the three pathogens. B) shows the number of the total number of transporters found the genomes of the three pathogens. The number of transporters for Puccinia triticina-pt76 and Ustilago maydis was retrieved from mycocosm database.

3.3.3. Expression analysis of genes encoding putative transporters in *P.brassicae*

A comprehensive analysis of the *P.brassicae* genome revealed a substantial number of transporters and their families (Figure 21 and Table S2). This finding raised the question whether these putative transporters are expressed during different stages of *P.brassicae* development.

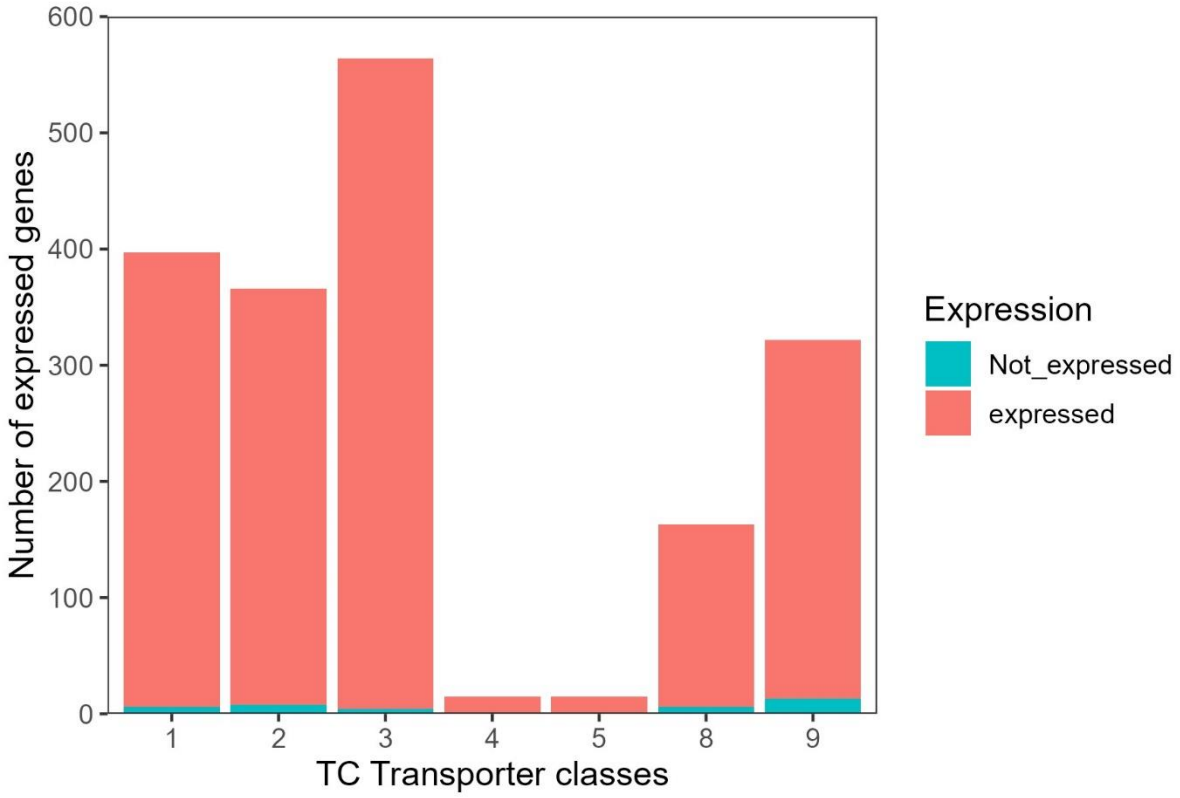


Figure 23: Normalised expression of P.brassicae transporters.
Transporter transcript counts from 14 RNAseq projects were normalised using DESEQ2 R package. Genes with thirty or less reads across the samples were removed before normalisation of the counts. The normalised read counts are expressed as fragments per kilobase of transcripts per million mapped reads (FPKM). Transporters were considered as expressed if they have ≥ 10 FPKM reads.

The results showed that 98% of putative transporter genes identified in the *P.brassicae* genome exhibited expression, indicating their likely role in encoding functional proteins (Figure 23). Notably, lowering the expression threshold to max FPKM to two or less reads increased the expressed genes to 99.7% (Figure S2), suggesting functional relevance even for genes with lower transcript abundance. Manual inspection of these lower-expressed genes using IGV (Figure 24) further corroborated their expression.

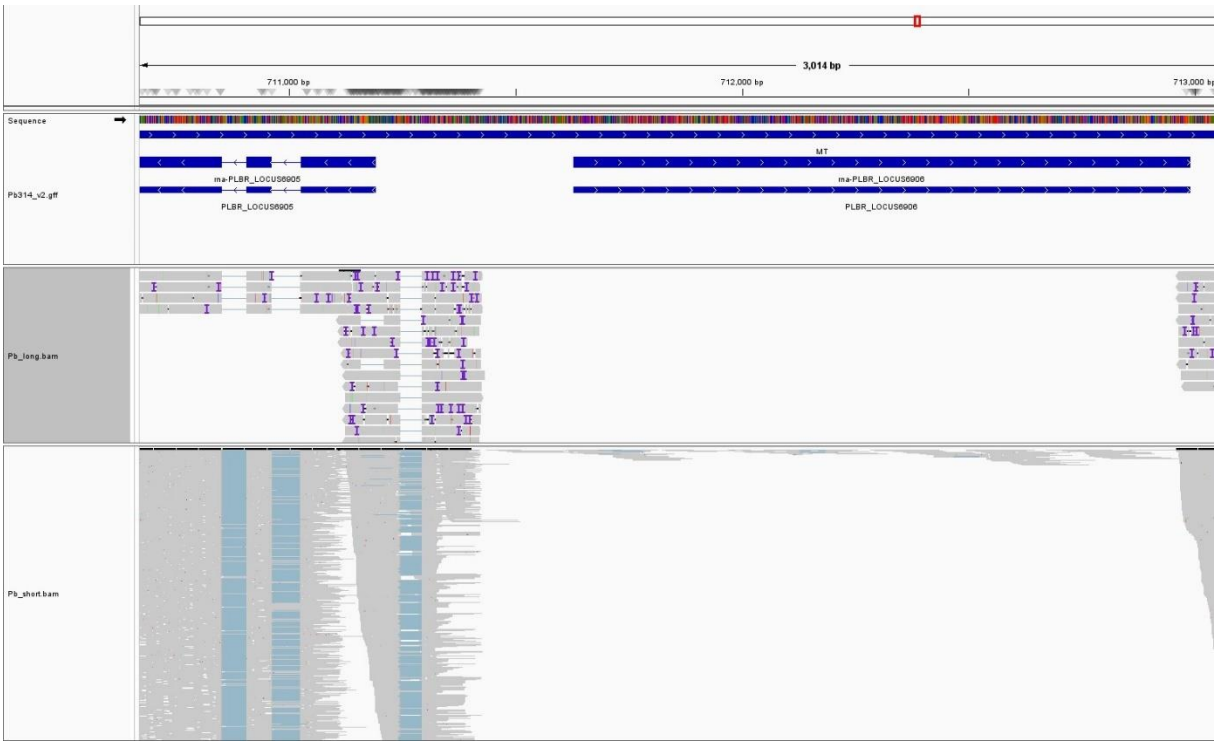


Figure 24: Manual inspection of expression for transporters with low transcript abundance. For example, gene PLBR_LOCUS6906 was initially deemed unexpressed when applying an FPKM cutoff of ≥ 10. However, manual inspection using IGV confirmed its expression. This confirmation was based on evidence from both short read RNAseq data and full-length/long read data, both of which were mapped to the P.brassicae reference genome.

3.3.4. Putative substrates of *P.brassicae* transporters

In TCDB, each transport system is further ranked progressively into more specific categories. This gives the opportunity to define *P.brassicae* transporter proteins at finer resolution up to the substrate level. *P.brassicae* transporter protein domains were compared with PFAM domains associated with cellular transport deposited in TCDB and all the putative substrates associated to each PFAM protein domain, and their families were retrieved from TCDB. This TCDB comparison returned multiple hits of putative substrates and families for each putative *P.brassicae* transporter (Table S3). The duplicated gene ids were removed, and the remaining transporters were manually assigned putative substrates. The putative *P.brassicae* transporters can transport wide range of substrates (Table 6). This analysis demonstrates the dependence of *P.brassicae* on its host for most of the nutrients it needs to sustain its growth.

Table 7: Putative substrates of *P.brassicae* transporters

Category	Substrate
Sugars	hexose, sucrose, glucose, galactose, N,N'-diacetylchitobiose, D-xylose, aldose, mannose, Saturated and unsaturated oligogalacturonide, N-Acetylglucosamine, N,N'-diacetyl chitobiose, Capsular polysaccharides, Trehalose, ribose,cellobiose, Maltooligosaccharide
Nucleosides and Nucleobases	adenine nucleoside, general nucleosides, nucleoside, nucleobases, purine nucleosides, pyrimidine nucleoside, queuosine
Nucleotides	UDP-sugar, GDP-sugars
Amino acids and Amino acid derivatives	General amino acids, amines, valine, L-glutamate, choline, proline, glycine-betaine, aromatic amino acid, Leucine, isoleucine, alanine, serine, glycine, histidine, Cystine, glutamate, methionine, basic amino acids
Vitamins	Folate, thiamine, Biotin, biopterin, thiamine precursor
Organic compounds	Organic cation, monocarboxylic acid, acetate, succinate Bicarbonate, glutamate, organic anions, 1-hydroxy-2-naphthoate, organic solutes, acetyl-CoA, butyrate, oxalate, formate carbonic acid, pyruvate
Lipids and fatty acids	Lysophospholipid, fatty acid, lipid, cholesterol, oleic acid, multiple long chain fatty acids, lipid intermediates, Glycerol
Nitrogen compounds	Ammonium,nitrate, nitrite, ammonia, urea
Phosphate compounds	inorganic phosphate, thiamine pyrophosphate, phosphate, organo-phosphate
others	Solutes, sulfate, metabolites, methylthiol, osmolites, metal-thiol Haem, anthocyanin, S-adenosylmethionine/S-adenosylhomocysteine(SAM/SAH), ATP, ADP, taurine, water, glutathione

3.3.5. *P.brassicae* nutrients transporters and their expression in different *P.brassicae* developmental stages

Through genome analysis, a comprehensive catalogue identified 1845 genes in *P.brassicae* predicted to encode transporters. From this catalogue, transporters potentially involved in substrate import and export were selected. The selection of substrate transporters was based on their predicted substrate. A total of 402 putative transporters were identified through this process (Figure 25). A significantly high number of these predicted transporters were found to transport proteins. The protein transporters are likely to be involved in the movement of proteins

between *P.brassicae* organelles rather than being involved in the trafficking of proteins across the *P.brassicae* plasmodia membrane.

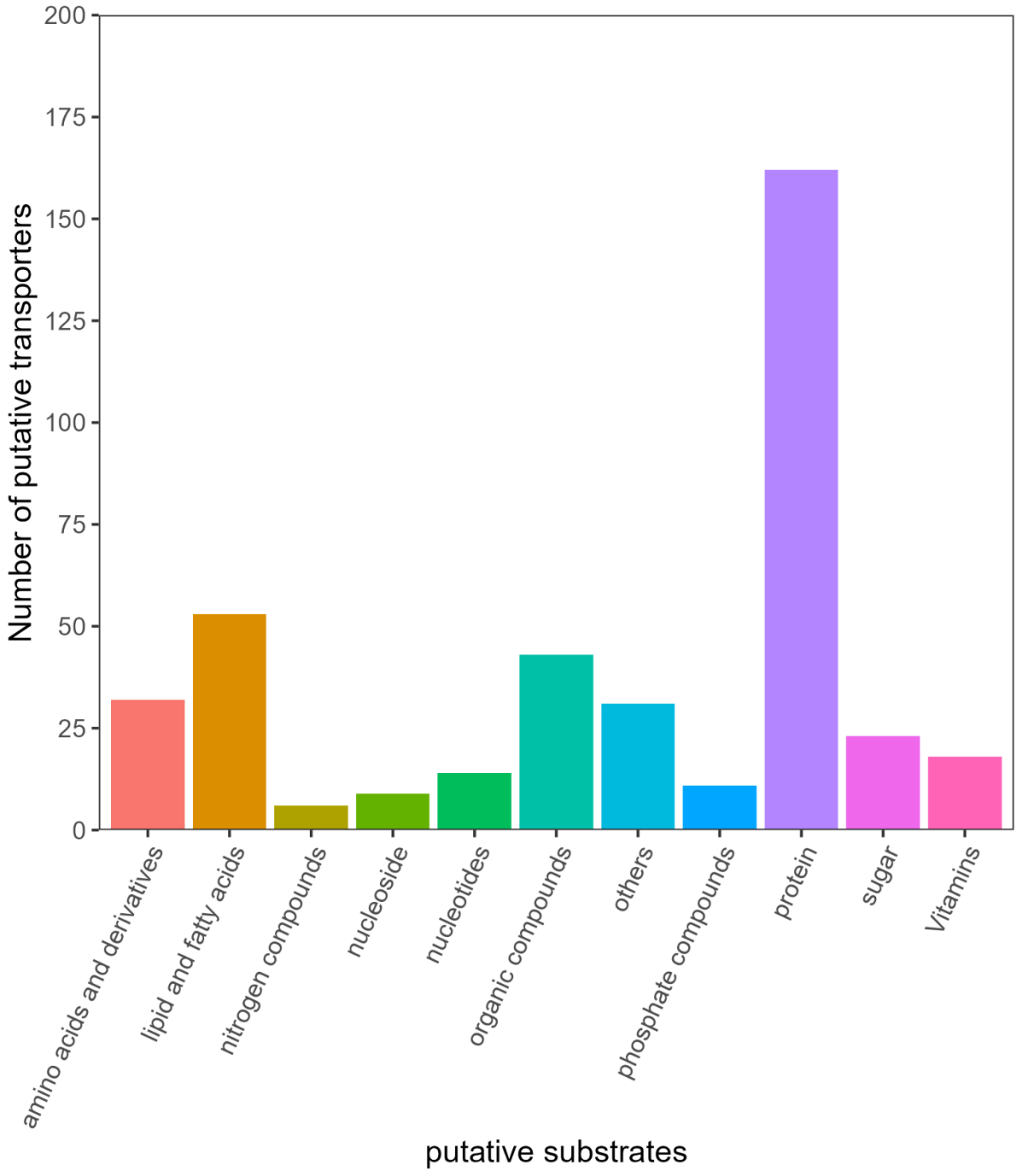


Figure 25:Putative P.brassicae substrate transporters.
Substrates were manually assigned to transporters based on the transporter family and putative substrates that members of the family in TCDB are known to or putatively transport.

To investigate the potential roles of putative substrate transporters in *P.brassicae* -host interactions, scRNAseq and whole tissue RNAseq transcriptomic approaches were employed. Sufficient protoplasts from both control and infected plants were successfully isolated. However, while cells from control samples were successfully loaded onto the 10X cell sorting machine and prepared for scRNAseq library preparation, the machine was repeatedly clogged by infected samples. The clogging of the 10X cell sorting machine by infected samples could have been caused by enlarged *P.brassicae* -infected cells and starch present in the sample. Publicly available whole tissue RNAseq datasets therefore were used to study their expression patterns by comparing normalised transcript levels across different developmental stages. In general, two different expression patterns for substrate transporters were observed across these stages. Transporters were exclusively expressed either in the cortical invasion stage or spore stages apart from a few transporters overlapping between two stages with weak expression levels (Figure 26). This expression trend was maintained even when expression of transporters for each substrate category as mentioned in table 7 above were analysed separately (Figure S3)

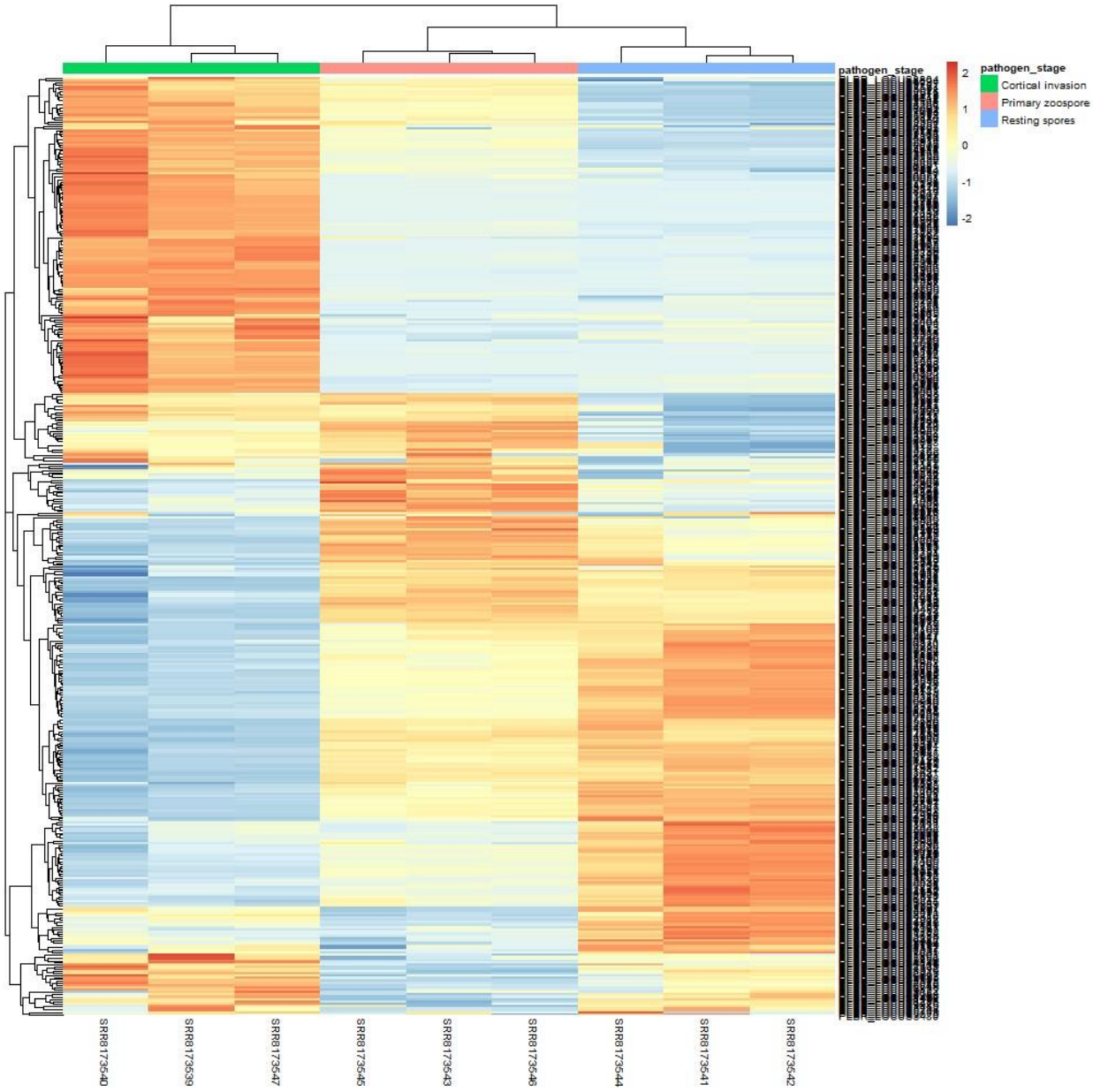


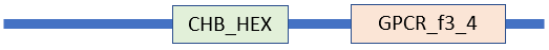
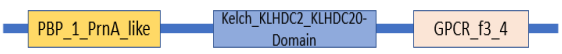
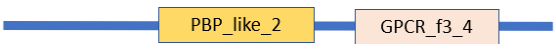
Figure 26: Expression of putative substrate transporters in different P.brassicae developmental stages .
The normalised transporter transcript counts from PRJNA331021 RNAseq project were used to compare transcript abundances of transporters in different stages.

3.3.6. Unveiling unusual transporters in the *P.brassicae* genome

Given the previous identification of eight PRO-EU transporters within the *P.brassicae* genome (Rolfe *et al.*, 2016), an investigation was conducted to find if *P.brassicae* possessed additional PRO-EU transporters, unique transporters, or transporters less commonly found in other organisms. These transporters are referred to as unusual transporters because they are not common in eukaryotes or are found in few eukaryotic organisms. PRO-EU transporters were identified by filtering transporters that contain bacteria substrate proteins(SBP) PFAM domains

(Table 5). 26 PRO-EU transporters that had bacteria SBP domains PF00497, PF01094, PF01547, PF02608, and PF04069 were found in the *P.brassicae* genome. Further inspection of these PFAM domains on InterProScan revealed that the PFAM domain PF01094 was widely distributed in eukaryotes and thus this PFAM domain and associated transporters were removed leaving a total of 12 PRO-EU transporters (Table 8). The remaining 383 PFAM domains were individually analysed for taxonomic distribution using InterProScan to identify transporters that were not common in eukaryotes or those restricted to specific organism groups. 56 PFAM domains and 173 *P.brassicae* transporters were found to be not common in eukaryotes or restricted to few organisms (Figure 27). These transporters belong to 53 transporter families and are predicted to transport various substrates (Table S4).

Table 8: Protein structure of putative PRO-EU transporters

PFAM Domain	SBP Family	Gene ID	Structure
PF00497	SBP_bac_3	PLBR_LOCUS3723	
		PLBR_LOCUS9199	
PF01547	SBP_bac_1	PLBR_LOCUS4174	
		PLBR_LOCUS7736	
		PLBR_LOCUS6651	
		PLBR_LOCUS2502	
		PLBR_LOCUS4177	
		PLBR_LOCUS8824	
PF02608	Bmp	PLBR_LOCUS1695	
		PLBR_LOCUS1720	
		PLBR_LOCUS1694	
PF04069	OpuAC	PLBR_LOCUS5658	

Note: Both 7TM3 and GPCR proteins have 7 transmembrane domains and they are very structurally related thus their annotations is unclear.

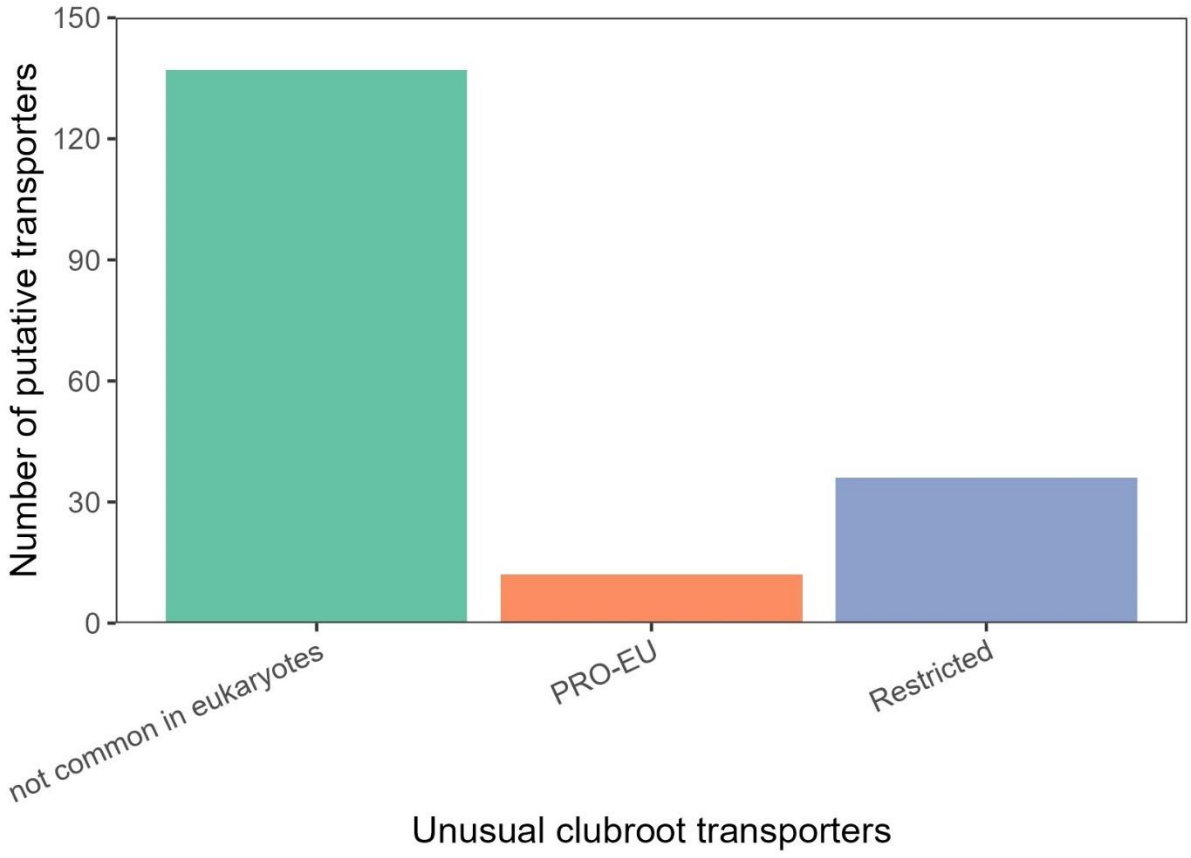


Figure 27: Unusual P.brassicae transporters.
PRO-EU transporters were identified by filtering P.brassicae transporters with bacterial SBP PFAM domain. Transporters that are restricted to few organisms and those that are not common in eukaryotes were identified based on taxonomic distribution of organisms using InterProScan.

3.3.7. Refinement of PRO-EU and SWEET transporters gene models

The analysis of protein families of *P.brassicae* PRO-EU with InterProScan revealed that the PRO-EU with PF00497 lacked the 7TM_3 domain. Because the *P.brassicae* gene models are incomplete, the correctness of PRO-EU gene models were analysed by comparing the gene models from three different *P.brassicae* genome annotations. Gene models that were supported either by short RNA or full length/long read mRNA reads were accepted as correct gene models. Six gene models were correct and another six were incorrect. Gene models of PLBR_LOCUS3723 and PLBR_LOCUS9199 which have the bacteria SBP domain PF00497 were both correct (Figure 28).

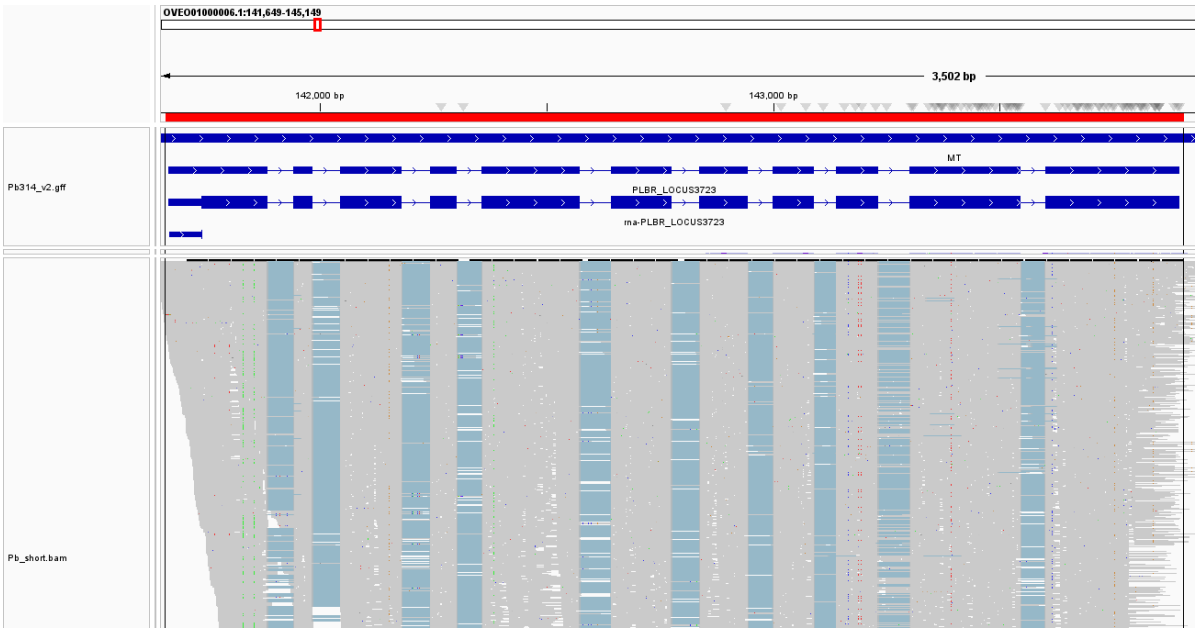


Figure 28: Transcript evidence for PLBR_LOCUS3723 gene model.
PLBR_LOCUS3723 gene model is supported by short RNAseq reads mapped to P.brassicae reference genome Pb314v2.

3.3.8. Transport analysis of sweet sugar transporters

To study the function of *P.brassicae* sweet transporter PLBR_Locus1241, PLBR_Locus1241 was heterologously expressed sucrose mutant yeast strain CSY4000. The yeast cells were able to grow on plates containing sucrose as the sole source of carbon. These results support the finding of SWEET transporters in the *P.brassicae* genome and confirm that PLBR_Locus1241 is a sucrose transporter as it was predicted by the TCDB database.

Figure 29: Growth of the yeast mutant strain CSY4000 expressing P.brassicae SWEET transporter.
The complemented strain was grown on SD (-Ura) solid media supplemented with 2% maltose and 2% sucrose at 300C for 4 days. PLBR_Locus1241 transport assay was done by Dr Gregory J Fowler

3.4. Discussion

Biotrophs express numerous transporter genes that enable diverse nutrients to be acquired from the host (Wahl et al. 2010; Chen et al. 2010). *P. brassicae* has lost the ability to synthesise many essential macromolecules (e.g. vitamins, amino acids) as they are biochemically expensive to make and acquiring them, 'ready-made' is energetically more favourable (Rolfe et al. 2016). In this work, a comprehensive bioinformatic analysis of *P.brassicae* genome was conducted to get an overview of transport systems, transporter families and putative transporter proteins present in *P.brassicae* . The transport systems present in the pathogen and the role of transporters in relation to nutrient salvage, manipulation of host sugar sensing and defence signalling mechanisms are discussed.

3.4.1. Transport systems and transporter families in *P.brassicae* .

Transporters in the TCDB are manually curated and organised in groups using a tier system based on their structures and function. There are seven categories (TC classes 1-5,8,9), the first five classes (TC 1-5) contain proteins whose role in transport is well defined and the last two (TC classes 8 and 9) being less clearly defined. The TCDB further ranks the transporters into increasingly specific groups: class, subclass, family or superfamily, subfamily, and the substrate they transport. By using the TCDB classification, all the 1845 putative *P.brassicae* transporters were assigned to one of the seven transport systems in the TCDB (Figure 21). 74% of the putative transporters were grouped into the well-defined classes (TC 1-5) and 26 % into the less defined categories.

The TC class 1 belongs to the channel/pores transporters. This class consists of transporters with an α-helical or β-strand-type transmembrane spanner that forms a channel. They transport substrates across the plasma membrane by means of facilitated diffusion but there is no evidence of carrier mediated mechanism (Saier et al., 2021). Channels or pores control the movement of ions across the plasma and organelle membranes whereby they regulate different cellular functions such as membrane potential, signalling, excitability, volume, and death. In host-pathogen interactions, channels are implicated in nutrient uptake, viability, infectivity, and replication (Jimenez & Mesones, 2022). *P.brassicae* possess eight subclasses of channels and pore transporters. This is significantly higher when compared to other plant biotrophic pathogens *Puccinia triticina-pt76* (3) and *Ustilago maydis* (3) (Figure 22). This difference in the number of subclasses of channel transporters could be related with the lifestyle of these plant pathogens where *Puccinia triticina-pt76* and *Ustilago maydis* live in the intercellular

spaces of host plants and obtain their nutrients via haustoria whereas *P.brassicae* resides and obtains nutrients inside the host cells. *P.brassicae* channels could help the pathogen to regulate intracellular pH by providing buffering capacity for proton mobilisation as observed in bacteria and speculated in mammalian pathogenic protists (Jimenez & Mesones, 2022). *P.brassicae* channels can also help the pathogen to increase host cell permeability to nutrients by targeting *P.brassicae* channel proteins to the host cell membrane a mechanism that is employed by pathogenic human protist *Plasmodium falciparum* (Desai *et al.*, 2000; Nguitragool *et al.*, 2011).

TC class 2 consists of the electrochemical potential-driven transporters. TC 2 transporters, also known as secondary carrier facilitators, are a type of transporter that utilize a carrier molecule to facilitate the movement of molecules across the membrane. These transporters can function in three different ways: uniport (movement of a single molecule in one direction), antiport (exchange of one molecule for another in opposite directions), or symport (movement of two or more molecules in the same direction) (Figure 7). These transporters are essential for the movement of substrates from the extracellular compartment into the cell for their cellular metabolism (Saier et al., 2021). Transporters of subclass 2.A in pathogenic protists *P. falciparum* and *P. berghei* have been described as essential or non-mutable highlighting their importance for the normal growth and survival of these pathogens (Martin, 2020). These transporters can function within these pathogens due to the formation of a new membrane, the parasitophorous vacuole membrane. This membrane arises during host cell invasion and surrounds the growing parasite. The parasitophorous vacuolar space creates ion gradient that is used by the pathogen plasma membrane-localised carrier transporters for solute trafficking. It is not known if *P.brassicae* does form the 'parasitophorous vacuole membrane' during the invasion process, but if it does, transporters of this subclass could be important to *P.brassicae* because the uniporters would help the pathogen to transport nutrients against their concentration gradient as there would be more nutrients in the host cell cytosol than in the *P.brassicae* .

TC 3 consists of primary active transporters that use chemical, electrical or solar energy to actively move a solute against its concentration gradient (Saier et al., 2021). *P.brassicae* has three subclasses in five that are present in this class of transporters. Transporters in TC 3 could help the pathogen to maintain its membrane potential, efflux of metabolites and the uptake of nutrients such as sugars. TC class 4 belongs to group translocators. Transporters in this class move and modify molecules at the same time. In this transport system, the molecule cannot be effectively transported without being changed. In other cases, the transport and modification are loosely coupled, meaning the molecule can be moved without being changed, but it's more

efficient if they happen together (Saier et al., 2021). Class 4 *P.brassicae* transporters are grouped into 3 subclasses as compared to one subclass in *P. triticina-pt76* and *U. maydis* (Figure 22). Subclass 4.C contains transporters that are involved in fatty acid transport. *P.brassicae* contains abundant lipid droplets throughout its life cycle, but their origin remains unclear. Bi et al. (2016) reported that *P.brassicae* can synthesize lipids, however, the evidence supporting this claim is limited. Schwelm et al. (2015) demonstrated that *P.brassicae* lacks fatty acid synthase, a key enzyme complex essential for de novo fatty acid synthesis (Remize *et al.*, 2021). This suggests that *P.brassicae* cannot synthesize fatty acids *de novo*. Instead, it likely acquires short-chain carbon molecules from the host and utilizes the microsomal elongase pathway an alternative fatty acid synthesis pathway employed by pathogenic trypanosomes (Lee *et al.*, 2006) to elongate and modify these molecules into longer fatty acid chains. The presence of fatty acid transporters in this study further supports the view that *P.brassicae* can transport and metabolise short fatty acid from the host. Unlike *P. triticina-pt76* and *U. maydis* (Figure 22), *P.brassicae* has transporters in the subclass 4.D and 4.G which are involved in the synthesis and transport of extracellular polysaccharides and the cleavage of transmembrane domains of type I membrane proteins (Saier *et al.*, 2021).

TC class 5 is made of transmembrane electron carriers. This group contains transporters which directly impacts the cellular energetics by influencing transmembrane electron flow systems. These systems move electrons from donor molecules on one side of the membrane to acceptor molecules on the opposite side. This directed flow of electrons directly impacts the magnitude of the membrane potential either contributing to or subtracting from it depending on the direction. Consequently, these systems play a critical role in cellular energy dynamics (Saier et al., 2021). *P.brassicae* lipid bodies are rich in arachidonic acid which is a signalling molecule that is found in animals and some pathogens of eukaryotic SAR clade (Savchenko *et al.*, 2010). The presence of arachidonic acid affects the activity of transporters in the TC 5 subclass 5.B there by influencing the direction of proton flux (Meves, 2008; Saier *et al.*, 2021). Presence of transporters of this class could give *P.brassicae* an advantage to control its cellular pH and that of the host cell.

TC 8 consists of accessory factors involved in transport. This class of transporters is unique because some members work alongside or even combine with other known transporters. Transporters in this class are involved in different cellular processes that affect development and metabolism of organisms. Auxiliary transporters are required for the export of peptides and functioning of some ABC efflux transporters in gram-positive bacteria (Lorca *et al.*, 2007),

transport of auxin in plants (Bouchard *et al.*, 2006), and switching of substrates by Glut1from glucose to L-dehydroascorbic acid (DHA) in humans (Montel-Hagen *et al.*, 2008). The presence of this class of transporters may help *P.brassicae* to finetune its development and metabolism in relation to the intracellular host cell environment. For instance, *P.brassicae* has 62 transporters in the family 8.A.104 that is involved in regulation of cellular energy metabolism and glucose homeostasis (Mcgee *et al.*, 2008; Egan *et al.*, 2011). TC class 9 is composed of incompletely characterized transport systems. This class consists of transporters with unknown functions that awaits further classification. How transporters of this class help *P.brassicae* is unclear because there is no information on how transporters in this category function.

While both prokaryotes and eukaryotes share the seven major transporter classes, the presence or absence of transporter classes, subclasses and families varies between organisms and species (Lorca *et al.*, 2007) and the environment and niches where the organisms live (Russum *et al.*, 2021; Zhang *et al.*, 2024). The analysis of transport capabilities of 11 gram-positive bacteria showed that some of the bacteria did not have all the seven transporter classes and the number of subclasses, families and transporters in same families varied between the bacteria (Lorca *et al.*, 2007). Studies on transport systems of *Plasmodium falciparum* revealed that this pathogenic protist lacks transporters in the transporter class 4 and 5 and has a small number of transporter subclasses (Wunderlich, 2022). *P.brassicae* has all the seven transporter classes, but not all the transporter subclasses or families present in the TCDB are present in *P.brassicae* . Comparison of *P.brassicae* transport system with those of *P. triticina-pt76* and *U. maydis* (Figure 22) shows that *P.brassicae* has the greatest number of transporter subclasses and the greatest number of transporters. However, it should be noted that the *P.brassicae* genome used in this study includes mitochondrion genome whereas the genomes of *P. triticina-pt76* and *U. maydis* do not. Since there is no consensus about the number of transporter subclasses or families that should be present in prokaryotes or eukaryotic organisms, the *P.brassicae* transport systems cannot be judged as complete or incomplete.

3.4.2. *P.brassicae* genome encodes and expresses large number of transporters

P.brassicae genome contains 1845 predicted transporters , a number that is higher than 1484 in *P. triticina-pt76* and 567 *U. maydis.* The number of transporters in *P.brassicae* is exceptionally high when compared to *Plasmodium falciparum* an intracellular pathogenic protist which has 200 transporters (Wunderlich, 2022) despite the genomes of the two protist pathogens being almost the same size (Gardner *et al.*, 2002). Almost all the predicated transporters were

expressed (Figure 22) further indicating that the putative transporters may be functional. The large difference in transporter repertoire of these pathogens may be related to different environments they live in and their life cycles. The high number of transporters in *P.brassicae* is also supported by the compact nature of its genome characterised by overlapping of genes (Rolfe *et al.*, 2016) which could in turn reflect a higher number of genes per Mb of genomic DNA. *P.brassicae* transporter genes have a genomic share of 18% of total *P.brassicae* genes. This number is exceptionally high compared to 4.5% in *P. triticina-pt76*, 8.36% in *U. maydis* and 2.5% in *Plasmodium falciparum* (Martin, 2020). The proportion of transport proteins in *P.brassicae* remains high even when compared to multicellular eukaryotes such as *Arabidopsis thaliana* 4.90%, *Homo sapiens* 3.87% and the unicellular prokaryote *E. coli* whose transport proteins account for 14.20% of all genes (Elbourne, LD *et al.*, 2017). Nonetheless, it should be noted that *P.brassicae* genome is poorly annotated with gene models missing, joined together, split or with varying start and stop sites among different genomes. This could affect the total number of genes present in the genome and ultimately the number of transport proteins. For example, in their *P.brassicae* genome annotation Rolfe et al., (2016) found a novel class of transporters designated PRO-EU in this study which combines prokaryotic and eukaryotic transporter features, but these genes were either not annotated or annotated as different genes in the *P.brassicae* Pbe3_h15 genome (Schwelm *et al.*, 2015). Another reason of high transporter numbers in *P.brassicae* may arise from the multicomponent nature of transporters. The presence of protein motif associated with transport in a protein does not necessarily mean that that protein is a transporter. A *Plasmodium falciparum* protein annotated as a putative ABC transporter because of the presence of ATP-binding motif was found to encode a homolog of the *S. cerevisiae* GCN20 ATPase, which is part of a protein complex that controls protein synthesis in amino acid deficient cells (Martin *et al.*, 2005). Transporters also evolve to have a dual function as transporter and sensor or sensors. Such transporters are termed as transceptors (Thevelein & Voordeckers, 2009). Non-transporting transceptors act as nutrient sensors and induce the expression of their homologous proteins which transports the nutrients across the membrane (Thevelein & Voordeckers, 2009). It is likely that some of PRO-EU transporters may not be transporters but nutrient sensors (Cerna-Vargas *et al.*, 2023).These mis-annotations of proteins as transporters due to part of the protein being associated with transport could inflate the number of putative *P.brassicae* transport proteins.

3.4.3. *P.brassicae* genome encodes unusual and transporters of unknown transporter classification

The presence of unusual transporters in the *P.brassicae* genome first came to light in 2016 during the *P.brassicae* genome sequencing project (Rolfe *et al.*, 2016). Rolfe et al (2016) found that 7 *P.brassicae* transport proteins with eukaryotic transmembrane domains matched the substrate binding domains (SBP) of prokaryote transporters. Following on this study, we identified 12 PRO-EU transporters and 173 transporters that are not common in eukaryotes or restricted to few organisms in eukaryotic kingdoms (Figure 27, Table S4). The SBP domains are present in eukaryotic membrane proteins however, the SBP domains coupled with ABC transporter systems appear to be limited to prokaryotes (Seppala *et al.*, 2016). The presence of these PRO-EU in eukaryotes is restricted to those eukaryotes that inhabit oxygen deprived environments. PRO-EU have so far been found in anaerobic gut fungi (Seppala *et al.*, 2016) but an analysis of distribution of PRO-EU transporters in our lab has found PRO-EU to be present in 48 eukaryotic organisms. Small multidrug resistance (SMR) transporters of prokaryotic origin have also been found in early branching fungi (Seppala *et al.*, 2023). Given that *P.brassicae* lives inside the host cell, these transporters could give an advantage to *P.brassicae* plasmodia to compete for sugar with host cells. Seppala et al (2016) found that 60% of putative sugar transporters of the anaerobic gut fungi were PRO-EU transporters, however, their function remains unknown. They speculated that these PRO-EU could communicate with fungal transporters, sequester sugar produced by extracellular digestive enzymes of the fungi which could increase sugar concentration around the fungi thus increasing sugar uptake or they could direct fungal zoospores to nutrient sources by unknown mechanism. The malaria parasite *Plasmodium falciparum* also has 12 unusual transporters although whether these are PRO-EU is not known (Martin *et al.*, 2005). Presence of transporters that are not common in eukaryotes and those that are restricted to few organisms in eukaryotic kingdoms could help *P.brassicae* to adapt to its intracellular host cell environment as they have a wide range of substrates they can transport (Table S4). For instance, 15 of the 23 predicted *P.brassicae* sugar transporters belong to the unusual transporters. These unusual sugar transporters may have higher sugar affinities than the host counterparts thus allowing *P.brassicae* to outcompete the host cells for sugar uptake. Expression of high affinity sugar transporters in relation to host sugar transporters is a strategy used by plant biotrophic pathogens to successfully colonise and proliferate in the hosts. *Ustilago maydis* which causes corn smut expresses a novel sucrose transporter which has higher sucrose affinity than the maize sucrose transporters (Wahl *et al.*, 2010).

3.4.4. Role of *P.brassicae* transporters in nutrient acquisition

P.brassicae infection of the host is characterised by the formation of galls on the infected roots which provide *P.brassicae* a space for shelter and nutrient acquisition. Like other biotrophic pathogens, *P.brassicae* has lost the ability to synthesise important metabolites (Rolfe *et al.*, 2016). *P.brassicae* genome contains a repertoire of transport proteins (Figure 25) which are expressed in different pathogen developmental stages (Figure 26) that allows it to acquire different nutrients from its host. Carbohydrates are the main source of energy and carbon for many pathogens. Carbon acquisition is crucial for the development and proliferation of pathogens, providing energy and building blocks of cellular components and signalling (Goulet & Saville, 2017). *P. brassicae* redirects the host carbohydrates in the infected roots by stimulating phloem proliferation and recruiting host sugar transporters in the developing galls (Li *et al.*, 2018; Walerowski *et al.*, 2018; Zhang *et al.*, 2019). However, the form in which these carbohydrates are taken by *P.brassicae* is unclear. A recent study (Kong *et al.*, 2022) identified 17 putative sugar transporters in the *P.brassicae* genome. Eight of these transporters which exhibited hexose sugar transport were significantly expressed in the plasmodial and resting spore stages when a large amount of energy is needed by the pathogen. In this study, 23 putative sugar transporters that are expressed in all stages of the pathogen life cycle but with a varying expression level in different stages were identified (Figure S3A). Heterologous expression of PLBRA_LOCUS124, one of the two SWEET *P.brassicae* sugar transporters that were identified in this study, restored the growth of CSY4000 sucrose transporter yeast mutant strain on media containing only 2% sucrose as a carbon source (Figure 27). These findings indicate that *P. brassicae* uptake host carbohydrates as hexose sugars, but it also has the capacity to directly take up sucrose from host. However, if sucrose is taken up, there must be rapid metabolism of the acquired sugar to 'trap' it and allow further uptake. This could be hydrolysis by sucrolytic enzymes (e.g. invertase) or conversion to other forms. *P.brassicae* has an extensive trehalose metabolism (Brodmann et al. 2002) but pathways from sucrose to trehalose have not been established nor have potential sucrolytic enzymes been identified (Rolfe *et al.*, 2016).

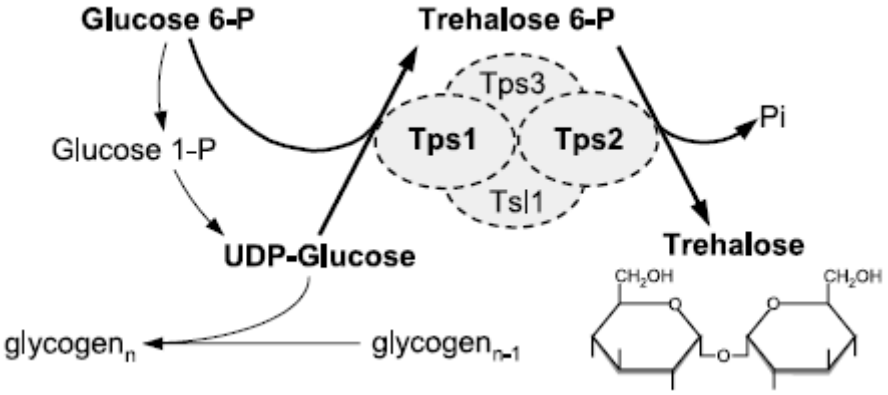


Figure 30. Yeast Trehalose biosynthesis pathway.
Tps enzymes Catalyse the reaction between Glucose 6-P and UDP-Glucose to make intermediate compound trehalose 6-P.Trehalose 6-P is further acted upon by another Tps enzyme to produce trehalose (Gancedo & Flores, 2004).

P.brassicae contains trehalose biosynthesis genes that are expressed in different pathogen developmental stages (Brodmann *et al.*, 2002; Schwelm *et al.*, 2015). This leads us to speculate that *P.brassicae* synthesise trehalose which is present in large amounts in resting spores to divert carbon away from the plant metabolism and create carbon sink in infected organs (Brodmann et al. 2002). *P.brassicae* genome has putative hexose transporters which supply the pathogen with glucose and has fructokinase (EC 2.7.1.4) and glucokinase (EC 2.7.1.2) which convert fructose and glucose into glucose 6-P, a starting compound for trehalose biosynthesis (Gancedo & Flores, 2004). UDP-glucose is another important component that must be available for trehalose biosynthesis (Gancedo & Flores, 2004). *P.brassicae* has 23 nucleobases, nucleosides and nucleotide sugar transporters that are expressed in different stages of the pathogen (Figure S3B). 11 of *P.brassicae* nucleosides and nucleotide transporters were expressed in the cortical invasion stage thus helping the pathogen to accumulate UDP-sugars that are prerequisite material for trehalose synthesis. Nucleoside and nucleotide transporters may provide *P.brassicae* an addition source of energy. Nucleosides have been confirmed to be metabolised to produce energy in cancer cell lines that are completely deprived of glucose (Skinner *et al.*, 2023). The researchers proposed that energy production from nucleosides occurs in three steps which involves production fructose-6-phoshate by the non-oxidative branch of Pentose phosphate pathway. *P.brassicae* has this pathway, and it was speculated by Schwelm et al.,(2015) as one of the major pathways for energy production in germinating spores and plasmodial (cortical invasion) stages of *P.brassicae* .

The *P.brassicae* genome has incomplete, or lacks, pathways that are necessary for the synthesis of some the amino acids and thus obtaining these nutrients from its host (Schwelm *et al.*, 2015;

Rolfe *et al.*, 2016). *P.brassicae* has 32 amino acids and amino acids derivative transporters with 9 of them exhibiting high level of expression in the cortical invasion stage (Figure S3C). PLBR_LOCUS555, a putative valine transporter, and PLBR_LOCUS5765, a putative transporter for leucine, leucine/isoleucine, and valine, were among the 9 most highly expressed amino acid transporters during the cortical invasion stage. PLBR_LOCUS8337 a putative histidine transporter was also expressed during cortical invasion and resting spores. This finding aligns with the absence of histidine and valine/leucine/isoleucine biosynthesis pathways in the *P.brassicae* genome (Rolfe *et al.*, 2016). *P.brassicae* has been reported to have reduced or lost genes essential for thiamine biosynthesis (Schwelm *et al.*, 2015; Rolfe *et al.*, 2016), suggesting that it cannot synthesize thiamine on its own. In this study, 18 putative thiamine transporters were identified, with varying expression profiles across different *P.brassicae* developmental stages. Notably, five transporters were exclusively expressed during the cortical invasion stage, while three were expressed only in the spore stage (Figure S3D). This pattern of expression supports the idea that *P.brassicae* acquires vitamins from its host. Nutrient uptake from the host by germinating spores has been observed in *Puccinia striiformis* f.sp. *tritici* (Pst) (Garnica *et al.*, 2013), further reinforcing the hypothesis that *P.brassicae* follows a similar strategy. One of the transporters expressed in the spore stage, PLBR_LOCUS8948, is predicted to transport thiamine precursors. However, it remains unclear why *P.brassicae* strongly expresses this transporter in the spore stage, as the pathogen is unable to synthesize thiamine due to the lack of several key enzymes in the biosynthesis pathway. Schwelm et al (2015) reported the absence of sulphur and nitrogen transporters whereas Rolfe et al (2016) found a sulfite oxidase but no nitrate reductase suggesting that the pathogen may acquire nitrogen in more complex forms. In this study, six putative nitrogen compound transporters with PLBR_LOCUS8716, PLBR_LOCUS2412 and PLBR_LOCUS9197 transporters exhibiting strong expression during the cortical invasion stage (Figure S3E) were identified. These transporters are predicted to transport nitrite, nitrate, ammonia and urea. Additionally, the analysis identified five putative sulphate and eleven phosphate compound transporters, with two of the sulphate and two of the phosphate transporters showing high expression in the cortical invasion stage respectively (Figure S3F, G). The presence of sulphate and phosphate transporters in our study agrees with the findings of Rolfe et al.,(2016) who also found phosphate transporters and sulphite oxidase indicating that the pathogen can acquire and metabolise sulphate and phosphate compounds from the host. However, the finding of organic and inorganic nitrogen compound transporters in this study contradicts with their findings which reported the absence of nitrogen reductase in *P.brassicae* genome.

P.brassicae resting spores, germinating spores and secondary plasmodia contains large amounts of lipid droplets. There are 53 putative fatty acid and lipid transporters in *P.brassicae* with fifteen showing high expression in the cortical invasion stage (Figure S3H) when c is thought to be actively taking up short chain fatty acids from the host (Schwelm *et al.*, 2015). There are two contradicting reports about fatty acid and lipid biosynthesis in *P.brassicae* . In one study, Schwelm et al.,(2015) reported that *P.brassicae* is unable to synthesise fatty acids *de novo* but can metabolise fatty acids of four or more carbons by elongating and modifying them via the microsomal elongase pathway. This hypothesis is more plausible because fatty acid synthase that is involved in *de novo* biosynthesis of lipids and fatty acid elongation (Jump, 2009) has so far not been identified (Schwelm *et al.*, 2015). Further support of this claim was provided by Bi et al.,(2016) who found long chain fatty acid synthetase (EC 6.2.1.3) in their study. However, Bi et al.,(2016) in their study reported that *P.brassicae* does not take up lipids from the host but rather synthesise them *de novo*. This claim is controversial because their experimental evidence supporting the claim is poor. Their hypothesis is based on Nile red staining of lipid droplets in plasmodia and spores. The Nile red staining can only detect the presence of fatty acid and lipid droplets but cannot give information on their biosynthesis. Furthermore, their analysis of proteins involved in *P.brassicae* fatty acid and lipid synthesis does not support *de novo* synthesis but rather fatty acid elongation pathway (Jump, 2009). The results from this study further supports the view of *P.brassicae* obtaining fatty acids from its host although the lengths of fatty acids transported by *P.brassicae* fatty acid and lipid transporters is unclear.

3.4.5. Role of transporters in manipulation of the host sugar sensing and defence signalling mechanisms

Trehalose and lipid droplets in secondary plasmodia may leak back into the host and regulate its development. Trehalose is usually found in plant symbiotic or pathogenic microorganisms though some plant species like Arabidopsis thaliana synthesises small amounts of the disaccharide (Brodmann et al. 2002). Research on the trehalose pathway in Arabidopsis revealed that the intermediate molecule trehalose-6-phosphate (T6P) of the trehalose pathway regulates how the plant uses carbohydrates and, consequently, its growth, by influencing glycolysis (Schluepmann *et al.*, 2003). In this study, the effect of different sugars on the growth of wild type and trehalose phosphate synthase (TPS) expressors (plants engineered to produce more T6P) was further evaluated. Wild type and TPS expressors were grown on agar supplemented with 100 mM glucose, fructose, or sucrose. TPS expressors consistently grew

faster than wild-type plants, exhibiting optimal growth on sucrose. It was found that glucose-6-phosphate (G6P) levels were inversely proportional to T6P levels: G6P accumulated in plants with low T6P but were depleted in those with high T6P content. However, a separate study on carbon utilisation in Arabidopsis presented a seemingly contradictory finding. Seedlings treated with either sucrose or trehalose accumulated high quantities of T6P. Yet, seedlings treated with trehalose showed retarded growth compared to those treated with sorbitol or sucrose (Schluepmann *et al.*, 2004). This suggests that while T6P plays a key regulatory role, the specific sugar source can also significantly influence growth outcomes.

Arabidopsis plants infected with *P.brassicae* accumulates high amounts of trehalose in the roots and hypocotyls and to a lesser extent in the stems and leaves (Brodmann *et al.*, 2002). The release of trehalose by these microorganisms could interfere with plant sugar sensing systems enabling the pathogens to divert carbon away from the plant metabolism and create carbon sink in infected organs (Brodmann et al. 2002). Two of the 23 putative *P.brassicae* sugar transporters are putative trehalose transporters. PLBR_LOCUS7736 is highly expressed during the cortical invasion stage and PLBR_LOCUS3479 is moderately expressed in germinating spores and to a lesser extent in resting spore stage (Figure S3A). The expression of PLBR_LOCUS7736 during cortical invasion stage supports the view that *P.brassicae* releases trehalose in the infected host cells to interfere with host sugar sensing mechanisms thus leading to accumulation of sugars in the infected tissue.

Plant pathogens exhibit a remarkable ability to manipulate the intricate network of plant hormone signalling pathways. This manipulation can target hormones crucial for disease resistance, such as salicylic acid, jasmonic acid, and ethylene, or hormones governing plant development, including auxin, cytokinins, and gibberellic acid (Chanclud *et al.*, 2016). *P.brassicae* is believed to interfere with the auxin and cytokinin homeostasis during disease development. The impact of *P.brassicae* on the host auxin homeostasis is contradictory. *P.brassicae* has been reported to increase the auxin level of infected plants (Ludwig-Müller *et al.*, 2009). However, reports of reduced auxin levels in *P.brassicae* infected plants which may happen as a result of conjugating free IAA to amino acids by the pathogen has been reported (Schwelm *et al.*, 2015; Smolko *et al.*, 2024). It has also been proposed that *P.brassicae* acts as sink for auxins (Devos *et al.*, 2006). There is no report that *P.brassicae* might produce auxins and based on the presence of Immunophilin like Prolyl: peptidyl Isomerase Regulator (INAPPI) Family, The Amino Acid/Auxin Permease (AAAP) Family and The ATP-binding Cassette (ABC) Superfamily in *P.brassicae* , it is tempting to speculate that reduction in auxin levels in

the host is the most likely scenario but whether this reduction happen through auxin uptake into the plasmodia and resting spores as sinks or uptake of free auxins and efflux of conjugated auxin forms is unclear.

Cytokinin is another plant growth hormone that is thought to be manipulated by *P.brassicae* during clubroot disease development. There is evidence that *P.brassicae* produces cytokinin (Dekhuijzen & Overeem, 1971; Muller & Hilgenberg, 1986; Rolfe *et al.*, 2016) but how *P.brassicae* produced cytokinin is released to host is unknown. In this study three transporters in the Equilibrative Nucleoside Transporter (ENT) Family and transporters of the ATP-binding cassette transporter family whose members transport cytokinin across plasma membranes (Hu & Shani, 2023) were found in *P.brassicae* . These transporters could leak the *P.brassicae* produced cytokinin into the host and thus interfering with host development and metabolism.

Plant defence hormonal signalling is at the centre of plant responses to fend off pathogenic and insect attacks. Responses to biotrophic pathogens often requires SA signalling, whereas necrotrophic pathogens or wounding mainly activate the JA/ET-dependent pathway. Thus, the crosstalk between these two pathways may lead to synergistic or antagonistic effects (Proietti et al. 2013). *P.brassicae* suppresses SA defence responses by transforming SA to methyl salicylate (MeSA) a mobile signal for systemic signalling (Ludwig-Müller *et al.*, 2015). *P.brassicae* interfere with JA homeostasis by conjugation to amino acids (Schwelm *et al.*, 2015; Smolko *et al.*, 2024). The possibility exists that *P.brassicae* may also interfere with the host JA signalling cascades by releasing Arachidonic acid (AA) into the host. Fifty-three putative fatty acid and lipid transporters were found in *P.brassicae* (Figure S3H) that are expressed at different developmental stages. The pathogen could use these transporters for the uptake of fatty acids and lipid from the host, the transport of fatty acid and lipids within the pathogen and the efflux of arachidonic acid to the host. In animals, arachidonic acid derivatives are involved in stress signalling cascades such as regulation of wound responses, inflammation, and immune responses. Plants possess a parallel, analogous Fatty acid signalling systems in which Linolenic acid is oxidized to produce plant oxylipins such as Jasmonate (JA) that modulate plant defence responses to wounding, insects and pathogens (Savchenko *et al.*, 2010). While plants do not naturally produce arachidonic acid, pathogens including *P.brassicae* (Bi et al. 2016) release these unusual fatty acids during infection and therefore plants have evolved to perceive and respond to them. Arachidonic acid has been reported to act as signalling molecule in plants. Transgenic Arabidopsis plants producing arachidonic acid or tomato plants treated with arachidonic acid had significantly high levels of jasmonic acid (JA) and significantly reduced

levels of salicylic acid (SA) with enhanced resistance to insect pest (aphid); the fungal pathogen *Botrytis cinerea* (Grape and B05.10), an oomycete pathogen *Phytophthora capsici* but were more susceptible to the bacterial pathogen *Pseudomonas syringae* pv tomato DC3000 [Pst] (Savchenko *et al.*, 2010). Plants infected with *P. brassicae* have been reported to contain reduced levels of SA in the early stages of infection (Siemens et al. 2006; Ludwig-Müller et al. 2015; Rolfe et al. 2016) thus the presence of AA in *P. brassicae* spores may also perturb plant SA-JA/ET signalling cascades.

3.5. Supplementary information

3.5.1. Supplementary figures

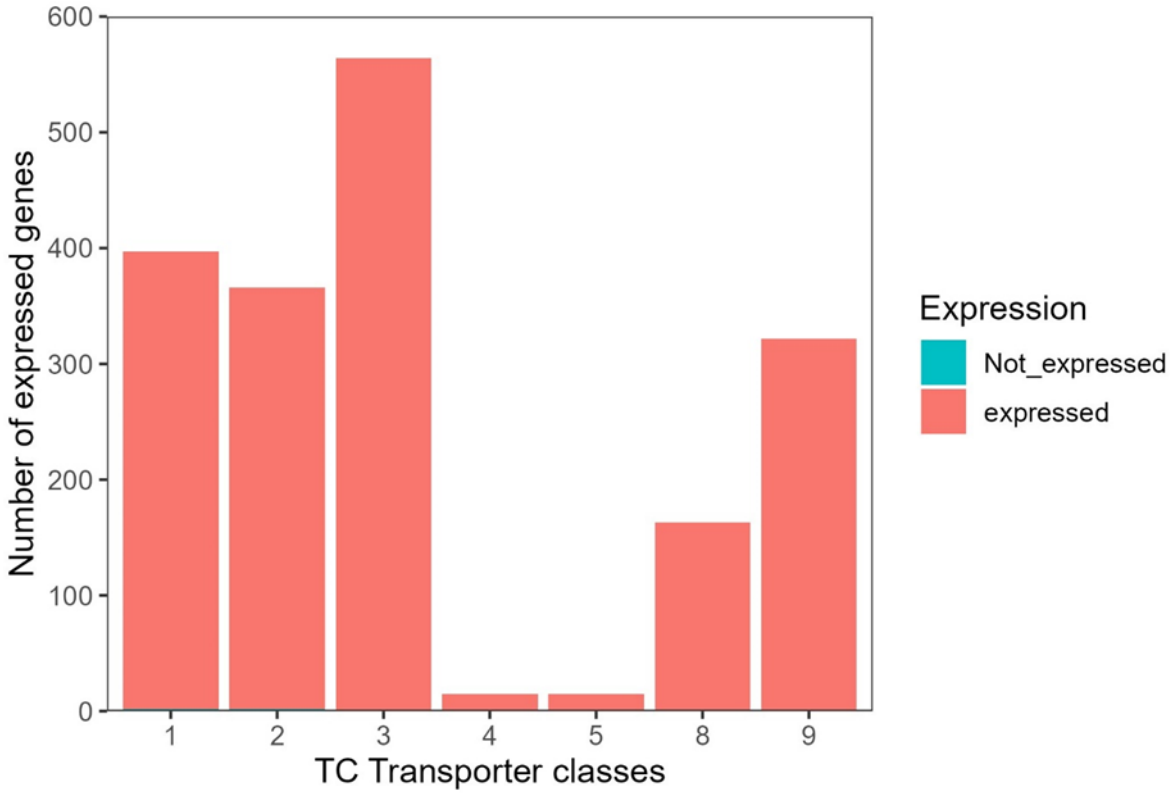
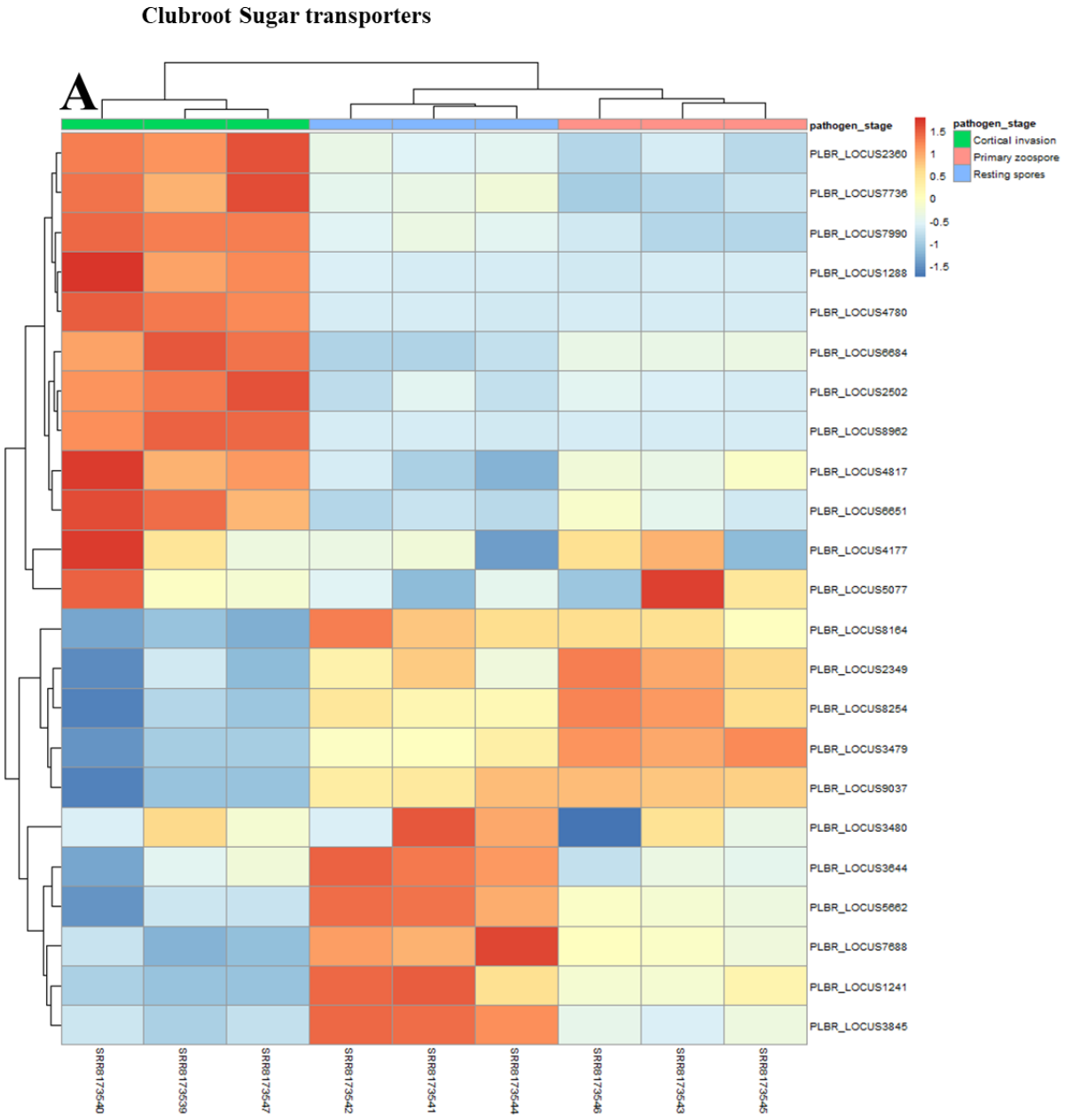
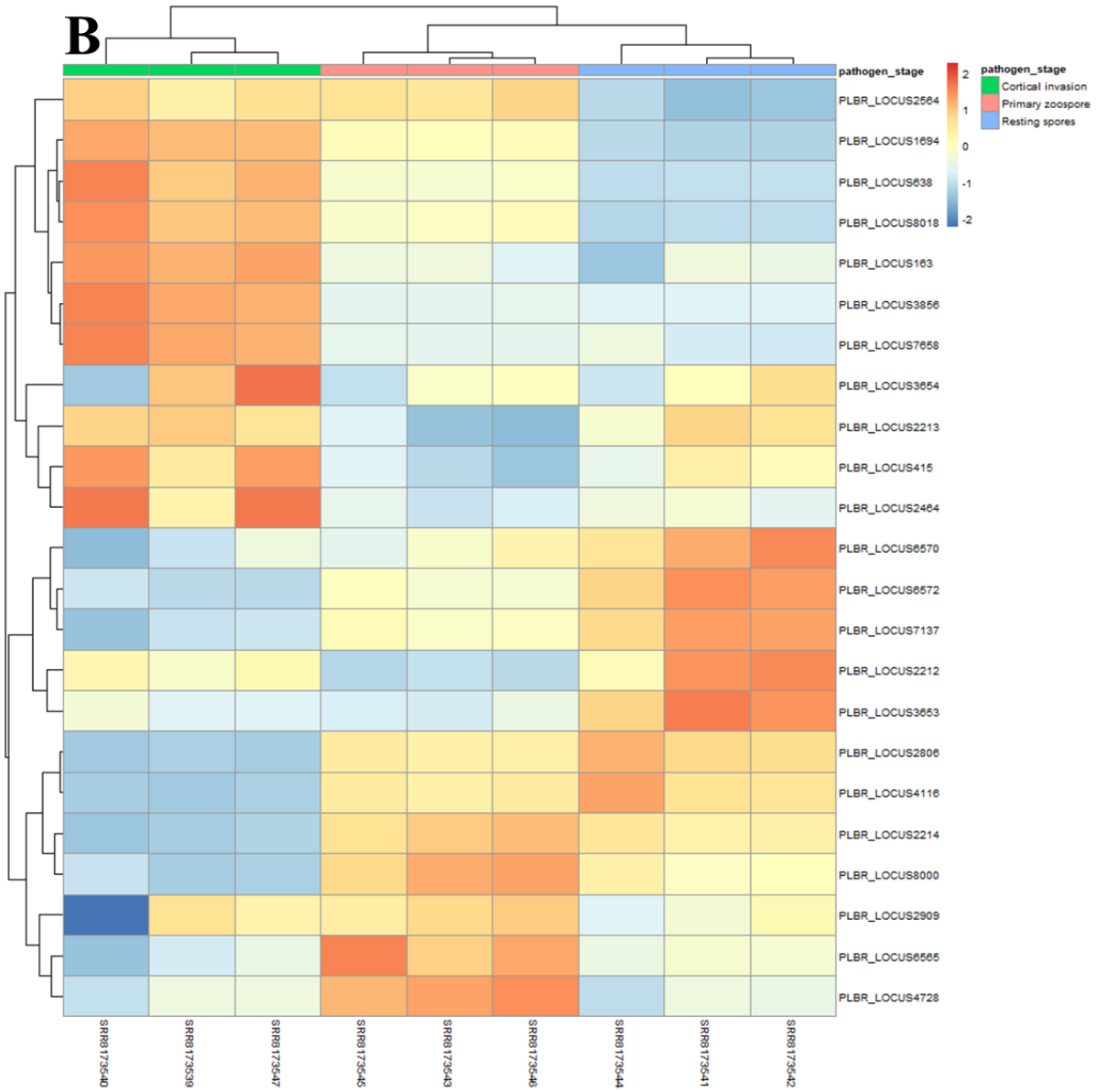


Figure S 2: Expression of transporters with expression cut off set to ≥ 2 FPKM reads. Normalised expression of P.brassicae transporters. Transporter transcript counts from 15 RNAseq projects were normalised using DESEQ2 R package. genes with thirty or less reads across the samples were removed before normalisation of the counts. The normalised read counts are expressed as fragments per kilobase of transcripts per million mapped reads (FPKM). Transporters were considered as expressed if they have ≥ 2 FPKM reads.

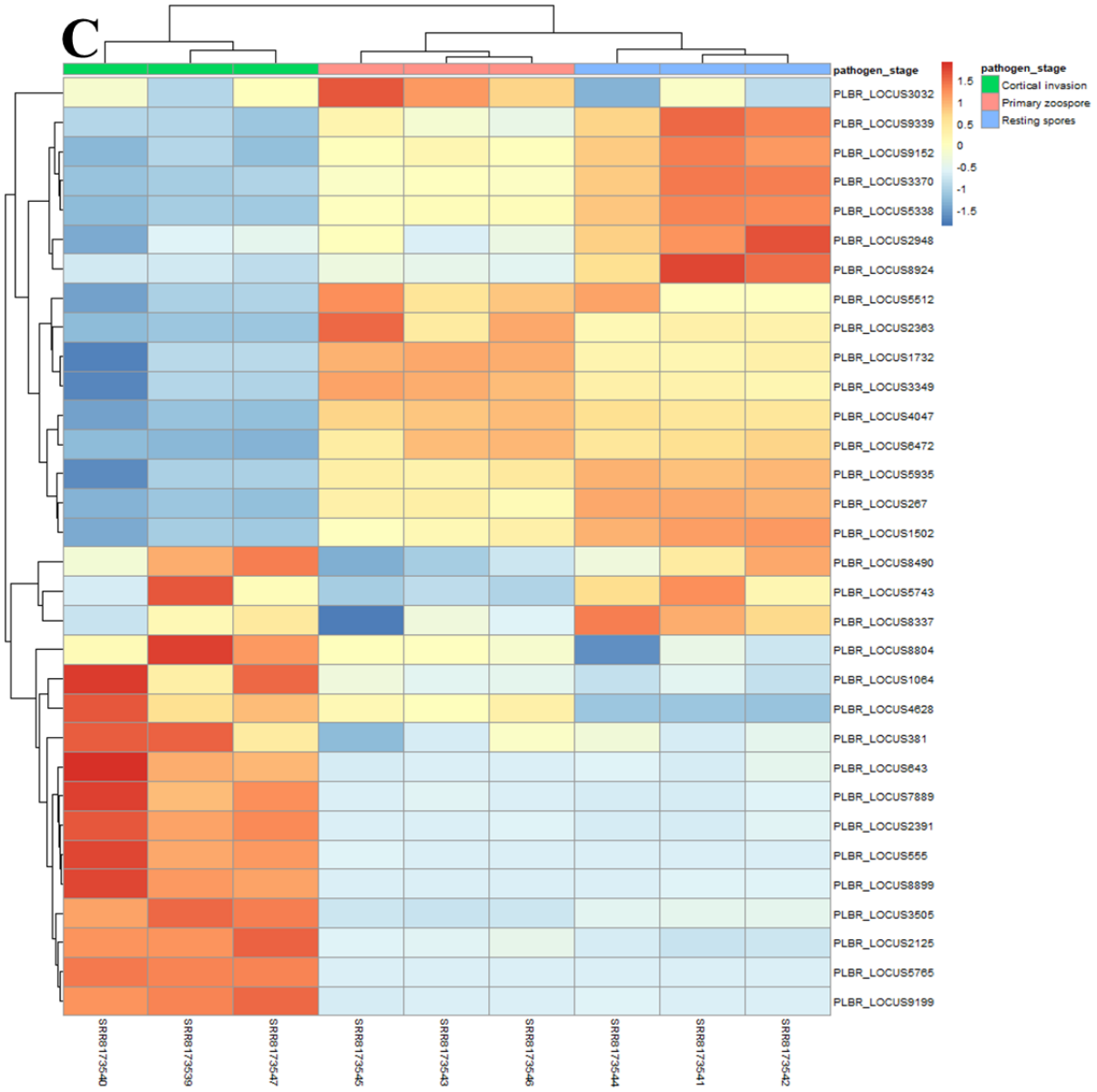
Figure S 3: Expression of putative substrate transporters in different P.brassicae developmental stages.
The normalised transporter transcript counts from PRJNA331021 RNAseq project were used to compare transcript abundances of transporters in different stages



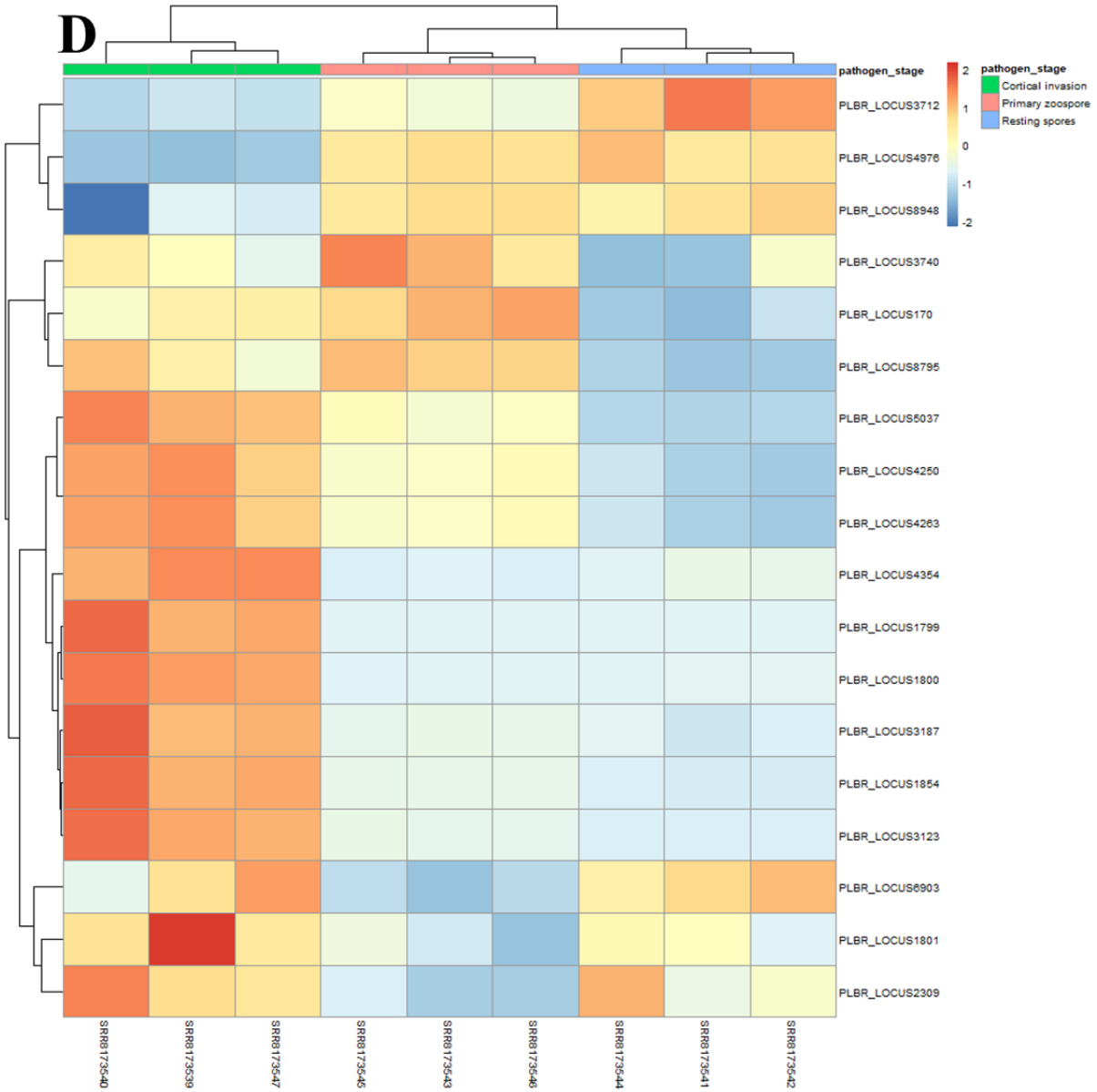
Clubroot nucleosides and nucleotides transporters



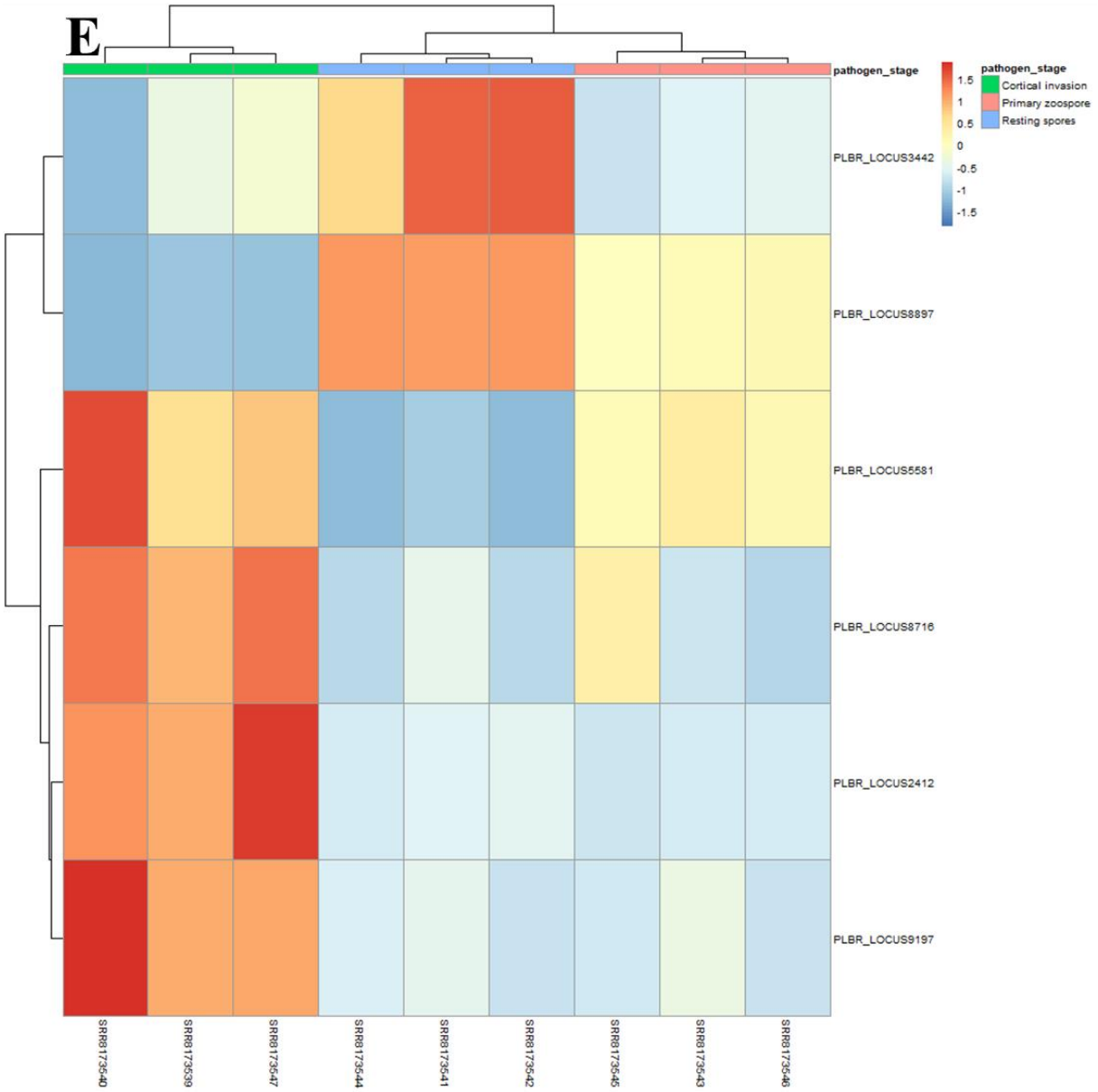
Clubroot amino acids and amino acid derivatives transporters



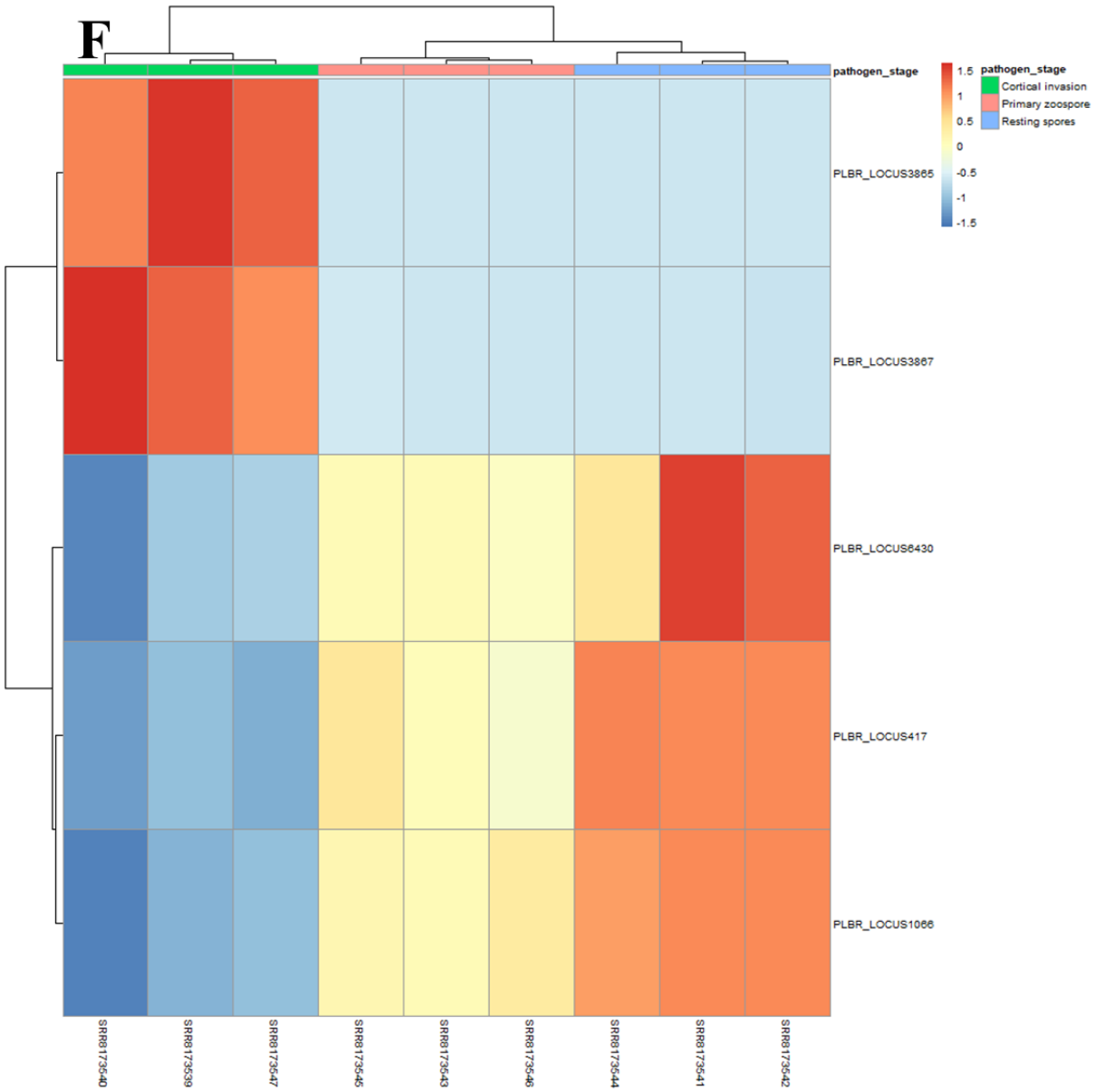
Clubroot vitamins transporters

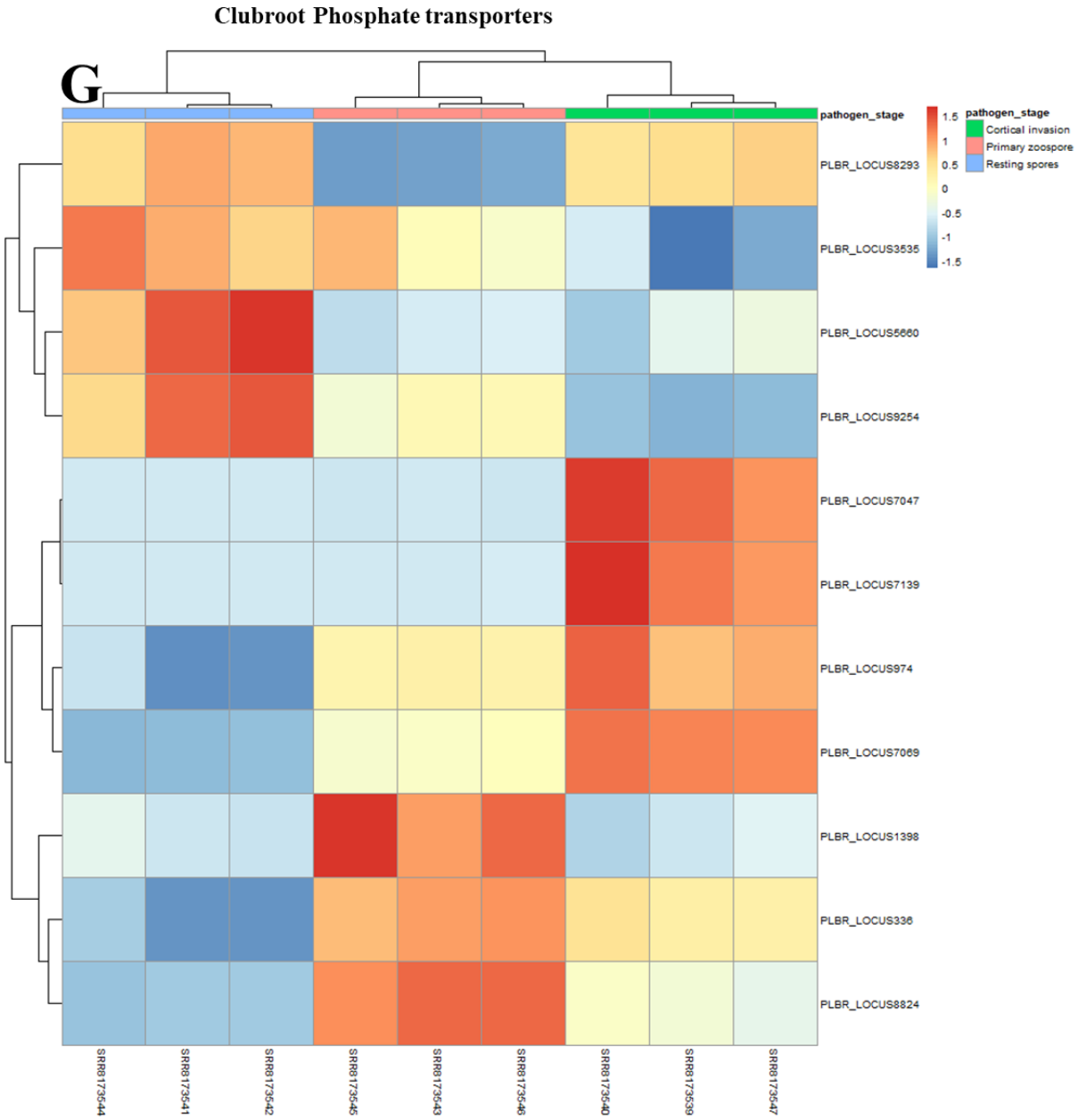


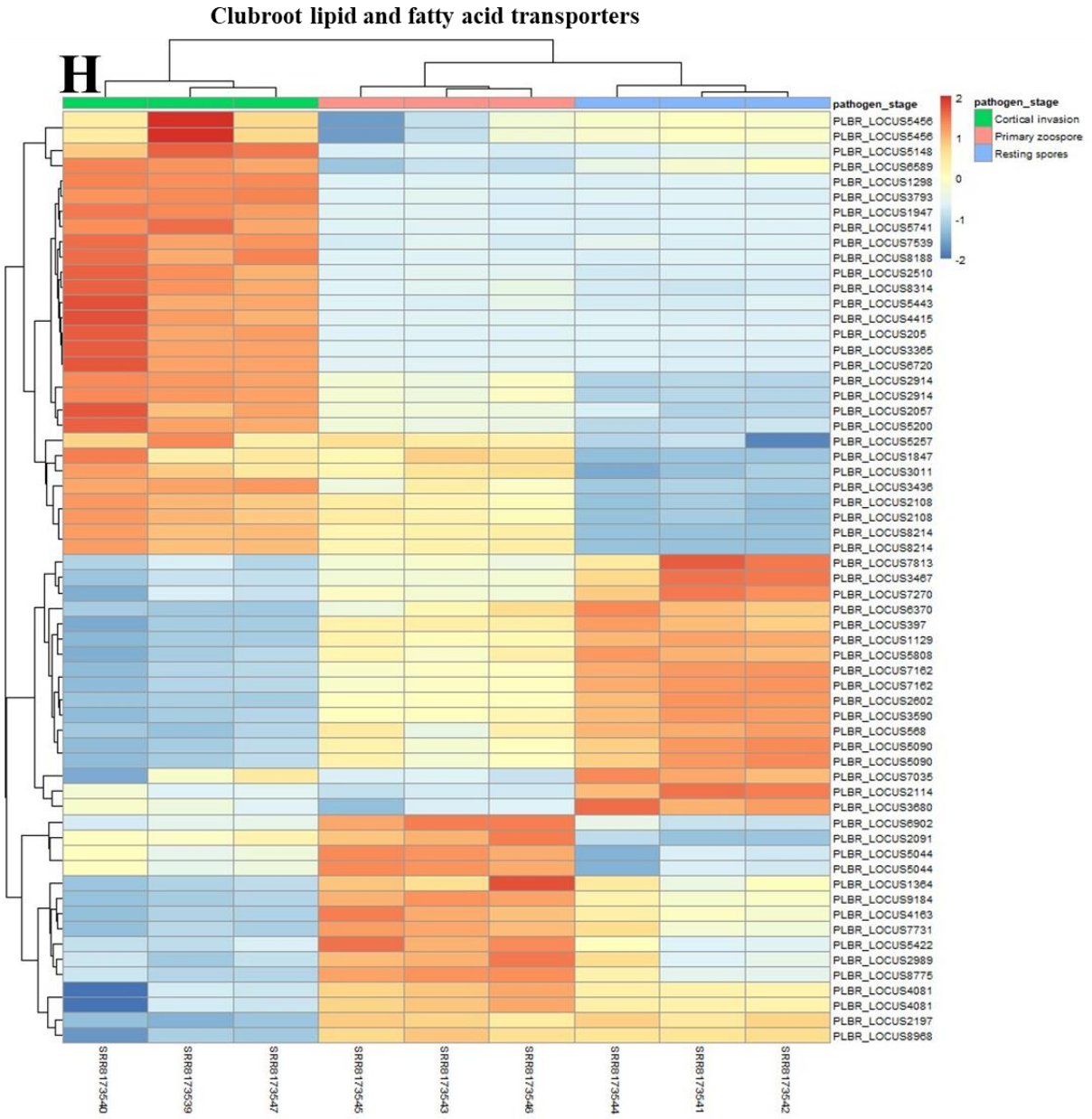
Clubroot nitrogen transporters



Clubroot Sulphate transporters




Clubroot Phosphate transporters
G
pathogen_stage
PLBR_LOCUS8293
PLBR_LOCUS3535
PLBR_LOCUS5660
PLBR_LOCUS9254
PLBR_LOCUS7047
PLBR_LOCUS7139
PLBR_LOCUS974
PLBR_LOCUS7069
PLBR_LOCUS1398
PLBR_LOCUS336
PLBR_LOCUS8824
SRR8173544
SRR8173541
SRR8173542
SRR8173545
SRR8173543
SRR8173546
SRR8173540
SRR8173539
SRR8173547
1.5
1
0.5
0
-0.5
-1
-1.5
pathogen_stage
Cortical invasion
Primary zoospore
Resting spores



3.5.2. Appendices

Appendix 1

>SPQ94026.1 (PLBR_LOCUS1241, PbSWEET)

MSASHVITDVGVVGGAPGYLKAICAMASMMTTLSPSSVLLRMLRRKRDTGEIAWITL
WTSALAQTLWVTYYFLRPNDDTVALWLSVFGAVTFFGYLFVHYSVSSYRDRPQRRR
MLYYLLVTIVLWGTAFTTIVPKKRRADYAGVVCNIVNLIMLMAPLSTLGEVFRTQSA
RTIPSLPALASTFGAACWLLWGYLDNDIFVCVPSVIALLLGSVQLSLIGFYGSGAPKHP

CHAPTER 4: EXPERIMENTAL CHARACTERISATION OF PLASIMODIPHORA BRASSICAE ISOPENTENYL TRANSFERASE GENES

Abstract

Isopentenyl transferase (IPT) enzymes regulate a rate limiting step in *de novo* biosynthesis of cytokinins (CK) in plants. IPTs can produce CKs using adenine phosphates or tRNA as substrates. Adenine-IPTs produce IP-type CKs whereas tRNA-IPTs produce *cis*-zeatin. Also, tRNA-IPTs modify tRNA, which is important in ensuring the fidelity of protein synthesis. Consequently, tRNA-IPTs are widespread and not necessarily involved in CK-induced changes in host development. Many plant pathogens produce plant growth regulators to manipulate host development, physiology and defence responses. Cytokinins (CK) can be produced by pathogenic bacteria, fungi and insects. It has been proposed that *Plasmodiophora brassicae*, pathogen responsible for clubroot disease of Brassicas, manipulates host developmental processes such as cell division, expansion and differentiation by manipulating host CK homeostasis. However, the type of CKs produced by *P.brassicae* , the extent to which this occurs, and its importance in pathogenesis is unclear. The *P.brassicae* genome contains two IPT genes which are expressed during gall formation. In this chapter, I experimentally characterised the *P.brassicae* IPTs to determine their importance in CK production during pathogenesis. The results confirm that *P.brassicae* IPTs function as tRNA-isopentenyl transferases and may modify tRNA to perform their canonical functions. Overexpressing one of the *P.brassicae* IPTs in an Arabidopsis mutant lacking both tRNA-IPT genes (atipt2,9) restored the shoot phenotype to wild-type. Additionally, *P.brassicae* -infected Col-0 and some *P.brassicae* IPT-overexpressing lines in atipt2 and atipt9 mutant backgrounds exhibited axillary bud outgrowth, a phenotype that is observed in response to increased CK levels. To validate these results, samples from these *P.brassicae* IPT overexpressing lines have been sent to a collaborator for CK quantification. Nevertheless, these findings support the view that *P.brassicae* -derived cis-zeatin can manipulate host development.

4.1. Introduction

Isopentenyl transferases (IPT) enzymes regulate a rate limiting step in *de novo* biosynthesis of cytokinins (CK) in plants (Kieber & Schaller, 2018). Two major classes of IPTs exist – adenylate IPTs that catalyse the transfer of an isopentenyl group from dimethylallyl pyrophosphate (DMAPP) to ATP or ADP and tRNA IPTs that catalyse the transfer of a dimethyl allyl group from DMAPP onto an adenine at position 37 on transfer RNAs (Sakakibara, Hitoshi, 2006). The adenylate IPTs are the rate limiting step in the synthesis of active plant cytokinins such as trans-zeatin (tZ) and isopentenyl adenine (iP). The tRNA IPTs are responsible for the synthesis of cis-zeatin (cZ). Whilst originally cZ was considered to have limited activity in many plant species including Arabidopsis (Gajdosová *et al.*, 2011), there is evidence for a role. cZ is the dominant form of CK in monocots, particularly members of the Poaceae, where it accumulates significantly at specific timepoints and triggers cytokinin responses (Yonekura-Sakakibara et al., 2004; Gajdosová et al., 2011). In Arabidopsis, cZ are the major cytokinin forms in dry seeds and during senescence (Gajdosová et al., 2011) and mutants of AtIPT9, an IPT that produces cZ, shows a visible phenotype (Miyawaki *et al.*, 2006).

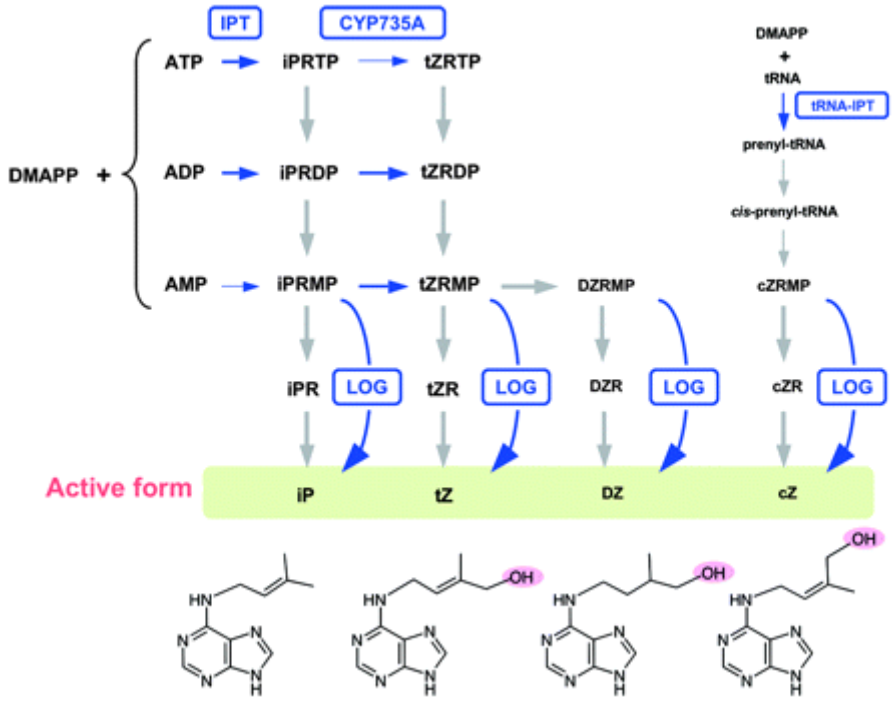


Figure 31: Simplified current model of cytokinin biosynthesis pathway in plant. DMAPP, dimethylallyl-pyrophosphate; ATP, Adenosine triphosphate; ADP, Adenosine diphosphate; AMP, Adenosine monophosphate; iPRTP, iP riboside 5'-triphosphate; iPRDP, iP riboside 5'-diphosphate; iPRMP, iP riboside 5'-monophosphate; tZRTP, tZ riboside 5'-triphosphate; tZRDP, tZ riboside 5'-diphosphate; tZRMP, tZ riboside 5'-monophosphate;

DZRMP, DZ riboside 5'-monophosphate; cZRMP, cZ riboside 5'-monophosphate;Ipr, iP riboside; tZR, tZ riboside; DZR, DZ riboside; cZR, cZ riboside;tZ, trans-zeatin; iP, N^6-(Δ^2-isopentenyl) adenine; cZ, cis-zeatin; and DZ, dihydrozeatin, LOG; lonely guy. Taken from (Hirose et al., 2008).

Although many studies in plants have focused on the role of tRNA IPTs in the production of cytokinins, the modifications that they make to tRNAs are important in diverse organisms. tRNAs are found in both the cytoplasm and mitochondria of eukaryotes and both mitochondrial and cytoplasmic tRNAs, are subjected to modifications by tRNA IPT generating the i6A37 modification. Modifications to the body of the tRNA facilitate folding whilst those in the anticodon loop improves anticodon-codon pairing, increases translation efficiency and decrease frameshifts (Khalique et al., 2020).

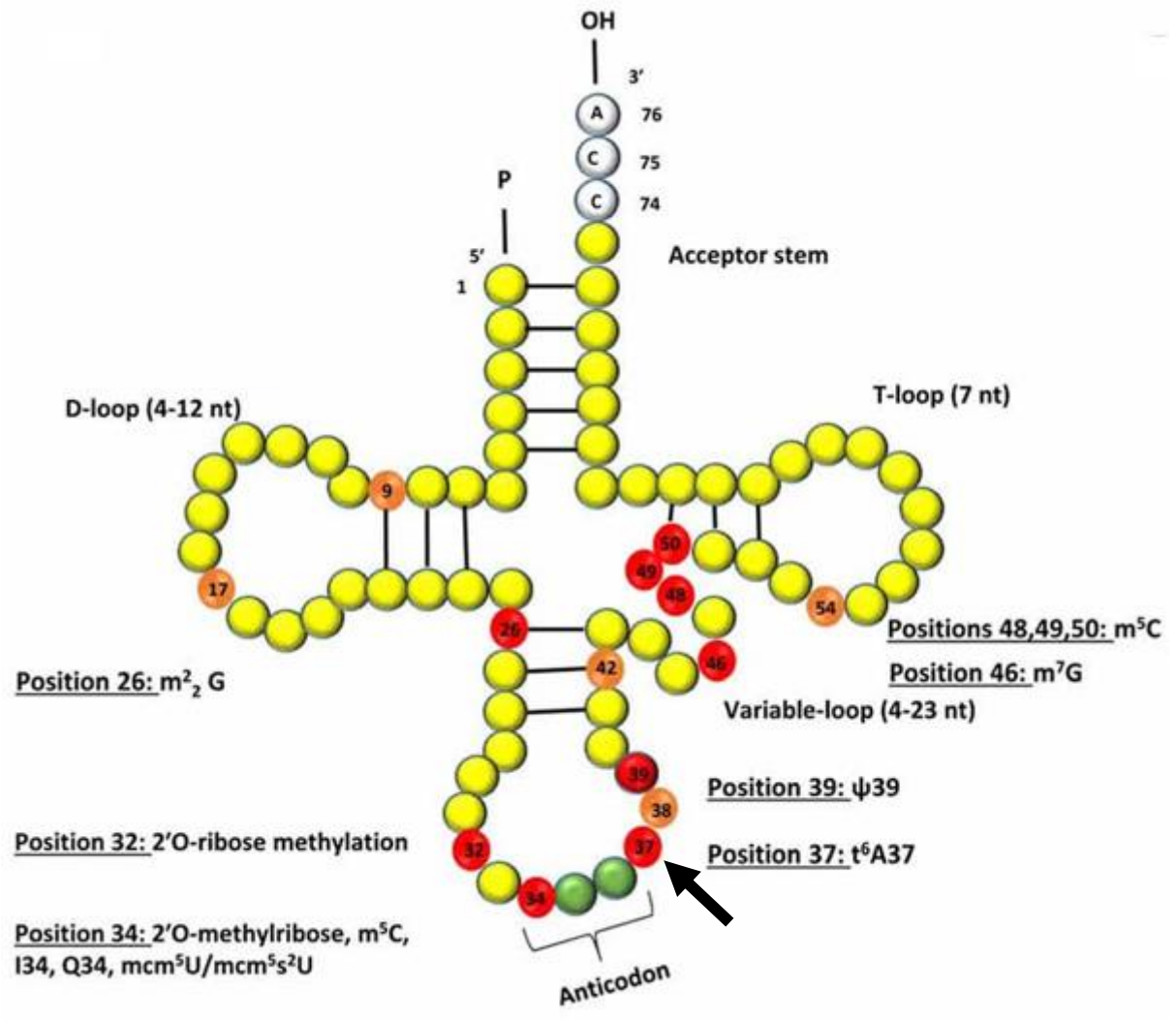


Figure 32: Schematic diagram showing post transcription modifications of tRNA by different tRNA modifying enzymes.

Modified residues are shown in light orange and red. the red residues are important for human diseases. The A37 modification catalysed by tRNA-IPT is indicated by the black arrow. Taken from (Bednarova et al., 2017).

IPT genes in the genomes of many plant pathogens are associated to the generation of phytohormone signals to modify their host development and metabolism (Spallek, Thomas *et al.*, 2018). For example, bacterial pathogens can manipulate auxins and cytokinins leading to the formation of undifferentiated callus (Sakakibara et al., 2005), shoot galls (Goethals et al., 2001) or root galls (Nilsson & Olsson, 2006). As well as affecting development, alterations in auxins and cytokinins can influence defence responses, promoting susceptibility (Chanclud *et al.*, 2016; Spallek, T *et al.*, 2018) and increase local sink strength facilitating access to nutrients (Mothes & Engelbrecht, 1961; Walters et al., 2008). Pathogens can alter hormone levels by interfering with host hormone metabolism or by directly producing hormones or mimicking compounds and thereby inhibit defences, modify nutrient flows and/or induce symptom development Chanclud et al. (2016). *P.brassicae* which causes clubroot disease of Brassica plants has been proposed to produce CK during the course of disease development (Ingram, 1969a; Ingram, 1969b; Williams *et al.*, 1969). The development of galls in the roots and hypocotyls of infected plants is proposed to involve an increase in CK content following *P.brassicae* infection (Dekhuijzen & Overeem, 1971). Support for CK production during clubroot disease development is given by the observation that callus from *P.brassicae* infected cabbage could grow in tissue culture independent of exogenous CK supply (Williams *et al.*, 1969). Muller and Hilgenberg (1986) provided evidence for cytokinin synthesis in *P.brassicae* by feeding radioactively labelled adenine to isolated plasmodia. The incorporation of the labelled adenine into a compound tentatively identified as trans-zeatin suggested that *P.brassicae* can synthesize this cytokinin.

The overall picture of changes in phytohormone content in *P.brassicae* -infected Brassicas is complicated (Ludwig-Müller *et al.*, 2009). Dekhuijzen and Overeem (1971) found that *P.brassicae* infected turnips had 100-fold higher CK content compared to the controls. Devos et al. (2006) reported increased levels of CK in *P.brassicae* -infected Arabidopsis at 4 days post inoculation (dpi) which they proposed led to *de novo* meristem formation. However, Malinowski et al (2016) found that this concept was incorrect as it is existing vascular cambium activity that is stimulated, rather than new meristems being formed. Malinowski et al (2016) found that infected Arabidopsis plants had low CK levels compared to controls at 16 dpi, a time point that corresponds with the onset of gall formation in Arabidopsis. In the same study, they found that *P.brassicae* infected *Arabidopsis thaliana* ipt1;3;5;7 quadruple mutant, which is

impaired in the biosynthesis of trans-zeatin and IP-type cytokinins had increased CK response gene expression but there was no rescue of the ipt1;3;5;7 phenotype and no CK increase. However, the activation of host CK response genes showed that *P.brassicae* -infection results in the production of a small amount of CK which are able to stimulate host CK signalling cascades; this may be of pathogen origin. Recently, it was reported that phloem sap extracted from leaf petioles of *P.brassicae* infected *B. napus* had high contents of IP-type cytokinins from 7 dpi (Blicharz et al., 2025). These apparently conflicting results arise from a number of factors. Measurement of CKs in their different forms is technically challenging (Tarkowski *et al.*, 2009) and some studies lack differentiation between precursors, active forms, conjugates and degradation products. Also, changes in CK content may be highly localised, affected by disruption in the host vasculature which may reduce CK influx into a region or reduce efflux, vary by time point, tissue type and host plant (Del Bianco *et al.*, 2013; Antoniadi *et al.*, 2015). Moreover, molecular studies of *P.brassicae* -infected plants generally indicate a repression of host cytokinin biosynthesis and degradation genes upon infection (Siemens *et al.*, 2006; Malinowski *et al.*, 2016) , though increased expression of CKX4 and some cytokinin response factors has also been reported (Siemens *et al.*, 2006; Blicharz *et al.*, 2025).

The cytokinin landscape of *P.brassicae* -infected plants is further complicated by conflicting reports of the abundance of different CK species. Early work by Dekhuijzen and Overeem (1971) identified 2-isopentenyladenine (2iP), zeatin, zeatin riboside, and isopentenyladenosine (2iPA) in both healthy and *P.brassicae* -infected turnips, with zeatin and zeatin riboside being the most active forms. Subsequently, Dekhuijzen (1981) found higher concentrations of zeatin, zeatin riboside, and zeatin riboside glucosides in 21-day-old *P.brassicae* plasmodia isolated from turnip callus compared to the host cytoplasm. The presence of glucose-6-phosphate derivatives of zeatin and zeatin riboside in the host cytoplasm, but not the plasmodia, led to the hypothesis that CKs are synthesized within the plasmodia and then released into the host tissue, where they are modified by host enzymes (Dekhuijzen, 1981). In contrast, Muller and Hilgenberg (1986) reported that *cis*-isomer CKs were the predominant forms in both healthy and 23-day-old *P.brassicae* -infected Chinese cabbage. More recently, elevated levels of isopentenyl adenine (iP) and isopentenyl adenosine (iPA) were observed in *P.brassicae* -infected *Arabidopsis* roots and hypocotyls (Devos *et al.*, 2006), and high concentrations of iP-type CKs have also been measured in phloem sap extracted from leaf petioles of *P.brassicae* -infected *B. napus* (Blicharz *et al.*, 2025). Cytokinin composition varies across plant species (Gajdosová *et al.*, 2011) and its function depends on its temporal and spatial distribution within

the plant (Sakakibara, 2003; Hirose *et al.*, 2008). The observation that *P.brassicae* infection in the *Arabidopsis thaliana ipt1;3;5;7* quadruple mutant elicited a cytokinin-signalling response similar to wild type, yet failed to restore meristematic function, suggests that *P.brassicae* -produced cytokinin species, while perhaps not crucial in *Arabidopsis*, may play a significant role in other Brassica species. Therefore, understanding *P.brassicae* cytokinin biosynthesis pathways is essential, as the host's response to infection is likely influenced by the specific cytokinins produced by the pathogen.

The *P.brassicae* genome has been sequenced (Schwelm et al., 2015; Rolfe et al., 2016; Stjelja et al., 2019). The analysis of *P.brassicae* genome revealed that *P.brassicae* contains two Isopentenyl transferases (IPTs) genes (Rolfe et al., 2016). The two *P.brassicae* IPT genes are expressed during gall formation and development (Malinowski et al., 2016) implying that they could be involved in CK production (Figure 1) and in fidelity of protein translation (Figure 2). Whether these utilise AMP/ADP/ATP (synthesising iPA and tZ) or tRNA (synthesising cZ and tRNA modification) as substrates is not known. Determining the substrate specificity of PbIPTs based solely on sequence homology is unreliable. Both adenylate and tRNA IPTs share a conserved GxxxxGKT(S) motif, similar to the P-loop nucleotide-binding motif. Additionally, tRNA-IPTs possess an arginine-rich motif resembling the RxxR sequence in tRNA-guanine transglycosylase (TGT) but lack the DDxxD diphosphate-binding motif typical of prenyltransferases and have low sequence homology to other isoprenoid prenyltransferases (Soderberg & Poulter, 2001) thus necessitating experimental validation of their substrate specificity. The aim of this study was to experimentally characterise *P.brassicae* IPT genes. The main objectives were to determine the types of cytokinins produced by the pathogen (given the species-specific nature of cytokinin activity and response) and to investigate whether these IPT genes might have functions beyond cytokinin biosynthesis.

4.2. Material and methods

4.2.1. Plant materials and growth conditions

Arabidopsis thaliana ecotypes used in this study were Col-0, atipt2 (Salk_042050C), atipt9 (Salk128927C) and atipt2,9. The Arabidopsis mutants were in Col-0 background. Seeds of Col-0, atipt2, and atipt9 were generously provided by Prof Jurrian Ton, University of Sheffield, United Kingdom. Seeds of atipt2,9 mutant were generously provided by Prof. Dr Thomas Schmülling, Freie Universität Berlin, Germany and Dr David Zalabák , Palacký University, Czech Republic. Chinese cabbage Wong Bok was used to propagate the *P.brassicae* inoculum. In all experiments, seeds were sterilised by placing them in an open Eppendorf tube within a vacuum chamber with a beaker containing a bleach and hydrochloric acid solution. The chamber was sealed, and the seeds were exposed to the chlorine gas for 3 hours. For seedling growth, seeds were sown on half-strength MS agar (10g sucrose, 5g MES, 2.2g MS with vitamins, 12g agar, pH 5.7, 1L dH2O) under sterile conditions. For root phenotyping and cytokinin quantification experiments, the quantity of agar was increased to 2% (w/v). Seeds were stratified at 4^{o}C for 3 days. For growth in compost, seeds were sown in F2+S soil substrate (Levington advance) and placed in a growth chamber with irradiance 100 µmol $m^{-2}sec^{-1}$, 21°C/18°C (day/night) and 60% humidity. Arabidopsis plants were grown under short day conditions (10 hours light/14 hours dark) during vegetative growth for six weeks after which the daylength was switched to long day conditions (16 hours light/8 hours dark) to induce flowering.

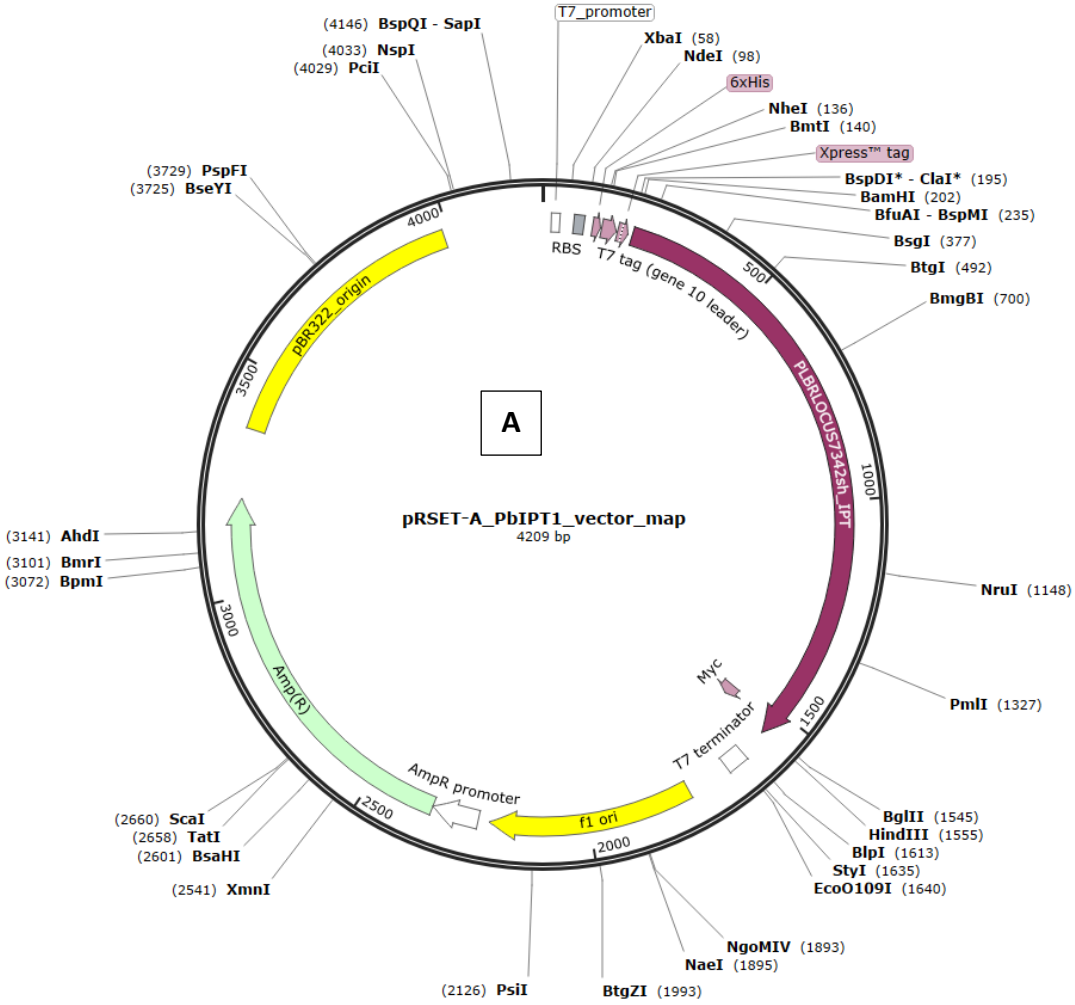
4.2.2. Phylogenetic analysis of *P.brassicae* IPTs

Protein sequences of adenylate and tRNA isopentenyl transferases from various organisms, including protists, bacteria, plants, algae, fungi, and animals (Frebort *et al.*, 2011; Nishii *et al.*, 2018; Khalique *et al.*, 2020) were used for phylogenetic analysis. Identifiers of all proteins used in this study are listed in Table S5. Using default settings and a 500 bootstrap, a maximum likelihood phylogenetic tree was created in MEGA X, version 10.2.4 software (Tamura *et al.*, 2021). The resulting tree was exported as a Newick file and saved for better presentation and annotation using the Interactive Tree of Life online tool (Letunic & Bork, 2024).

4.2.3. *P.brassicae* tRNA isopentyl transferase gene synthesis

PbIPT1 and PbIPT2 protein sequences (Appendix 2) as predicted in Pb314v2 reference genome were used to create PbIPTs DNA coding sequences were then synthesized and cloned by Thermo Fisher Scientific Inc, United Kingdom. For *E. coli* expression, the constructs were

synthesised with a C-terminal MYC tag and were cloned into the pRSET-A vector between the BamHI and HindIII restriction sites, resulting in an N-terminal 6xHis tag fusion. For *Arabidopsis thaliana* expression, the constructs were codon-optimized and cloned into the pDNR221 gateway entry vector.



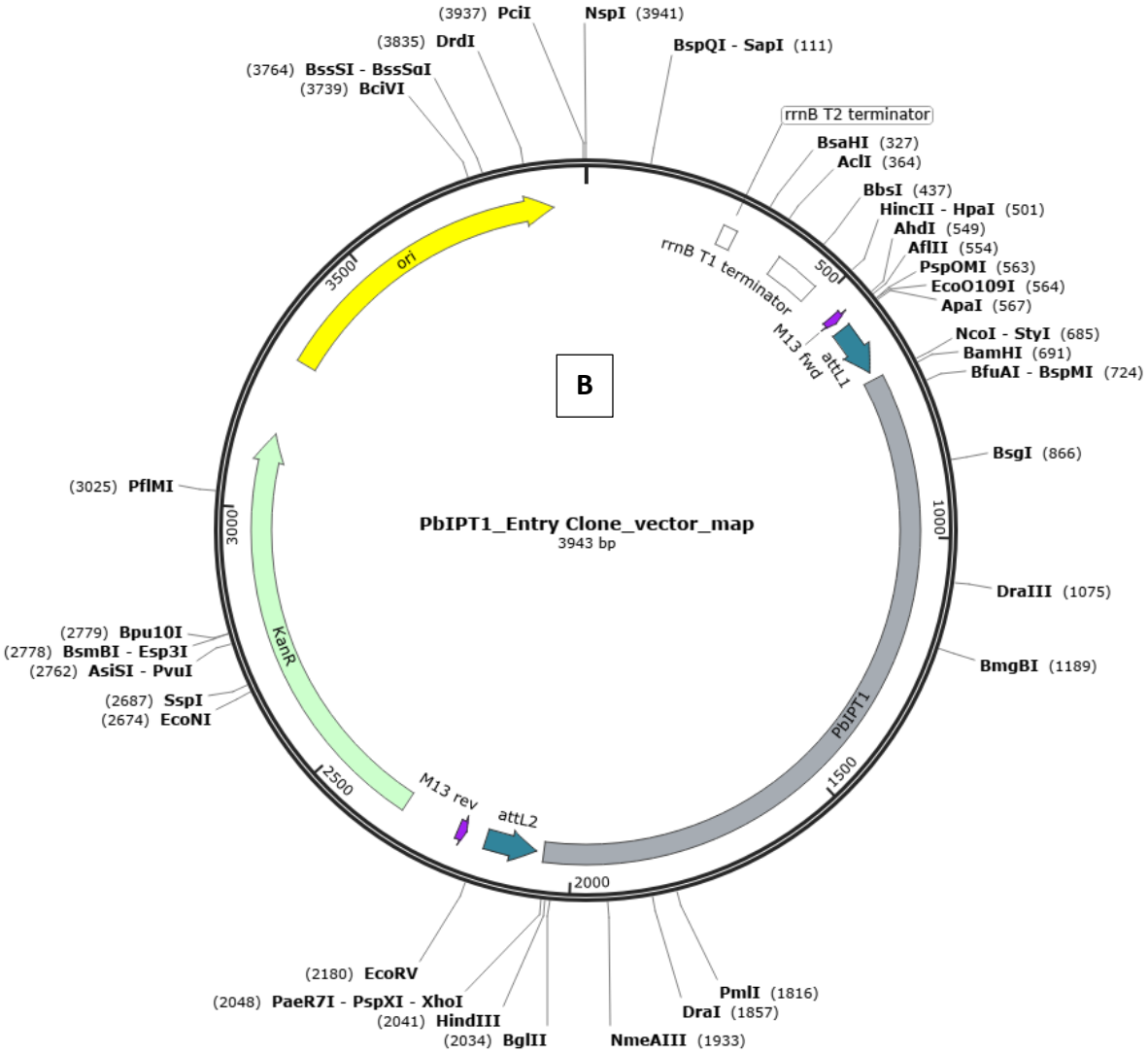


Figure 33: Vector map showing the cloning of P.brassicae IPT genes in different expression vectors.
A) protein expression vector pRSET-A and (B) pDNR221 gateway entry vector

4.2.4. Plasmids construction and Cloning

ATIPT4, which is an adenylate isopentyl transferase, and MiaA, which a tRNA isopentyl transferase, were selected to be used as positive controls for the PbIPTs enzyme assays. Full length sequences of ATIPT4 and MiaA genes were produced by polymerase chain reaction (PCR) using genomic DNA (gDNA) of *Arabidopsis thaliana* Col-0 and gDNA of *E.coli* BL21(DE3) cells (Thermo Fisher Scientific Inc) using MyTaq DNA polymerase (Bioline/Meridian Bioscience). The primers used to amplify are listed in table S6. Bands of ATIPT4 and MiaA PCR products were cut out from the gel and purified using QIAquick gel extraction kit (Qiagen, Germany) following the manufacturer's instructions. The purified DNA of ATIPT4, MiaA and pRSET-A empty vector were separately digested with high fidelity SacI and HindIII restriction enzymes for 30minutes. Direct cloning of ATIPT4 and MiaA into the

pRSET-A protein expression vector was unsuccessful. All PCR products therefore were cloned using TA cloning system in the pGM-T cloning vector (Promega corporation, UK) and gene constructs were cut from pGM-T for restriction enzyme cloning into different expression vectors. Purified PCR products of these genes were cloned using TA cloning system in the pGM-T cloning vector (Promega corporation, UK). Purified DNA of ATIPT4 and MiaA were put in different reaction tubes and cloning vector pGM-T vector DNA, T4 DNA ligase and ligation buffer (Promega corporation, UK) were added to the tubes and were left overnight at 15°C for ligation.

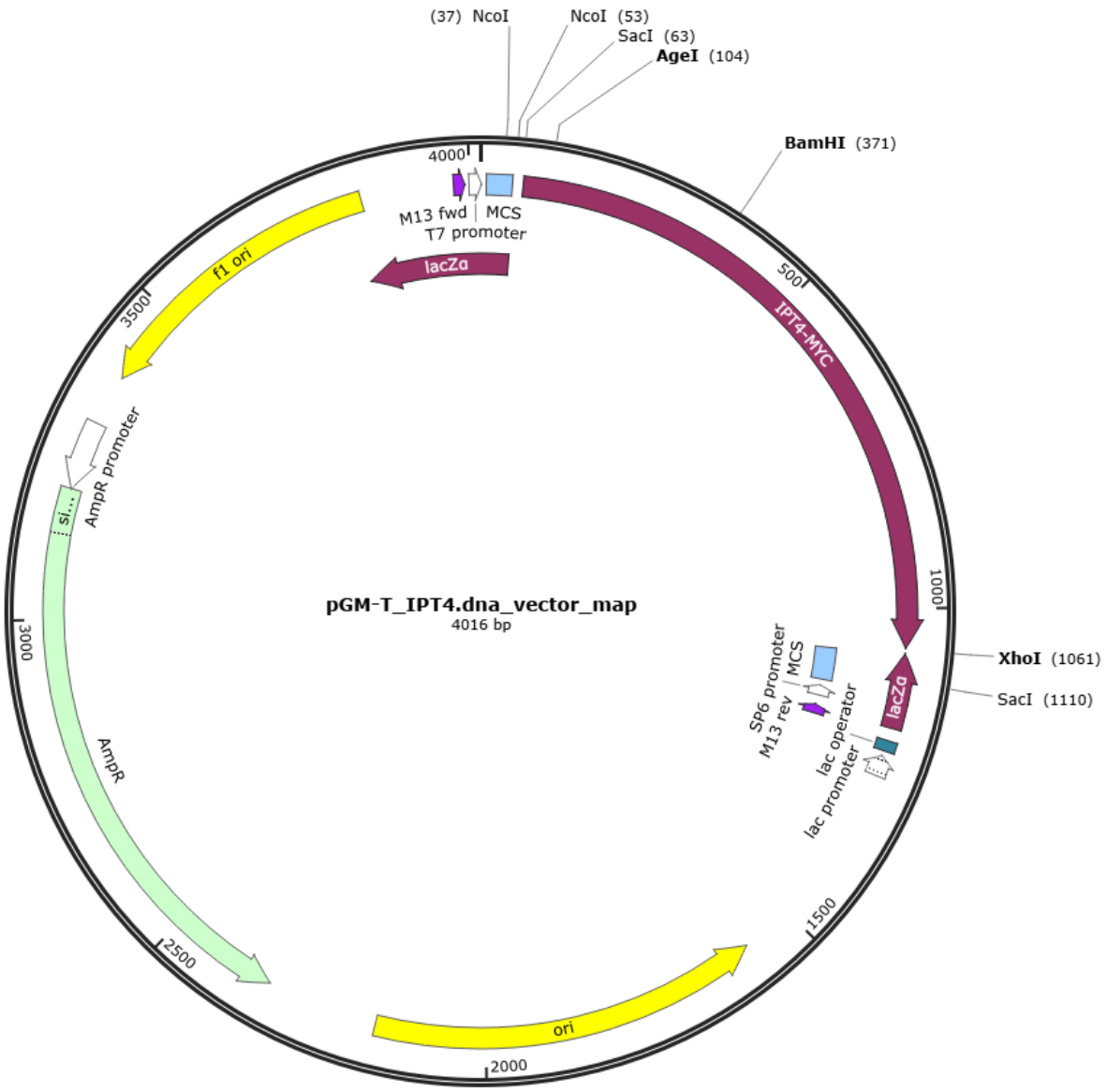


Figure 34: Vector map showing a gene construct cloned in pGM-T cloning vector system. Gene constructs of ATIPT4 and MiaA were amplified from respective genomic DNA using an enzyme that creates TA overhangs at the ends of the PCR product.

The ligation mixture was heat-shock transformed into 50 µL of DH5α competent cells (Thermo Fisher Scientific Inc.) by incubating the mixture in a 42°C water bath for 90 seconds, followed by a three-minute incubation on ice. After transformation, 1 mL of Luria-Bertani (LB) broth without antibiotics was added to the cells, which were then placed in a shaker at 200 rpm and 37°C for 90 minutes to allow growth. A stock solution of ampicillin was prepared by dissolving 50 mg of ampicillin powder in 1 L of deionized water. Similarly, the X-gal stock solution was made by adding 20 mg of X-gal to 1 mL of dimethyl sulfoxide (DMSO). To prepare the LB agar plates, 100 µL of ampicillin and 500 µL of X-gal were added to autoclaved LB agar media that had been cooled to approximately 55°C. The mixture was then poured into Petri dishes, with 20 to 25 mL of agar per dish. Transformants were selected on the LB agar plates using the Blue-white screening method (Sigma-Aldrich). Positive transformants were grown overnight in liquid LB media with ampicillin and plasmid DNA was extracted using Qiaprep spin Miniprep kit (Qiagen, Germany). To confirm if the cloned genes were correct and without unintended mutations, plasmid gene constructs DNA were sent for sequencing at MRC PPU DNA sequencing facility, University of Dundee, Scotland after each cloning step.

4.2.5. Protein expression and purification

Plasmid constructs of ATIPT4, MiaA, LacZ, PbIPT1 and PbIPT2 cloned in pRSET-A expression vector were transformed into expression cells E. coli BL21(DE3) pLysS (Invitrogen). Transformants were selected on plates containing 50mg/L chloramphenicol antibiotic. One colony of each gene construct was used to inoculate 15 mL of LB media containing 15 µL of ampicillin (50 µg/mL) and 15 µL of chloramphenicol (35 µg/mL). The cultures were then incubated overnight at 37°C with shaking at 200 rpm. ATIPT4, MiaA, LacZ, PbIPT1 and PbIPT2 proteins were expressed as described in (Kettle *et al.*, 2015) with modifications. 1ml of the overnight culture was used as a starter culture to inoculate 100ml LB/Amp/CM, in a 250ml flask. The bacteria were grown to an OD_{600} 0.4at 37^{o}C, and then 0.4 mM IPTG added to induce protein expression. After the addition of IPTG, the flasks were kept at 37^{o}C and at 200rpm for 4 hours. Cells were harvested by centrifugation at 4000 x g for 10min, 4^{0}C in 50ml tubes (Falcon) and the pellets were stored at -20^{o}C for protein extraction. Pellets were resuspended in 5 mL of buffer 1 (non-denaturing Lysis Buffer; 50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) containing 1mM of Phenylmethylsulfonyl Fluoride (PMSF) proteinase inhibitor. Cells were sonicated four times for 20seconds, with the tube put on ice for a minute after each sonication. Sonicated samples were spun down at 5000 x g, 30 min, 4^{o}C and the supernatants were used immediately or stored at 4^{o}C not more than two days.

The pellets were collected and resolubilized by adding 1 mL of SDS buffer, followed by boiling on a heating block for 10 minutes. The SDS buffer was prepared by combining 25 mL of 0.5 M Tris-HCl (pH 6.8), 50 mL of glycerol, 4 g of SDS, and 5 mg of bromophenol blue, then adding deionized water to a final volume of 100 mL. An SDS-PAGE gel or Western blot was run to check if the proteins were in solution before purification on the His-column. Recombinant proteins were purified using His-select nickel magnetic agarose beads (Sigma Aldrich) following manufacturer's instructions.

4.2.6. Yeast strains, media and growth conditions

Yeast strain MT-8 (MT-8 MATa, SUP7,ura3-1, his5-2, leu2-3,112, ade2-1, trp1, lysi-1, lys2-1, can1-100, mod5: TRP1) (Pratt-Hyatt et al., 2013) was provided by Dr R.J. Maraia, National Institutes of Health, USA on approval from Dr A.K. Hopper, Ohio state university, USA. MT-8 strain does not grow on media lacking adenine because it is unable to synthesise adenine due to a disrupted tRNA-IPT gene and a premature UAA stop codon in the ADE2 gene. Due to the accumulation of adenine biosynthesis intermediates (Bharathi et al., 2017), the strain forms red colonies when grown on rich media. Introduction of a functional tRNA-IPT gene into MT-8 enables SUP7 to suppress the ADE2 mutation by inserting a tyrosine at UAA codons, restoring adenine biosynthesis and leading to growth of white colonies. MT-8 strain was maintained on yeast extract peptone dextrose (YPD) medium. YPD medium was prepared by adding 1 g yeast extract, 2 g peptone and 3g of agar in 100 mL of deionised water. Synthetic dropout media (SD) for Leucine selection was prepared by adding 0.16 g of Yeast Synthetic Drop-out Medium Supplement without Leucine and 0.17g of yeast nitrogen base without amino acids (YNB) in 100 mL (3 g of agar were added for solid agar plates). Synthetic dropout media (SD) for Adenine selection was prepared by adding 0.16 g of Yeast Synthetic Drop-out Medium Supplement without histidine, leucine, tryptophan, and adenine, 0.17g of YNB, 76 mg of Histidine, 380 mg of leucine and 76 mg of tryptophan amino acids in 100 mL (3 g of agar were added for solid agar plates). The media were cooled to approximately 60°C after autoclaving and 2% filter sterilised glucose was added to the media as source of carbon. Yeast cells were grown at 30°C.

4.2.7. Yeast complementation assay

Yeast expression vector pYX-242 and pYX-242_TRIT1 (Khalique *et al.*, 2020) were generously provided by Dr R.J. Maraia, National Institutes of Health, USA. Gene constructs of PbIPT1, PbIPT2 and Delta_PbIPT1 were cloned in the NcoI and XhoI sites of the pYX-242 vector.

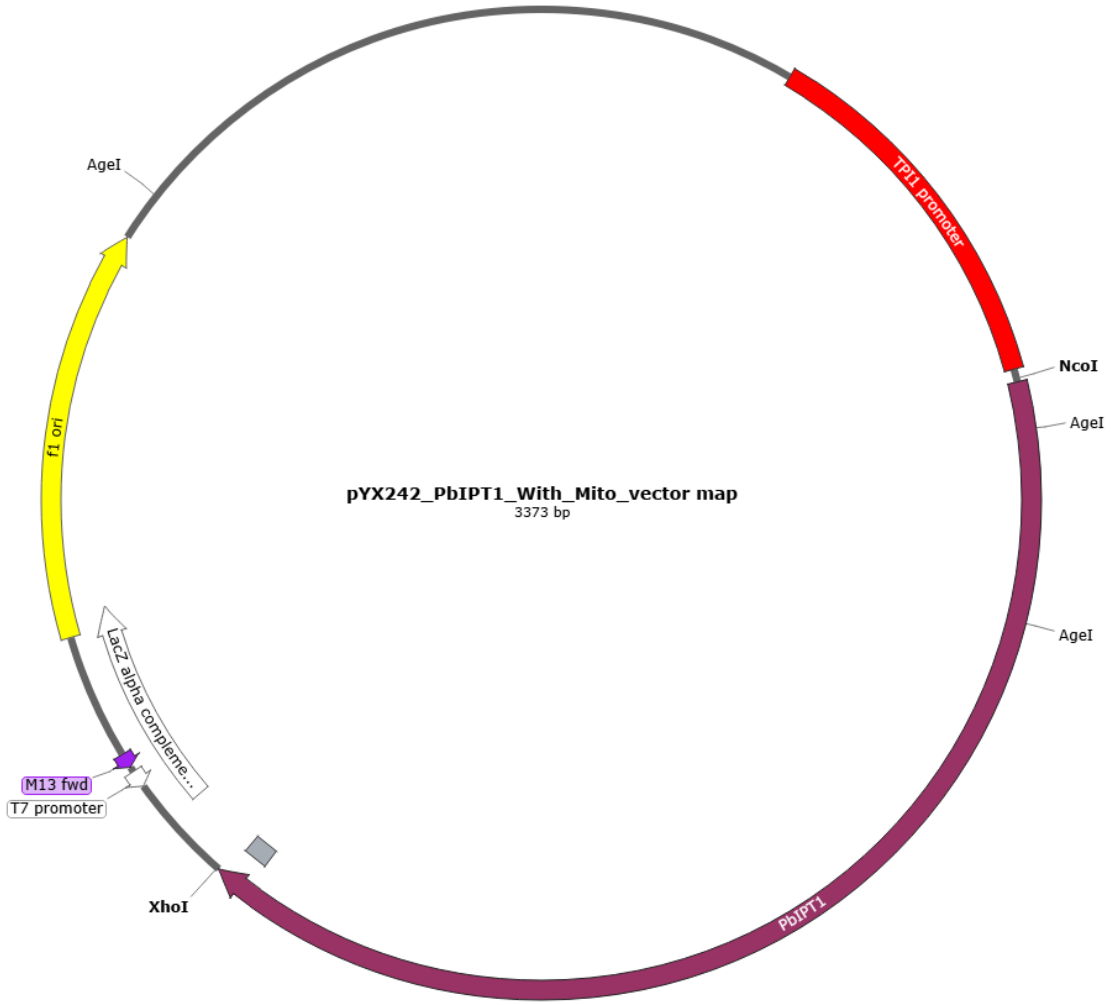


Figure 35: Portion of pYX242 vector map showing PbIPT1 cloned.
The full vector sequence and map was not provided. This vector map sequence was obtained by sequencing the vestor plasmid using T7 promoter primer together with forward and reverse primers of TRIT1 gene that was already cloned in the plasmid.

All gene constructs were transformed into MT-8 yeast strain by by the LiAc/SS Carrier DNA/PEG Method (Gietz & Woods, 2006). MT-8 strain was used for tRNA-IPT complementation assay as described in Hinsch *et al.,* (2016). Selection of plasmid expression was done using selective Leucine SD media. pYX-242 empty vector and pYX-242_TRIT1were used as negative and positive controls respectively. A gene construct complementation was

reported only on the gene constructs that were able to restore white colony colour on YPD media and growth on adenine selective media.

4.2.8. Arabidopsis atipt2, atipt9 and atipt2,9 complementation

Plant expression vectors pB7FWG2 (Karimi *et al.*, 2002) and pGWB508 (Nakagawa *et al.*, 2007) were provided by Prof A. Fleming and Prof J. Ton, University of Sheffield, UK. Constructs cloned in pB7FWG2 vector produce a C-terminal GFP translational fused protein and the ones in pGWB508 produce a C-terminal HIS-tag translational fused protein. PbIPT1 and PbIPT2 were cloned in plant expression vectors using gateway cloning method (Invitrogen). Plasmid constructs were transformed by electroporation into *Agrobacterium* strain C58C1. 5ml starter culture for each construct was grown overnight at 28^0C and 220rpm. Overnight cultures of each construct were inoculated into 200 mL of Luria-Bertani media containing appropriate antibiotics, then grown at 28°C and 220 rpm for four hours. The bacterial cultures were subsequently centrifuged at 4,000 rpm for 30 minutes at 4°C to pellet the cells. The *Agrobacterium* transformation mixture was prepared by resuspending the pellets in 100ml of transformation buffer (0.22g MS (M5519-50L, Sigma), 5g Sucrose, 1ml MES (25.6mM), 50µl silwet L-77). Plants were transformed with floral dip method (Clough & Bent, 1998) with minor modification. The *Agrobacterium* transformation mixture was poured into small beaker or large weighing boat and plants were dipped into the Agrobactrerium culture with all the flowers being submerged. Plants were gently moved around into the media for 1-2 minutes. After transformation, plants were gently removed from the media and allowed to drain for a few seconds over the culture. Opened flowers and siliques were cut off to increase the proportion of transformed seeds. Dipped plants of the same construct were placed horizontally in the tray lined with moist tissue and covered with tray covers to increase humidity. Plants were moved to a growth chamber, and they were uncovered and placed upright after 24 hours. Plants were moved to growth chamber with 16 hours lighting photoperiod. To identify transformed seeds, individual seeds transformed with pGWB508 were surface sterilized using chlorine gas as described above and were plated on plates containing 15 mg/L hygromycin. Hygromycin resistant plants were selected using a protocol for selecting transformants following floral dip transformation(Harrison *et al.*, 2006). Transformants were transplanted to pots and the presence of transgene was further confirmed by PCR. Transgenic plants, generated by transforming plants with the pB7FWG2 vector carrying the glufosinate-ammonium resistance marker, were selected by spraying a 1:1000 dilution of 150 g/L Basta onto soil-grown seedlings that were 7 days old. Transgenic plants expressing constructs in pB7FWG2

expression vectors were analysed under fluorescence stereomicroscope (Leica M165 FC) microscope for the expression of GFP. T3 transgenic plants in atipt2, atipt9 and atipt2,9 backgrounds were analysed for the complementation of cis-zeatin biosynthesis, root and shoot phenotypes.

4.2.9. Extraction and quantification of CK of Arabidopsis plants overexpressing PbIPTs

Table 9: Genotypes used for cytokinin quantification

Vector	*P.brassicae* IPT	Salk ID	At mutant background	Genotype
pGWB508	PbIPT1	SALK042050C	atipt2	PbIPT1_Line 4-8
pGWB508	PbIPT1	SALK042050C	atipt2	PbIPT1_Line 6-2
pGWB508	PbIPT2	SALK042050C	atipt2	PbIPT2_Line 13-2
pGWB508	PbIPT2	SALK042050C	atipt2	PbIPT2_Line 17-8
pGWB508	PbIPT1	SALK128927C	atipt9	PbIPT1_Line 6-2
pGWB508	PbIPT1	SALK128927C	atipt9	PbIPT1_Line 11-2
pGWB508	PbIPT2	SALK128927C	atipt9	PbIPT2_Line 12-3
pGWB508	PbIPT2	SALK128927C	atipt9	PbIPT2_Line 14-6
pB7EWG2	PbIPT1	atipt2,9	atipt2,9	PbIPT1_Line 3-5
pB7EWG2	PbIPT1	atipt2,9	atipt2,9	PbIPT1_Line 4-5
pB7EWG2	PbIPT2	atipt2,9	atipt2,9	PbIPT2_Line 3-9
pB7EWG2	PbIPT2	atipt2,9	atipt2,9	PbIPT2_Line 4-4
control		SALK042050C	Aatipt2	wild type
control		SALK128927C	atipt9	wild type
control		atipt2,9	atipt2,9	wild type
control		col_0	col_0	wild type

Seeds from the genotypes listed in table 9 were sterilized , stratified and grown in growth chambers as described in the plant material and growing conditions section above. Plates were moved to growth chamber, placed upright to ensure that the roots grew on the surface of the agar and the seedlings were left to grow for 9 days after stratification. Roots were excised from the seedlings with sterile surgical blade and their fresh weight (10-15mg) measured before being immediately frozen in liquid nitrogen. Samples were freeze-dried with a vacuum and dried samples were sent for cytokinin quantification at Laboratory of Growth Regulators, Palacky University & Institute of Experimental Botany AS CR, Czech Republic.

4.2.10. Root and shoot phenotyping

To measure primary root length, 50 seeds from each transgenic line in the *atipt9* and *atipt2,9* backgrounds, as well as their respective wild-type controls, were sown on antibiotic-free agar plates. The plates were kept vertically with a small inclination angle to allow root growth on

the agar surface for 7 days after stratification. Pictures of the plates with seedlings were taken in the lab with a camera mounted on a stand. Root length of seedlings were measured using FIJI software (Schneider *et al.*, 2012).

4.2.11. Subcellular localisation of PbIPTs

For observation of the subcellular localization of PbIPT1 and PbIPT2, transgenic lines expressing PbIPT1 and PbIPT2 fused to GFP were used. MitoTracker Red CMXRos (M7512) was used for staining mitochondria while propidium iodide (PI) was used as a counter stain for cytoplasmic localisations. MitoTracker Red stain was freshly made by dissolving 50µg of lyophilised MitoTracker Red in 94µl of dimethyl sulfoxide (DMSO). Five-day-old roots were harvested, submerged in 1 mL of half-strength MS(pH 5.7) media containing 1µL of the 1mM MitoTracker Red stock solution, and vacuum-infiltrated for 30 minutes. Propidium iodide staining solution was prepared by diluting 0.1 µL of 20 mM stock solution into 1 mL of water. Roots were stained by incubating them in 1 mL of half-strength MS media containing 0.1 µL of 2 mM PI for 5 minutes at room temperature. MitoTracker red and PI stained roots were mounted on microscope slides and were visualized by confocal microscopy with a Zeiss LSM880 AiryScan Confocal microscopy. The microscope was set to measure eGFP signal with exitation at 488nm and emission at 475 to 522nm. MitoTracker red was measured with excitation of 579nm and emission of 599 to 615nm and PI was exited at 535nm and emission was 600 to 620nm. The captured images were processed using FIJI software(Schneider *et al.*, 2012). Consistent fluorescence patterns were observed across multiple plants per construct.

4.3. Results

4.3.1. The *P.brassicae* IPTs genes are predicted to localise to different cellular compartments.

P.brassicae has two IPTs genes, PbIPT1 and PbIPT2. However, the gene models for *P. brassicae* are uncertain. Therefore, due to variations in gene length at the 5' and 3' ends, comparisons of *PbIPT* gene structures were conducted across available *P.brassicae* gene models. PbIPT1 is encoded by PbPT3Sc00040_Am_6.202_1(Rolfe *et al.*, 2016). This gene model is identical to PBRA_004561 from the first annotation of the Pb314 genome (Schwelm *et al.*, 2015), but disagrees with new model, PLBR_LOCUS7342 from the updated genome (Stjelja *et al.*, 2019). The gene models PbPT3Sc00040_Am_6.202_1 (PBRA_004561) appear to be correct, the model PLBR_LOCUS7342 has a longer 5' end but this is not supported by both RNASeq and long read data (Figure 14). PbIPT1 does not contain introns and is likely to be mitochondrial - it has a predicted signal sequence (Figure 36a).

PbIPT2 is more problematical. The gene models are PbPT3Sc00024_Sm_2.181_1, PBRA_005817 and PLBR_LOCUS5459. Both PbPT3Sc00024_Sm_2.181_1 and PLBR_LOCUS5459 start at the same position and have introns in the same position, the difference is at the 3' end. Inspection shows that long reads do not support PbPT3Sc00024_Sm_2.181_1 whereas the support for PLBR_LOCUS5459 is good (Figure S1), although the long read depth of coverage is low. This is supported by BLASTP analysis as the predicted protein from PbPT3Sc00024_Sm_2.181_1 matches very well over first 340 amino acids, but then has no match with any other tRNA IPT. In contrast, the predicted protein from PLBR_LOCUS5459 matches very well over its entirety with numerous other IPTs. Therefore, the PLBR_ LOCUS5459 is accepted as correct. PbIPT2 does not have a predicted signal peptide and is therefore likely to be cytosolic (Figure 36b).

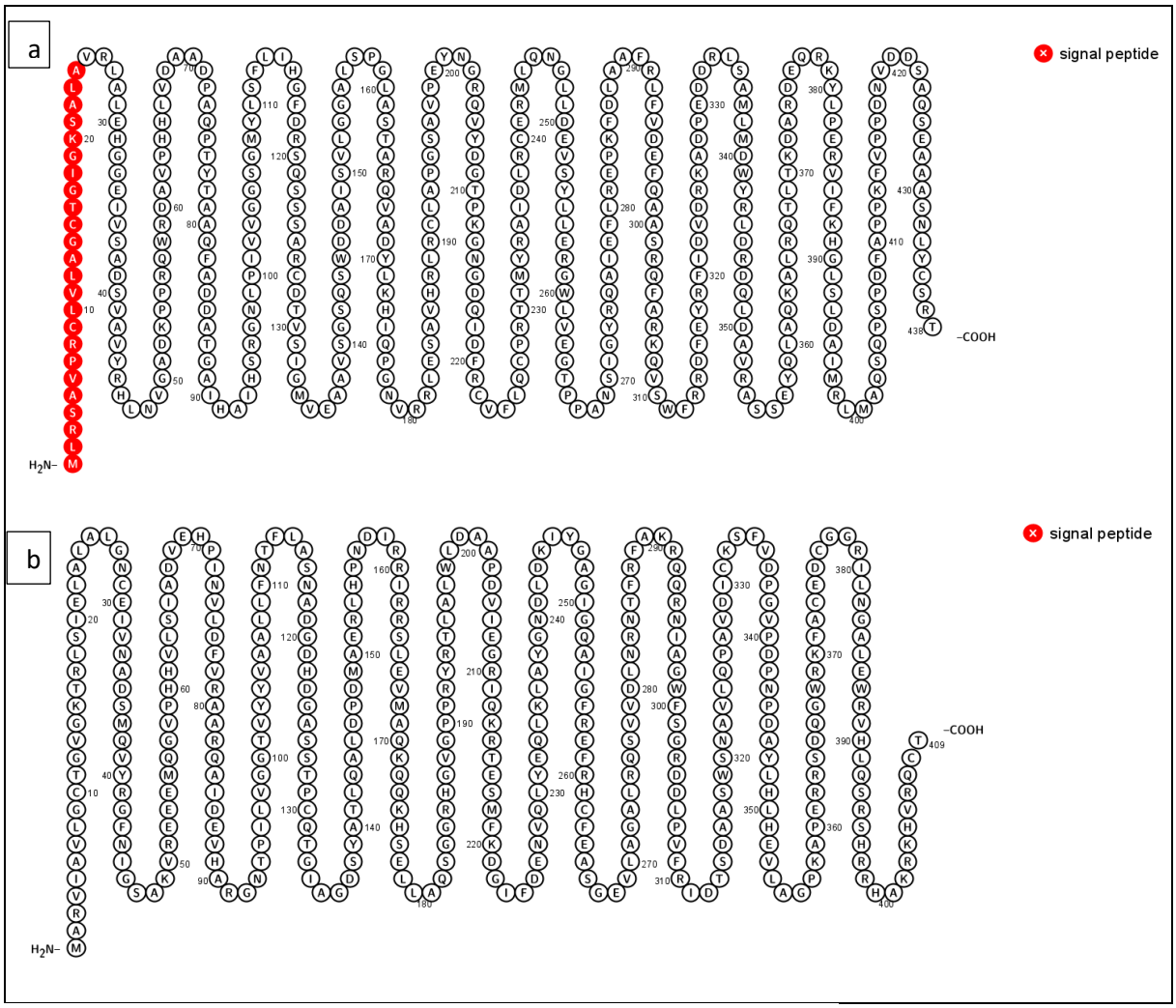


Figure 36:Predicted signal peptide sequences on P.brassicae IPTs.
a) PbIPT1 and b) PbIPT2. Images created in Protter with signal peptide determined using Phobius. The signal peptide on PbIPT1 was predicted to be mitochondrial by DeepMito.

To validate the subcellular localization of PbIPTs, 35S::PbIPT1-GFP and 35S::PbIPT2-GFP constructs were generated and individually overexpressed in *Arabidopsis thaliana* atipt2,9 double mutant background. As expected, PbIPT1-GFP was localised in the mitochondria (Figure 37) and PbIPT2-GFP was found to localise in the cytoplasm but was also found in the nucleus. Nuclear and cytoplasmic localisation of PbIPT2 was further confirmed by using propidium iodide (PI) as a counter stain (Figure S4).

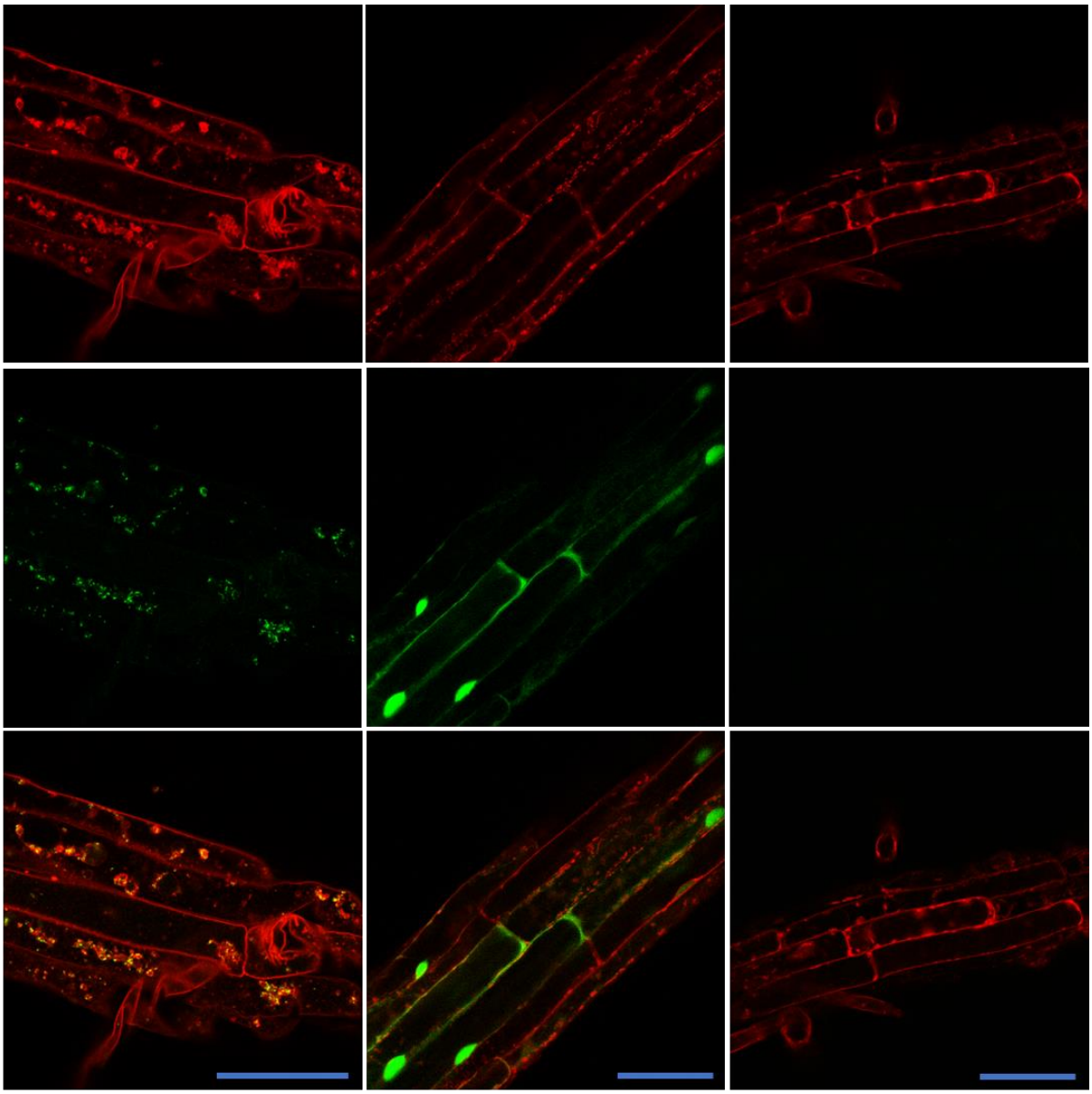


Figure 37: Subcellular localization of PbIPT1 and PbIPT2 in Arabidopsis thaliana root. Representative confocal pictures of 35S::PbIPT1-GFP, 35S::PbIPT2-GFP and atipt2,9 plants. Top panel is Mito-tracker red signal, middle panel is GFP signal, and the bottom panel is GFP merged with Mito-track red. The punctate red dots in the upper panels shows mitochondria staining. The green punctate dots in the left middle panel show PbIPT1::GFP expression and the yellow dots shows co-localisation of PbIPT1::GFP and Mito-tracker red in the mitochondria. PbIPT2::GFP shows nuclear and cytoplasmic localisation. Scale bars = 50μm.

4.3.2. Phylogenetic analysis places PbIPT1 and PbIPT2 in same clades with tRNA-IPTs

To explore the phylogenetic relationship of PbIPTs with tRNA-IPTs and adenylate IPTs, phylogenetic analysis of PbIPTs proteins was performed using 66 protein sequences from bacteria, protists, fungi, plants, and animals which have been described as either tRNA-IPT or adenylate IPTs in different studies. This analysis placed PbIPTs into two separate clades of tRNA-IPTs suggesting that *P.brassicae* IPTs are tRNA-specific IPTs.

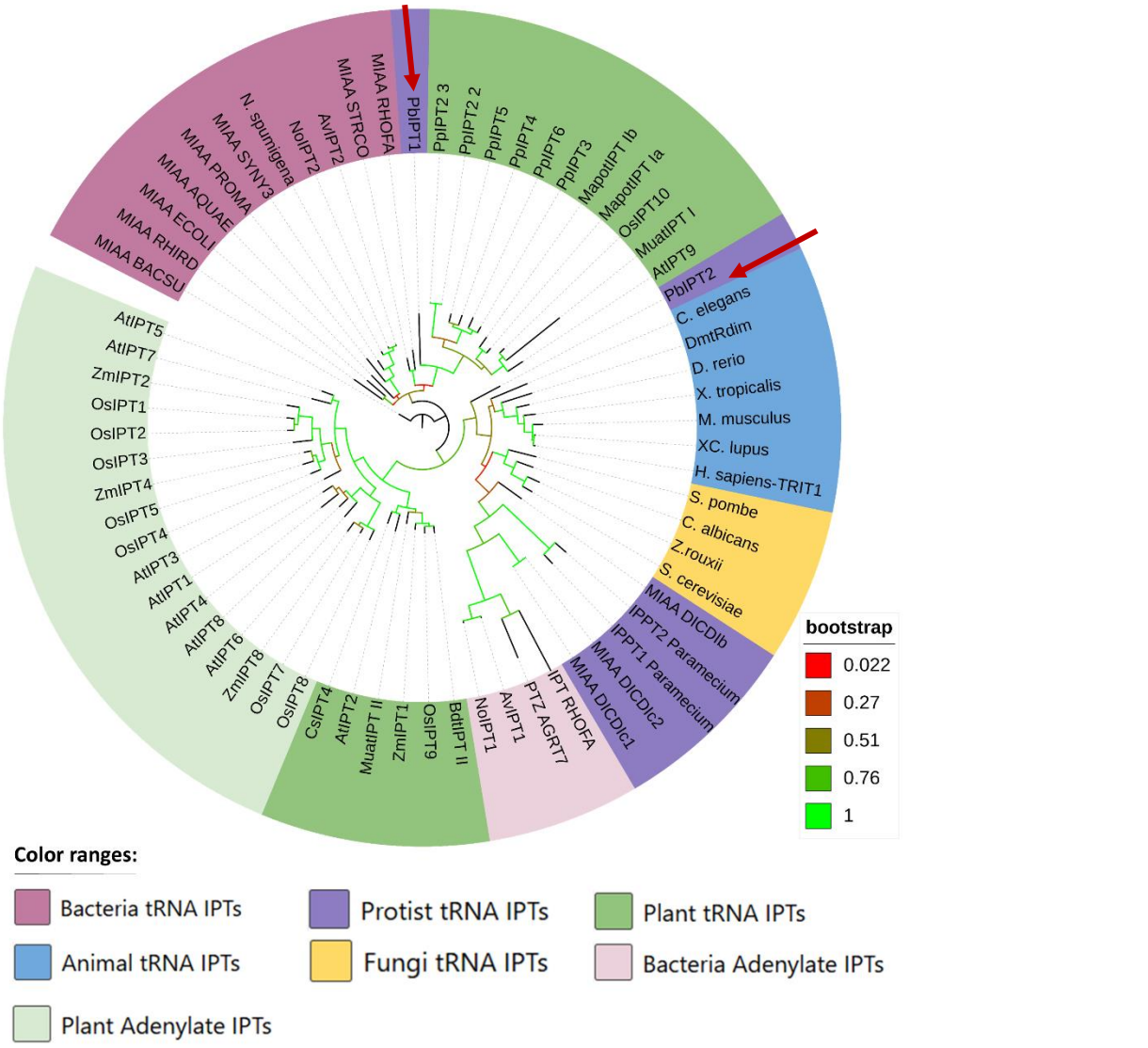


Figure 38: Maximum likelihood Phylogenetic tree of P.brassicae IPTs and other IPTs from different organisms.
P.brassicae IPTs are indicated by red arrows

4.3.3. Predicted protein structures of PbIPTs are similar to known tRNA-IPTs protein structures

Structural features of proteins are important for a protein to carry out a wide range of functions within a cell. To further examine if PbIPTs are tRNA-IPTs, structures of PbIPT1 and PbIPT 2 were predicted using Alphafold. The predicted structures of PbIPTs were aligned with experimentally determined structure of tRNA-IPT protein from E.coli. both PbIPT1 and PbIPT2 had a good alignment (RMS 1.907 A) to the equivalent E coli protein (PDB: 3FOZ). PbIPT2 has the 'arm' (red circle) characteristic of tRNA-IPTs that capture the tRNA molecule.

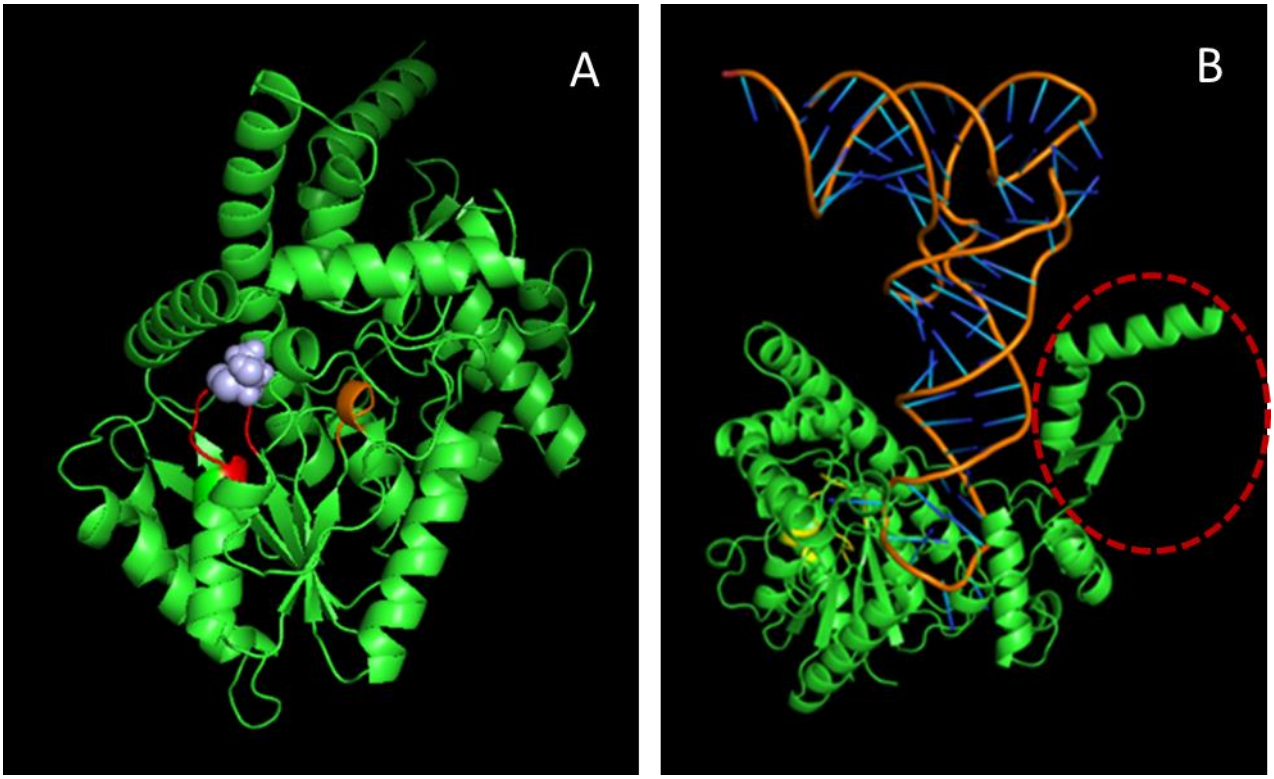


Figure 39: Prediction of PbIPT1and PbIPT2 protein structures using Alphafold.
A) shows predicted protein structure of PbIPT1 with conserved amino acid residues involved in the binding and catalysis of tRNA shown in red, brown and grey colours. B) indicates alignment of PbIPT2 protein to the equivalent E coli protein (PDB: 3FOZ) bound to tRNA. The red circle indicates the characteristic 'arm' of a tRNA-IPT.

Sequence alignment of PbIPTs fungi and animal tRNA IPTs sequences revealed PbIPT2 has conserved residues involved in binding to tRNAs and catalysis (Khalique *et al.*, 2020). These are an invariant threonine in the P-loop and a conserved DMSQ motif involved in interactions with A37 and catalysis (Figure S5). Structural prediction using Alphafold (Fig39A) shows that these are in close proximity to each other. These conserved residues are missing in PbIPT1 (even after checking all possible open reading frames in the gene models). However, note that the 3FOZ structure from E coli (Seif & Hallberg, 2009) does not have the DSMQ (it has DSAL)

(Figure S6) therefore this is not an absolute requirement. Both PbIPT1 and PbIPT2 align to 3FOZ well.

4.3.4. PbIPT1 and PbIPT2 encodes functional tRNA-IPTs

To confirm whether PbIPT1 and PbIPT2 indeed encode a functional tRNA-specific IPT, a yeast-based assay was used. The assay is based on three modifications which cause accumulation of red pigmentation and adenine autotrophy when grown on complete medium yeast extract peptone dextrose (YPD) and adenine deficiency medium respectively. However, when the strain is complemented, the colony colour changes from red to white on complete medium and the strain is able to grow on adenine-deficient medium. PbIPT1, PbIPT2,and TRIT1 constructs were cloned in yeast expression vector pYX-242 and transformed in the MT-8 yeast strain. TRIT1 was used as a positive control whereas empty vector was used as negative control. MT-8 cells transformed with PbIPT2 and TRIT1showed change of pigmentation from red to white and did grow on adenine deficient medium as the positive control (TRIT1) indicating complementation of the strain but cells transformed with PbIPT1 and empty vector in the upper panel did not complement the yeast mutant indicating that it was not functional as shown by the inability to change pigmentation to grow on adenine deficient medium (Figure 40). The inability of PbIPT1 to complement MT-8 phenotype was thought to have been caused by mis-localisation of the protein as PbIPT1 had a predicted mitochondrial localisation signal but PbIPT2 did not. The amino acid residues at the N-terminal of the protein that were predicted to encode the mitochondrial signal peptide were removed to make the protein cytoplasmic ((delta- pYX-242-PbIPT1). MT-8 cells transformed with the deleted version of the protein (delta- pYX-242-PbIPT1 in the lower panel) proliferated on adenine deficient medium and showed a loss of red pigmentation on complete medium indicating that the deleted version of the protein which is cytoplasmic localised complemented the strain. Taken together, both *P.brassicae* IPTs were confirmed to be tRNA-specific IPTs.

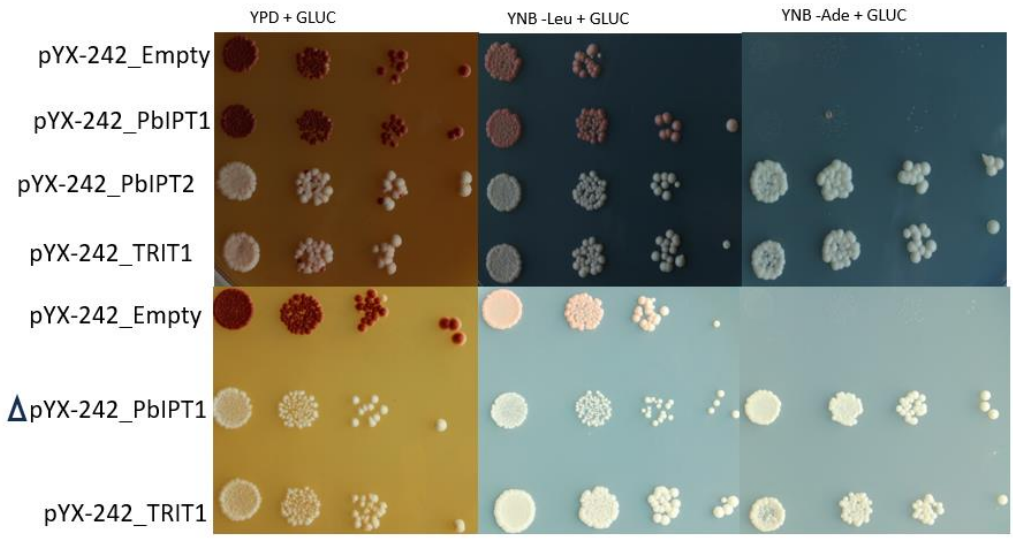


Figure 40: Complementation of MT-8 yeast strain with PbIPT genes.
Different yeast lines were grown on control complete medium YPD and on selective Synthetic Defined medium (YNB) lacking Leucine (YNB -Leu) and adenine YNB–Ade). pYX-242-Empty vector was used as a negative control, pYX-242-TRIT1 as positive control, PbIPT1 and PPbIPT2 are putative P.brassicae tRNA-IPTs. delta- pYX-242-PbIPT1 is PbIPT1 without the first 26 amino acids of N-terminus. The assay is based on the strain MT-8 which, due to a disrupted tRNA-IPT gene and a premature UAA stop codon in the ADE2 gene, cannot synthesize adenine, resulting in a red colony and adenine auxotrophy phenotype. Introduction of a functional tRNA-IPT gene into MT-8 enables SUP7 to suppress the ADE2 mutation by inserting a tyrosine at UAA codons, restoring adenine biosynthesis and leading to white colony growth

4.3.5. CK species produced by *P.brassicae*

P.brassicae is believed to produce cytokinin to influence host metabolism and development during host colonisation. To determine the types of CK compounds synthesized by *P.brassicae* IPT genes, pathogen tRNA-IPTs were expressed in *E. coli* cells to produce proteins for in vitro enzyme assays. While PbIPT proteins were successfully expressed in E. coli (Figure S7), purification using anti-His beads was only successful for the MiaA positive control. Subsequently, the PbIPTs genes were overexpressed in Arabidopsis thaliana atipt2, atipt9, and atipt2,9 cis-zeatin biosynthesis mutant backgrounds. Root samples from 7 days old seedling were sent for total CK content quantification to a collaborator in Czech Republic but the results are not yet available.

4.3.6. *P.brassicae* infection promotes Axillary bud outgrowth

To determine the influence of *P.brassicae* infection on host development, *Arabidopsis thaliana* Col-0 plants were inoculated with *P.brassicae* . Compared with control plants, growth retardation became visible from 10 dpi. In later stages of infection, infected plants exhibited

axillary bud outgrowth at the base and within the leaf rosettes (Figure 41). These axillary buds develop as shoots.

Figure 41: Shoot phenotype of P.brassicae-infected Arabidopsis thaliana Col-0 30dpi. Infected plants develop multiple axillary buds at the base of leaf rosette later during the infection before the root system start to rot.scalebar=1cm

Overexpression of PbIPT1 and PbIPT2 in Arabidopsis thaliana atipt2 and 9 single mutants also affected axillary bud outgrowth. Some of the plants overexpressing PbIPT1 and PbIPT2 had two or more additional stems that developed into inflorescences (Figure 42). However, the axillary bud outgrowth as leafy shoots observed at the base and within leaf rosettes of *P.brassicae* infected Col-0 were not found in PbIPTs overexpressing plants.

Figure 42: Multiple inflorescence stems on atipt2 and atipt9 single mutants over expressing P.brasicae PbIPTs.
The picture shows one of atipt9-PbIPT1 transgenic plants with two stems developing side by side.

4.3.7. PbIPT1 may perform tRNA-IPTs canonical functions

tRNA-IPTs are ubiquitously found in all living organisms where they play an important role in protein synthesis. In *Arabidopsis thaliana*, the *atipt2,9* double mutant, lacking both tRNA-IPT genes, exhibits pale green leaves and short root phenotypes. While the root phenotype is not rescued by CK treatment, it is restored by reintroducing the ATIPT9 gene into the double mutant (Antoniadi *et al.*, 2022). This suggests that ATIPT9 has functions beyond its involvement in the synthesis of *cis*-zeatin. To find whether PbIPTs have other functions beyond CK synthesis, PbIPTs genes were overexpressed in atipt2,9 double mutant individually.

Overexpression of PbIPT1 in atipt2,9 double mutant but not PbIPT2 restored the shoot phenotype (Figure 43).

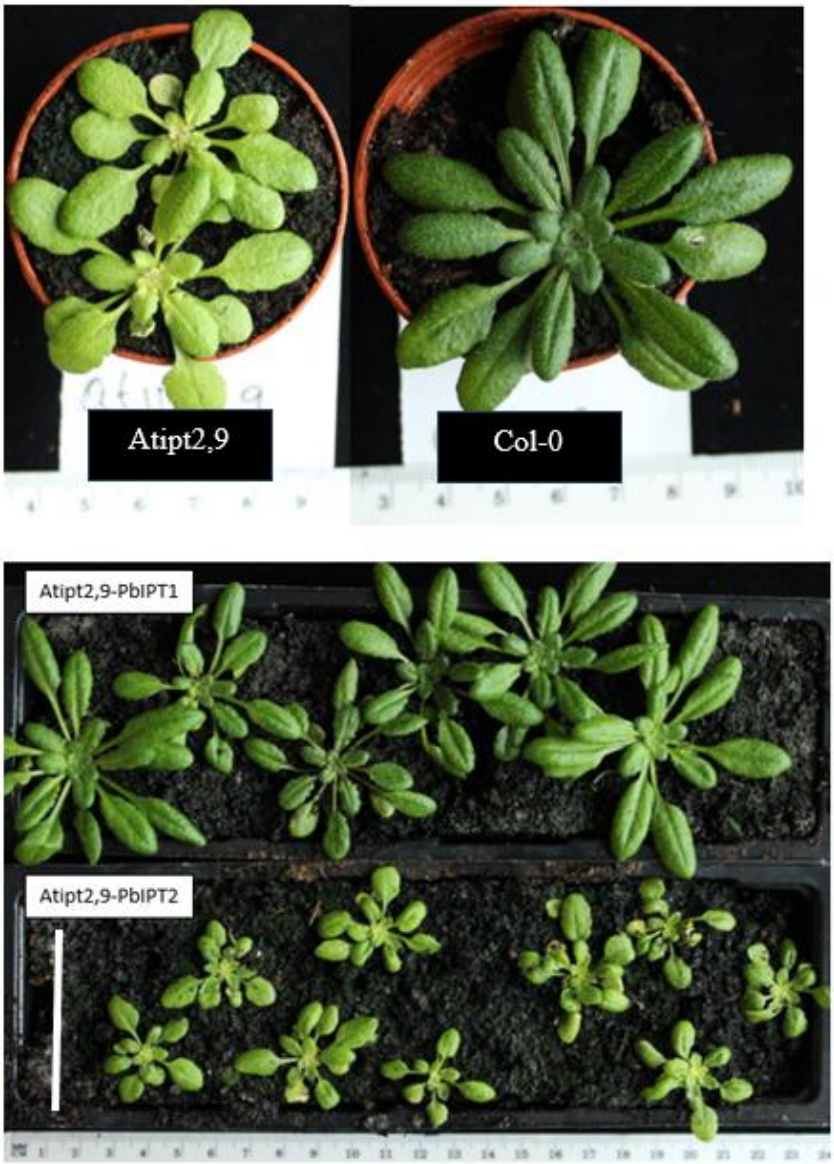


Figure 43: Shoot phenotypes of independent T1 lines expressing P.brassicae IPT genes. The parental plant (T0) was transformed by agrobacterium-mediated floral dip method. Seeds from each T0 plants were harvested and kept separately. Seeds were germinated on soill, sprayed with Basta and only transgenic seedling were able to grow and wild type seedling dried out.

One PbIPT1-overexpressing line exhibited increased primary root length (Figure 44a), although not to wild-type levels (Figure 44b). Another PbIPT1-overexpressing line, atipt2,9-PbIPT1_4_5_1, displayed a primary root phenotype similar to both the *atipt2,9* mutant and the PbIPT2-overexpressing lines (Figure 44b). This was observed despite atipt2,9-PbIPT1_4_5_1 exhibiting a restored leaf phenotype, resembling that of wild-type plants.

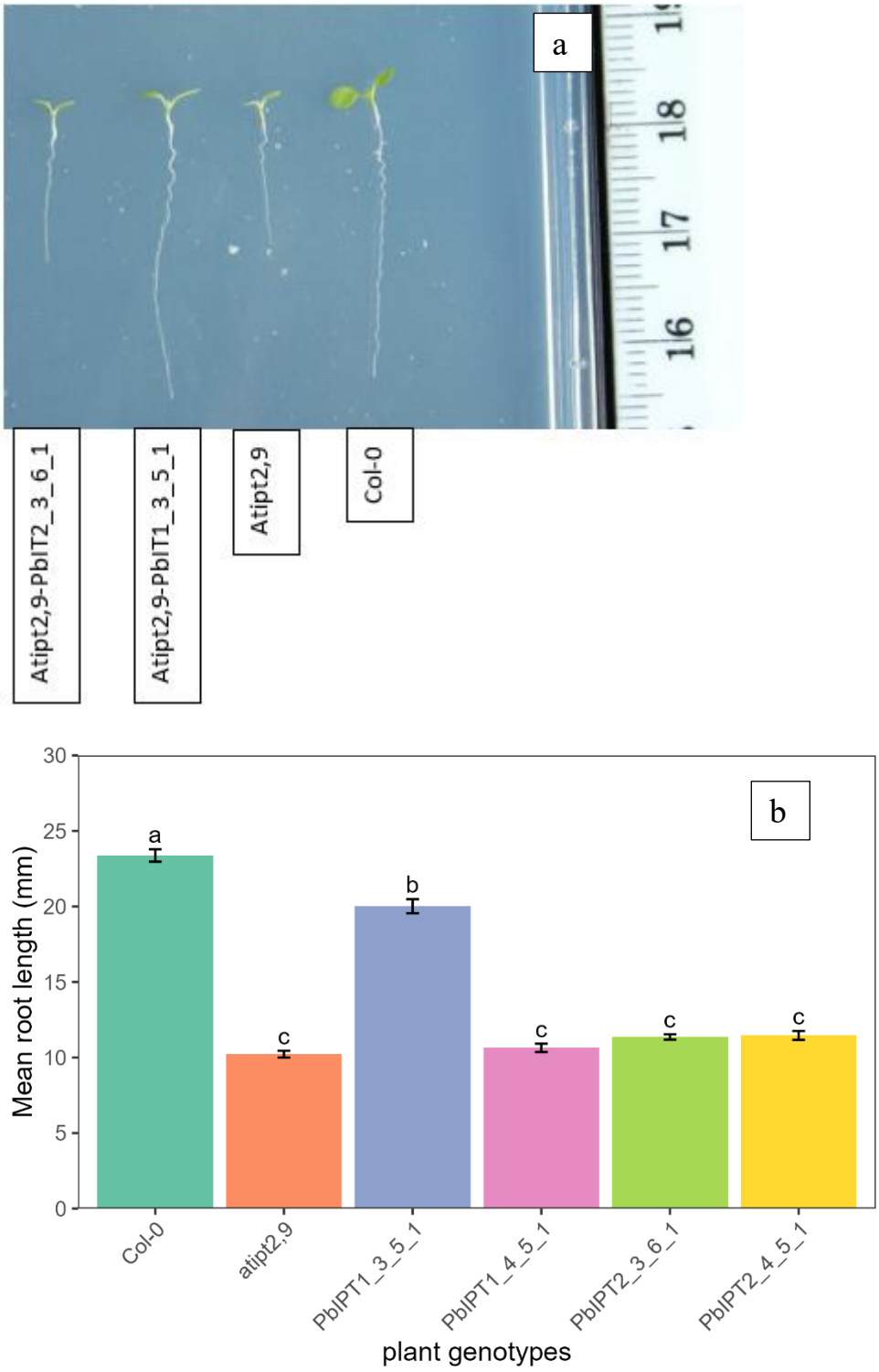


Figure 44: Complementation of atipt2,9 root phenotype by 35S::PbIPTs-eGFP overexpression. a) primary root phenotype of wild-type Col-0, atipt2,9, atipt2,9-PbIPT1_3_5_1and atipt2,9-PbIPT2_3_6_1 seedlings 12 days after sowing. (b) primary root length quantification of the wild-type Col-0, atipt2,9, atipt2,9-PbIPT1_3_5_1, atipt2,9-PbIPT1_4_5_1, atipt2,9-PbIPT2_3_6_1and atipt2,9-PbIPT2_4_5_1 seedlings 12 days after sowing. Error bars show standard error of the means and genotypes with the same letter indicates that their primary root length did not differ significantly (Tukey test, p>0.05), n=50.

The lines atipt2,9-PbIPT1_3_5_1 and atipt2,9-PbIPT1_4_5_1, both derived from the same parental plant (Figure 43), represent distinct outcomes of independent *Agrobacterium*-mediated transformation events. While the exact cause of the differing root phenotypes in these PbIPT1expressing lines is unclear, the observation that PbIPT1 unlike PbIPT2 restores both root and shoot phenotypes suggests that PbIPT1 may have a function beyond cytokinin synthesis.

4.4. Discussion

Plant hormones play important roles in regulation of processes such as growth, nutrient transport and interaction with their environment and thus plant hormonal pathways are often targeted by plant pathogens to promote colonisation and disease development (Vañó *et al.*, 2023). Plant pathogens manipulate host hormonal homeostasis by targeting or secreting enzymes involved in the target hormone metabolism or by producing hormonal mimicking compounds (Chanclud *et al.*, 2016). The uncontrolled cell division and cell enlargement which leads to the formation of galls on infected plants is thought to involve auxins and cytokinins but how *P.brassicae* manipulate the host hormone homeostasis remains unclear (Ludwig-Müller *et al.*, 2009). The *P.brassicae* genome contains two isopentenyl transferases which are key enzymes in the biosynthesis of cytokinins but also have other functions beyond CK biosynthesis. The substrate specificity of these IPT enzymes, the type of CK they produce, the effect of these *P.brassicae* produced CK and PbIPTs mediated tRNA modifications in the infection process remains unclear. Thus, this study was undertaken to determine the type of *P.brassicae* IPT enzymes and the role of *P.brassicae* CK in clubroot disease development.

4.4.1. *P.brassicae* IPTs belong to tRNA-IPT type

Phylogenetic analysis, alongside structural predictions, placed *P.brassicae* IPTs within the tRNA IPT family(Figure 38 and 39). This classification is supported by the presence of conserved motifs essential for tRNA recognition, binding, and catalysis in both PbIPTs, as identified in other tRNA-IPTs(Soderberg & Poulter, 2001; Hinsch *et al.*, 2016; Khalique *et al.*, 2020). The tRNA IPT enzymes have the GxxxxGKT(S) element that resembles the P-loop of NTP-binding proteins with an invariant threonine in the P-loop and a conserved DMSQ motif (Figure S5) involved in interactions with A37 and catalysis (Soderberg & Poulter, 2001; Khalique *et al.*, 2020). Both PbIPT1 and PbIPT2 have amino acid residues which are conserved in the Mia A tRNA IPT of E.coli except S21 and R306 corresponding to E.coli Mia A T24 and K280 which were conservatively substituted in PbIPT1 (Figure S6). The Mia A arginine (R170) was also non-conservatively substituted by a serine (S184) in PbIPT1. Conserved residues in Mia A are K23, T24, H67, and R217 which have been proposed to be involved in DMAPP binding, K56, R167, R170, and K280 for tRNA binding and T19 and Y47 for stabilization of developing positive charge on the dimethylallyl group during the alkylation reaction (Soderberg & Poulter, 2001).

Yeast complementation assay indeed proved that PbIPTs are tRNA IPT type as both gene constructs were able to restore adenine auxotrophy and white colony colour (Figure 40). tRNA-IPTs involved in the modification of cytoplasmic and mitochondrial tRNAs can be encoded by one or multiple genes. In yeast *S. cerevisiae*, a single tRNA-IPT MOD5 encodes both cytoplasmic and mitochondrial enzymes by having two in-frame translation initiation sites in the 5' end of the MOD5 mRNA (Gillman *et al.*, 1991). Like MOD5, the PbIPT1 gene model of the reference genome used in this study had a long 5'mRNA with two in frame translation initiation sites suggesting that PbIPT1 could encode tRNA-IPT with two different subcellular localisations. However, the PbIPT1 gene model of the reference genome was incorrect as it was not supported by full length/long reads mRNA sequencing data (Figure 14).

The short PbIPT1 supported used in this study had a predicted mitochondrial localisation signal in the N-terminus whereas PbIPT2 did not (Figure 36). The subcellular localisation of PbIPT1 and PbIPT2 was visualised by expressing 35S::PbIPT1::GFP and 35S::PbIPT2::GFP constructs in the roots of Arabidopsis thaliana atipt2,9 double mutant. The Arabidopsis plants expressing 35S::PbIPT1::GFP construct had GFP signal detected in the mitochondria (Figure 37) confirming mitochondrial localisation of PbIPT1. Plants expressing 35S::PbIPT2::GFP construct had GFP signal in the cytoplasm and the nucleus but not in the mitochondria (Figure 37 and Figure S4) indicating that PbIPT2 is localised in the cytoplasm and nucleus. The subcellular localisation of PbIPT1 and PbIPT2 are similar to that of MOD5 where the long protein which like PbIPT1 has a mitochondrial targeting sequence localise predominantly to mitochondria and the shorter protein which like PbIPT2 lacks the mitochondrial targeting sequence localises in cytosol and nucleus (Gillman *et al.*, 1991; Boguta *et al.*, 1994; Tolerico *et al.*, 1999). The subcellular localisation of PbIPTs is consistent with the role of tRNA-IPTs that catalyse the isopentenylation of ***N6*** of adenosine of both cytoplasmic and mitochondrial tRNAs that decode UNN codons at position 37 (i6A37) directly 3' to the anticodon loop (Khalique *et al.*, 2020).

4.4.2. *P.brassicae* produced CK influence host development

P.brassicae infected plants developed multiple shoots within and below the leaf rosettes in the late stages of infection. It is unclear why *P. brassicae* induces the formation of shooty outgrowths. These outgrowths act as new nutrient sinks, potentially competing with the sporulating pathogen. However, it can be argued that these newly formed shoots might develop into source tissues, maintaining photosynthesis in infected plants and ensuring a continuous supply of carbohydrates for the growing *P. brassicae* plasmodia in the roots. This outgrowth

could be because of accumulation of cytokinins in the upper parts of infected plants (Cardenas-Aquino *et al.*, 2022). However, how *P.brassicae* affects host cytokinin homeostasis remains unknown. Some studies have reported the repression of host CK metabolism machinery and low CK contents in infected plants (Siemens *et al.*, 2006; Malinowski *et al.*, 2016) but others reported increase in some CK species in infected plants (Dekhuijzen & Overeem, 1971; Devos *et al.*, 2006). Blicharz et al (2025) found that phloem sap collected from leaves of *P.brassicae* infected *B. napus* had increased isopentyl-type cytokinins between 10 and 14dpi but their concentration decreased as the disease progressed. The disparities in the reported CK contents of *P.brassicae* infected plants can be attributed to different CK analysis methods used and more importantly on sampling timepoints and samples used for the CK quantification. CK quantification on samples collected in the early stage of infection shows CK increases (Devos *et al.*, 2006) whereas those CK measurements on samples collected in the later stages of disease cycle shows CK decline (Malinowski *et al.*, 2016). The early CK increases reported by Devos et al (2005) is controversial due to inconsistencies in sample fresh weights, as hormone content was normalised per gram of fresh weight while samples were collected based on the number of plants thus variations in the initial sample fresh weight of control and infected plants. Furthermore, host cytokinin metabolism is repressed as early as 7dpi (Siemens *et al.*, 2006; Malinowski *et al.*, 2016). On the other hand, the report of increased CK in the leaf phloem sap could be true as CK measurements from whole plant or roots and hypocotyls which have more healthy cells will mask the subtle changes that may be detected if measurements were done on isolated infected tissues or cells. additionally, all CK quantification studies in *P.brassicae* infected plants have been made on roots and hypocotyls (Devos *et al.*, 2006; Malinowski *et al.*, 2016) and there is lack of information on CK contents in the leaves.

The general trend of host CK metabolism is that of decline but how *P.brassicae* infection triggers host responses like those of increased CK content is unclear. One way that *P.brassicae* may be able to induce these CK responses may be that *P.brassicae* is synthesising and releasing CK in the host. Since the host CK biosynthesis and degradation machinery is repressed, CK buildup because of continuous release of small CK concentrations by *P.brassicae* will reach the level they can activate host CK signalling cascades (Malinowski *et al.*, 2016). This, however, could lead to increased CK contents in infected plants which is not the case. *P.brassicae* may be producing a spectrum of CKs which are active in plants but are not detected in CK measurements of infected plants because cytokinin measurement in *P.brassicae* infected plants have been focused on the four active cytokinins: isopentenyladenine (iP), trans-zeatin

(tZ), cis-zeatin (cZ), dihydrozeatin (DZ) and their conjugates (Devos *et al.*, 2005; Devos *et al.*, 2006; Malinowski *et al.*, 2016). Like *P.brassicae*, *Rhodococcus fascians* bacteria that causes formation of leafy galls on infected plants was reported to either maintain or decrease host CK levels during symptom development (Galis *et al.*, 2005; Depuydt *et al.*, 2008) but it was found that *R. fascians* produces a spectrum of CKs that are able to activate host CK signalling cascades (Pertry *et al.*, 2009) and thus postulating that the shooty gall phenotype develops due to a continuous local secretion of less active, less toxic and synergistic action of different *R. fascians* produced CK.

P.brassicae produced CK could therefore be responsible for the induction of expression of cytokinin responsive genes in the roots and hypocotyls of infected plants thus mediating long-distance coordination of host physiological and developmental responses (Blicharz *et al.*, 2025). Based on the findings in this study, it is reasonable to propose a model in which *P.brassicae* likely modulates host metabolism and development programs by hijacking the host cytokinin signalling pathways (Figure 45). CK biosynthesis occurs in both above and belowground plant parts (Matsumoto-Kitano *et al.*, 2008) but their physiological roles may depend on the temporal and spatial differentiation of their contents and the direction of the signal (Sakakibara, 2003; Hirose *et al.*, 2008). In Arabidopsis tZ and tZ-type CKs are synthesised in the roots and transported via xylem whereas IP-type and cZ are the main CK species in the phloem (Takei *et al.*, 2001; Hirose *et al.*, 2008; Matsumoto-Kitano *et al.*, 2008).

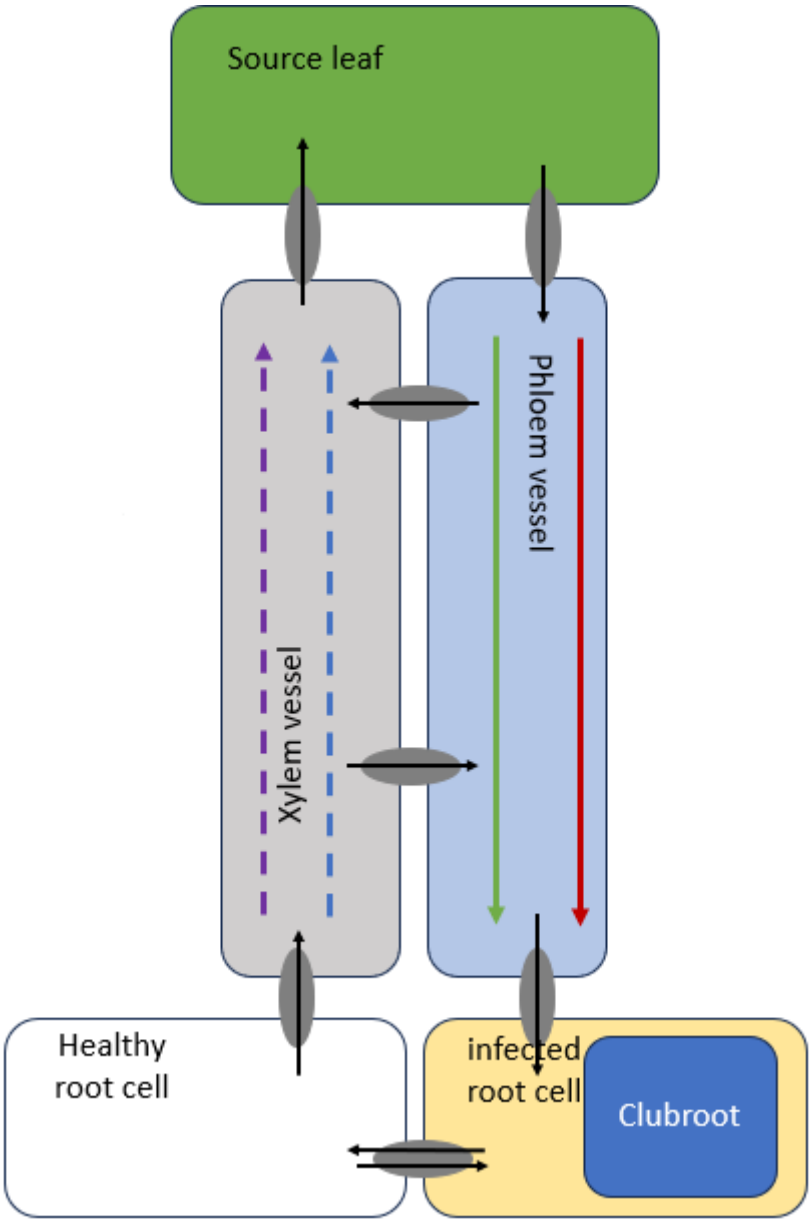


Figure 45: Proposed model of CK signalling during P.brassicae disease development. Black arrows indicate the direction of flow. Grey circles represent transporters involved in the transfer of water and nutrients between different plant tissues. P.brassicae infection may cause local increase of cytokinin in the infected roots cells (pale yellow). This increase could happen as a result of strong xylogenesis repression or produced by the pathogen itself. Xylem transport is reduced (Broken arrows) but nevertheless, root synthesised CKs (blue arrow) are transported together with water and minerals (purple arrow) via remaining xylem vessel to the leaves where they are involved in cytokinin mediated modulation of nutrient partitioning and developmental programs. The leaf synthesised CKs (red arrow) and nutrients (green arrow) are translocated to the infected root through phloem.

Arabidopsis responds to *P.brassicae* infection by redistribution of root CK to the leaves (Figure 45) (Blicharz *et al.*, 2025). In response, *P.brassicae* increases vascular cambium activity, increases phloem formation and represses xylogenesis (Malinowski *et al.*, 2012; Walerowski *et al.*, 2018). The cause for strong repression of xylogenesis in the early stages of infection is not

known but *P.brassicae* could cause xylogenesis repression either by releasing effector proteins in the host which target genes involved in the regulation of xylogenesis or by manipulating the balance of CK species which is important for regulation of xylem specification (Kollmer *et al.*, 2014). Disruption of xylem transport of cytokinin and CK production by *P.brassicae* will increase local CK content in the infection site which might be important to keep infected root cells dividing. Consistent to this, *P.brassicae* has been reported to delay cell cycle exit and maintaining the cells in a proliferative state by stimulating the formation of the E2Fa/RBR1 complex and increase the expression of MYB3R1, MYB3R4, and A- and B-type cyclins which are involved in the G2-M cell cycle check point (Olszak *et al.*, 2019) thus increasing the number of cells susceptible to invasion.

In the early stages of infection, an increase in leaf CK species will occur due to either the redistribution of root CK to the leaves or enhanced CK synthesis within the leaves themselves. Increase in leaf CK could then induce CK-mediated signalling pathways and probably modulate nutrient repartitioning (Figure 45) (Sakakibara *et al.*, 2006; Blicharz *et al.*, 2025). Plants have several sinks during their development which compete for the available carbohydrates. The movement of carbohydrates from the source to sink is dependent on the demand of carbohydrates in the sink tissue but plants tightly regulate the sink-source relationship for their fitness in response to changes in developmental stages and environmental stimuli (Lemoine *et al.*, 2013). As the gall develops, it becomes a strong sink and mass flow through the phloem to the gall removes both sugar and leaf synthesised CK and therefore the above ground growth declines. Phloem transport of IP-type and cZ is also important in the regulation of root development (Bishopp *et al.*, 2011). Taken together, these findings indicate that the CK increase in leaves either by local CK synthesis or its subsequent acropetal movement are crucial for redirecting nutrients to the pathogen, a process that is essential for its growth and development. Additionally, CK synthesized in the leaves and its basipetal movement via the phloem play a key role in reprogramming host development processes. This, in turn, is necessary for increasing the number of infected cells and spore density. Phloem transport has been shown to have an indispensable role in *P.brassicae* growth and in the number and gall size formed on infected plants (Walerowski *et al.*, 2018; Blicharz *et al.*, 2025).

4.4.3. *P.brassicae* tRNA-IPTs may be involved in other cellular metabolic pathways beyond cytokinin biosynthesis

tRNA-IPTs have been found in all living organisms except archaea (Frebort *et al.*, 2011). tRNA-IPTs are responsible for the i6A37 position directly 3' to the anticodon loop of tRNA that decode UNN codons. i6A37 modification is important for translational fidelity and efficiency of protein synthesis. tRNA modification was shown to influence mRNA proof-reading where by bacteria mutants that lacked i6A37 modification were found to have decreased or increased misreading of the first and third position respectively (Golovko *et al.*, 2002). The roles of tRNA modifications in cellular processes are continually being uncovered (Chou *et al.*, 2017) but those involving i6A37 tRNA modifications have been shown to affect expression of amino acid biosynthesis genes, bacterial virulence, nutrient uptake and response to environmental stress (Soderberg & Poulter, 2001). Plant and pathogen tRNA-IPTs are primarily studied for their role in cytokinin biosynthesis, particularly cis-zeatin production, often overlooking their potential involvement in other cellular processes. Mutation of Arabidopsis thaliana tRNA-IPT2 causes reduction in cis-zeatin but did not cause a visible phenotype whereas mutation of the tRNA-IPT9 has reduced cis-zeatin and a chlorotic phenotype and double mutant atipt2,9 lacks completely the cis-zeatin and also have chlorotic phenotype (Miyawaki *et al.*, 2006). It has been shown that the short root and chlorotic phenotypes of atipt2,9 double mutant are not caused by the lack of cis-zeatin (Kollmer *et al.*, 2014; Antoniadi *et al.*, 2022). Exogeneous application of CK and grafting with wild-type scion did not rescue the phenotype but overexpression of ATIPT9 in the double mutant rescued the root phenotype to wild-type leading to a conclusion that IPT9 may have roles that go beyond cytokinin synthesis thereby directly or indirectly impacting auxin homeostasis and plant growth (Antoniadi *et al.*, 2022).

Both ATIPT9 and PbIPT1 are localized to the mitochondria. Given that the atipt2,9 root phenotype is solely attributable to ATIPT9 mutation and can be rescued by reintroducing ATIPT9 (Antoniadi *et al.*, 2022), it was hypothesized that expressing PbIPT1 might also rescue this phenotype. Expression of PbIPT1 rescued the shoot phenotype (Figure 43) of atipt2,9 to wild-type in two independent transgenic lines PbIPT1-3-5-1 and PbIPT1-4-5-1 and significantly increased the primary root length in PbIPT1-3-5-1 (Figure 44) but not in PbIPT1-4-5-1. Variations between the two transgenic lines was expected as the two lines come from two independent transformation events (Figure 43) but the mechanism behind the rescue of shoot phenotype but not root phenotype in PbIPT1-4-5-1 cannot be explained. Partial

complementation of some of the functions of tRNA-IPT have been observed in yeast after re-introduction of a functional tRNA-IPT in tRNA-IPT deleted mutants (Hinsch *et al.*, 2016). Expression of PbIPT2 which, like ATIPT2, is cytoplasmic localised did not rescue both shoot and root phenotypes of atipt2,9 (Figure 43 and 44). Rescuing the root and shoot phenotype of atipt2,9 by PbIPT1 shows that PbIPT1 function is similar to that of ATIPT9. It can be hypothesised that both PbIPTs have other roles independent of CK synthesis and are needed by *P.brassicae* for its own growth and development and stress tolerance. Mutations in tRNA-IPTs causes growth and development abnormalities (Chen *et al.*, 2019), neurodevelopmental diseases (Khalique *et al.*, 2020) and in plant pathogens functional tRNA IPTs are necessary for pathogen virulence and for maintaining cellular function and viability during stress conditions (Chanclud *et al.*, 2016; Hinsch *et al.*, 2016; Eisermann *et al.*, 2020). The subcellular localization of PbIPT2, found in both the nucleus and cytoplasm, suggests a function independent of cytokinin (CK) synthesis, potentially similar to other nuclear and cytoplasmic tRNAs. MOD5 nuclear localised population is involved in tRNA mediated gene silencing (Pratt-Hyatt *et al.*, 2013) whereas cytoplasmic localised population is important in the formation of prion state, a strategy that gives the yeast fitness advantage when exposed to environmental stress (Suzuki *et al.*, 2012). Aggregation of tRNA-IPT in the cytoplasm of C. purpurea when treated with fungicide fluconazole was observed in strain overexpressing tRNA-IPT::mCherry and induced resistance against this fungicide was observed for all the strains lacking cptRNA-IPT, suggesting that cptRNA-IPT forms a prion state under harsh environmental conditions (Hinsch *et al.*, 2016)

P.brassicae IPT genes belong to the tRNA-IPT class, which catalyse the addition of an isopentenyl group to the adenine at position 37 of tRNAs that decode UNN codons. This modification, known as isopentenylation, is commonly referred to as i6A37 (isopentenylated adenine at position 37). The modification is essential for the proper functioning of tRNAs, as it helps maintain the fidelity of translation and stabilizes tRNA structure (Khalique et al 2020). Thus, *P.brassicae* IPT genes likely contribute to protein synthesis, supporting *P.brassicae* growth and development. During this process, the degradation of isopentenylated tRNA produces cZ zeatin (Figure 31), a cytokinin that plays a key role in regulating various plant growth and developmental processes. While total CK content in plants decreases following *P.brassicae* infection, it is reasonable to speculate that the role of *P.brassicae* IPT genes is primarily focused on protein synthesis, with small amounts of CK potentially leaking into the host cytoplasm and triggering CK signalling cascades. Interestingly, Arabidopsis plants

infected with *P.brassicae* displayed axillary bud outgrowth, resulting in a "shooty" phenotype characteristic of a plant's response to increased cytokinins (Cardenas-Aquino *et al.*, 2022), observed at 30dpi. This suggests that *P.brassicae* may release small amounts of CK, which accumulate as the disease progresses, ultimately inducing host CK responses to maintain aboveground plant structures. Furthermore, the observed plant responses, despite the reported decrease in CK levels, could be attributed to factors such as sampling time points, whole organ CK quantification, or the possibility that *P.brassicae* produces other CK species capable of activating host CK signalling. Breakdown of isopentenylated tRNA molecules could result in the accumulation of a novel class of small noncoding RNAs called tRNA-derived RNA fragments (tRFs). tRFs have specific roles in gene regulation, protein synthesis and they act as cross kingdom signalling molecules (Soprano *et al.*, 2018; Alves & Nogueira, 2021). The *P.brassicae* tRFs could therefore be involved in *P.brassicae* gene regulation and protein synthesis or be exported to the host to target host genes involved in regulation of growth and defence responses thus promoting host susceptibility. Concomitant accumulation of tRFs and down-regulation of predicted tRFs target host defence genes have been reported during the black pepper and *Phytophthora capsica* while high levels of tRFs were observed in *Phytophthora infestans* during host interaction suggesting a possible role of tRFs in gene regulation in the pathogen (Soprano *et al.*, 2018).

4.5. Supplementary information

4.5.1. Supplementary figures

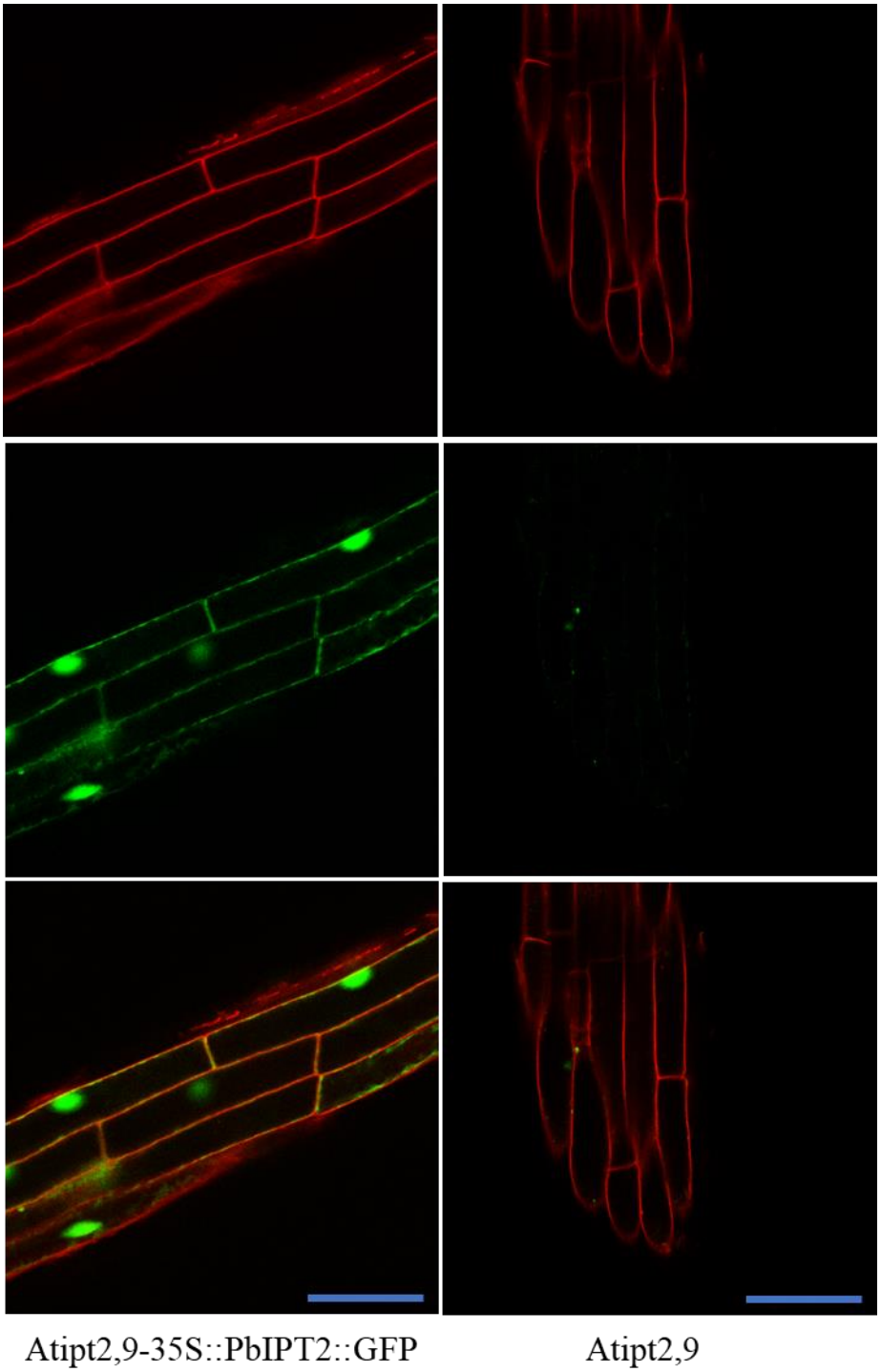


Figure S 4: Subcellular localization of PbIPT2.
Representative confocal pictures of 35S::PbIPT2-GFP and atipt2,9 plants. Top panel is propidium iodide red signal, middle panel is GFP signal, and the bottom panel is GFP merged with propidium iodide. Propidium iodide was used a counter stain for the plasma membrane generating a uniform red stain around the cell. The upper panels shows cell membrane staining,

the middle green show PbIPT2::GFP expression and the bottom panel shows a merged pictured from the two top panels. Consistent to the Mito-tracker staining results, PbIPT2::GFP was confirmed to be localised in the nucleus and cytoplasm.

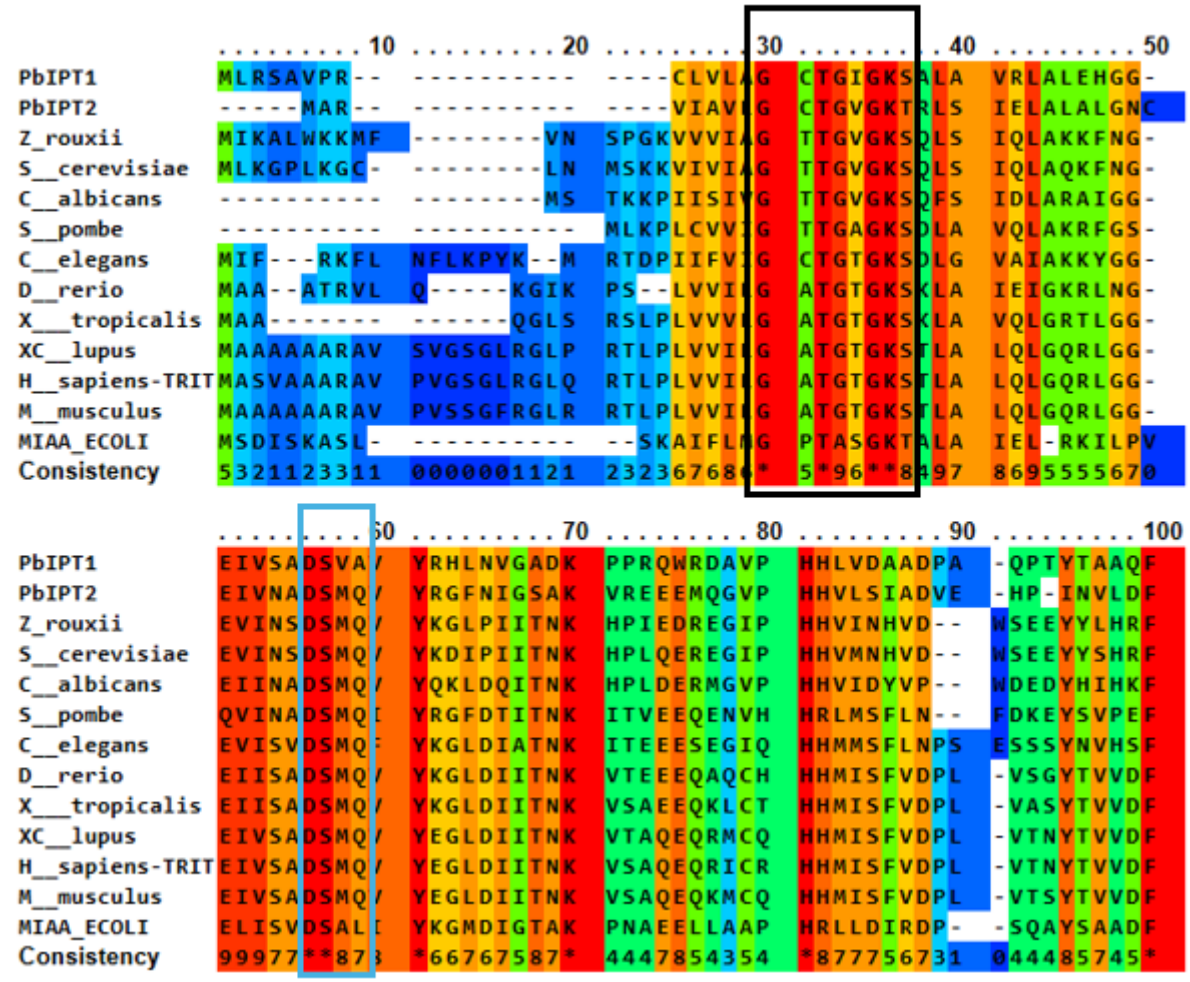


Figure S 5: N-terminal protein sequence alignment of tRNA-IPT homologs from different organisms.
PbIPT1 and PbIPT2 (Plasmodiophora brassicae), Z.rouxii (Zygosaccharomyces rouxii), S.cerevisiae (Saccharomyces cerevisiae), C.albicans (Candida albicans), S.pombe (Schizosaccharomyces pombe), C.elegans (Caenorhabditis elegans), D.rerio (Danio rerio), X.tropicalis (Xenopus tropicalis), XC. lupus (Canis lupus familiaris), H_sapiens-TRIT1 (Homo sapiens), M_musculus (Mus musculus) and MIAA_Ecoli (Escherichia coli). The alignment is color-coded for amino acid conservation (scoring by PRALINE) scoring from 0 (least conserved; blue) to 10 (most conserved; red). The P-loop and the DSMQ motif are indicated by black and blue lines respectively.

Unconserved 0 1 2 3 4 5 6 7 8 9 10 Conserved

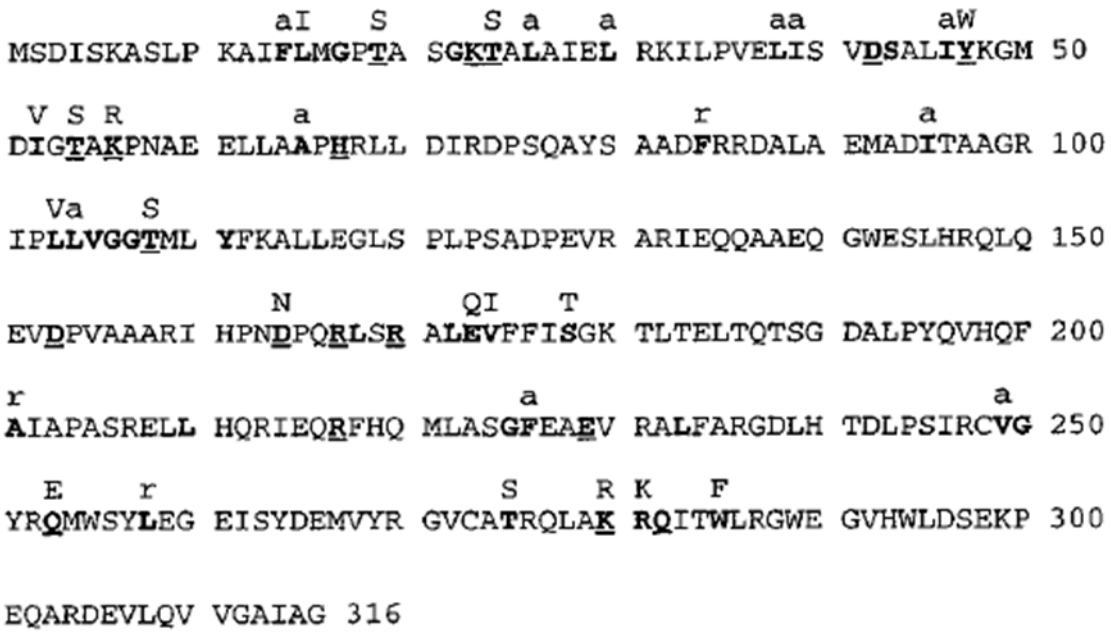


Figure S 6: Amino acid sequence of DMAPP-tRNA transferase from E. coli.
Bold residues represent those that are conserved or conservatively substituted. Underlined residues are those for which mutations produced large changes in enzyme kinetic parameters. Letters above residues specify conservative substitutions that are present in other DMAPP-tRNA transferases. The letter "a" is used for aliphatic residues (alanine, valine, leucine, and isoleucine), "r" for aromatic residues (phenylalanine, tyrosine, and tryptophan), and "b" when aliphatic and aromatic substitutions are seen (Soderberg & Poulter, 2001).

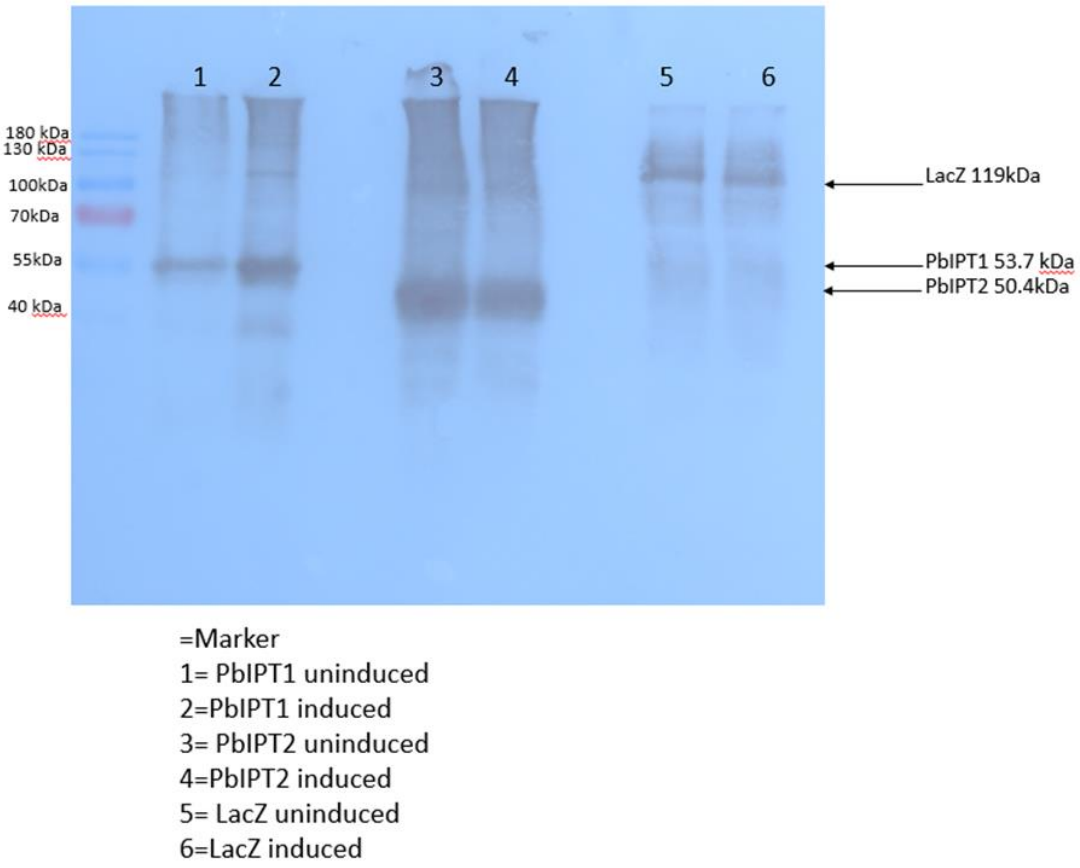


Figure S 7: Western blot analysis showing the expression of PbIPT1, PbIPT2, and LacZ proteins in E. coli.
Recombinant protein expression was induced by IPTG addition to the bacterial culture. Cells containing the induced proteins were harvested and lysed by sonication. The resulting cell lysate was separated by SDS-PAGE, and proteins were subsequently transferred to a membrane. Detection of PbIPT1, PbIPT2, and LacZ was then performed using anti-His antibodies.

4.5.2. Appendices

Appendix 2

>CEO95848.1 PBRA_004561(PLBR_LOCUS7342, PbIPT1)

MLRSAVPRCLVLAGCTGIGKSALAVRLALEHGGEIVSADSVAVYRHLNVGADKPPR
QWRDAVPHHLVDAADPAQPTYTAAQFADDATGAIHAIHSRGNLPIVVGGSGMYLSF
LIHGFDRSQSSSARCDTVSIGMVEAAVSGSQSWDDAISVLGGALSPGLASTARQVAD
YLKHIQPGNVRRLESAVHRLRCLAPGSAVPEYNGRQVYDGTPKGNGDQIDFRCVFLQ
CPRTTMYRAIDLRCERMLQNGLLDEVSYLLERGWLVEGTPPANSIGYRQAIEFLREPK
FDLAAFRLFVDEFQAASRQFARKQVSWFRRDFEYRFIDVDRKADPEDDRLSAMLMD
WYRLDRDQLDAVRASSEYQLAQKALRQTLTKDARDEQRKYLPERVIFKHGLSLDAI
MRLMAQSQPSPDFAPPKFVPPDNVDDSAQSEAAASNLYCSRT

>SPQ98244.1 (PLBR_LOCUS5459, PbIPT2)

MARVIAVLGCTGVGKTRLSIELALALGNCEIVNADSMQVYRGFNIGSAKVREEEMQG
VPHHVLSIADVEHPINVLDFVRAARQAIDEVHARGNTPILVGGTVYYVAALLFNTFLA
SNADGDHDGASSTPCQTGIAGDSYATLQALDPDMAERLHPNDIRRIRRSLEVMAQKQ
QKHSELLAQSGGRHGVGPPRYRTLALWLDAAPDVIEGRIQKRTESMFKDGIFDENVQ
LYEQLKLAYGNDLDKIYGAGIGQAIGFREFRHCFEASGEVLAGALRQSVVDLNRNTF
RFAKRQQRNIAGWFSGRDDLPVFRIDTSDAASWSNAVLQPAVDICKSFVDPGVPDPN
PDAYLHLHEVLAGPKAPERRSDQGWRKFACEDCGGRILNGALEWRVHLQSRSHRRH
AKRKHVRQCT

CHAPTER 5: GENERAL CONCLUSION

P.brassicae poses a significant economic threat to plants in the Brassicaceae family, a group that encompasses high-value vegetables and agricultural crops such as brussels sprouts and oilseed rape. Notably, the oilseed rape industry generated £877 million for the UK economy in 2022 (https://www.gov.uk/government/statistics/agriculture-in-the-united-kingdom-2022 accessed on 03/03/2025) and $14.5 billion CAD in exports for the Canadian economy in 2024 (https://www.canolacouncil.org/markets-stats/ accessed on 03/03/2025). *P.brassicae* has been reported to cause up to 50% yield loss and sometimes total crop failure in heavily infested fields (Strelkov & Hwang, 2014). Clubroot disease management is challenging due to the large quantities and long-lasting spores it produces, extensive host range and lack of effective chemical and biological control agents. This is further hampered by limited knowledge on the biology of *P.brassicae* due to its strict intracellular lifestyle. Genome information, however, can be leveraged to study the biology of these experimentally challenging important pathogens (Bryant *et al.*, 2023).

Decreasing sequencing costs coupled with advances in sequencing technologies are enabling researchers to sequence genomes that they are interested in. This is true for the clubroot research community where 51 *P.brassicae* genomes have been published since the first *P.brassicae* genome was published in 2015 (Schwelm et al., 2015). Due to the use of different gene prediction methods and supporting evidence, the current *P.brassicae* gene models in different annotations have disagreements. This study compared PbPT3 and Pb314v2 *P.brassicae* genome annotations and found that 45% of gene models in the two annotations matched at locus level and the number of correct gene models. Manual inspection of gene models using short and long reads further identified errors that could significantly reduce the number of correct gene models in these annotations. There were 13 different major problems in the gene models which are probably common in different annotations. These can be summarised as (i) inappropriately merged or split genes, (ii) genes with intron–exon number and boundary errors, (iii) genes with long/short 5' and 3' regions, (iv) missing gene models, (v) gene models that lack supporting evidence and (vi) genomic regions with expression evidence but lack predicted gene models. The gene model errors identified in this study highlight the need for manual curation of *P.brassicae* gene models. This process will not only improve the accuracy of existing models but also likely reveal additional genes within the *P.brassicae* genome. Manual curation however is labour and resource intensive and is mainly done by model organism databases.

Manual curation of *P.brassicae* gene models is further hampered by insufficient transcriptome data. The number of RNAseq studies on *P.brassicae* or *P.brassicae* infected plants are few and the number of *P.brassicae* specific sequences in most of these datasets is small which make it difficult to curate gene models with few or no reads. In addition, Illumina RNAseq reads are short and are not able to provide unambiguous resolution for gene models in regions with uniformly distributed reads which as a result makes *P.brassicae* genomemanual curation more challenging. However, it is hypothesised that the number of genes expressed increase when datasets from different treatments, tissues and genotypes are used in the analysis (Cheng et al., 2017). Thus, *P.brassicae* RNAseq studies are still valuable to improve the gene models. The use of RNAseq datasets in ongoing annotation improvements and regular updates has provided an unparalleled level of detail even in extensively annotated genomes like those of Arabidopsis, fruit flies, humans, mice, rice, and yeast (Cheng et al., 2017). Furthermore, long read sequencing technology has a great potential to improve *P.brassicae* genome annotations as shown in this study and as well as in genome annotations for sorghum, maize, mouse, and human (Cheng et al., 2017). Accurate interpretation of genomic function relies on ongoing annotation improvements and regular updates.

For a long period of time, studies about *P.brassicae* nutrition have focused on how the pathogen manipulates host metabolism in its favour and thus little is known about how *P.brassicae* transports these nutrients across its plasma membrane and the form in which they are transported. This is partly due to its strict obligate biotrophic lifestyle meaning it cannot be grown in culture, the lack of transformation systems and its unusual intracellular lifestyle compared with most (hemi-)biotrophs which are extracellular. Nevertheless, the availability of genome sequence of *P.brassicae* offers the opportunity to investigate how the pathogen manipulates host metabolism and development in its favour. Identification of membrane transporters is crucial as they shape the unique lifestyle of this economic pathogen of cruciferous plants. In this study, the Pb314v2 genome sequence of *P.brassicae* was searched for homologues in the TCDB to understand the transport capabilities of*P.brassicae* . The current data give a comprehensive view of *P.brassicae* transporters which are expressed in resting spores, germinating spores and cortical infection stages. The data showed that *P.brassicae* devote 18% of its total genes to transport. 74% of the *P.brassicae* transporters fall in classes 1-5 of the TCDB which contains transporters whose function as transporters is known and 26% fall in the classes 8 and 9 that contains accessory proteins involved in transport and those with incompletely characterised transport system. 30.62% of the total *P.brassicae*

transporters are primary active transporters indicating that *P.brassicae* depends heavily on active transport to transport solutes. Genomes exhibit a broad range of diversity both in number and type of transport systems (Elbourne, LDH *et al.*, 2017). Comparison of transport systems of *P.brassicae* with other obligate biotrophic pathogens indicate that *P.brassicae* have more channel/pore transport systems, more transporter families and transporters. Unlike other obligate biotrophic pathogens which colonises the intercellular space and obtain the nutrient from the host via haustoria (Garnica *et al.*, 2014), *P.brassicae* lives inside the host cell thus the presence of more transport systems, families and transporters might help the *P.brassicae* to adapt to intracellular host environment. The *P.brassicae* transporter repertoire gives the pathogen the ability to transport all nutrients that it needs from the host. Transporters for vitamin B and some of the amino acids whose biosynthetic pathways are incomplete or completely missing in *P.brassicae* genome (Schwelm *et al.*, 2015; Rolfe *et al.*, 2016) were identified. However, the range of putative substrates assigned to each transporter in this study is not definitive. Thus, to experimentally verify transport activity, all substrates associated with a transporter based on TCDB analysis should be tested. It is important to note that the *P.brassicae* genome annotation is still under development. Therefore, the high number of transporters reported in this study may change as more accurate and complete gene models become available. Nonetheless, the accuracy and the reliability of the findings of this study are supported by expression data from 14 publicly available transcriptome studies and heterologous expression studies (Kong *et al.*, 2022).

P.brassicae disease symptom development is largely due to uncontrolled cell division and cell enlargement in the infected tissues. This characteristic growth indicates the involvement of cytokinin and auxin, the key plant hormones that regulate many plant growth developmental processes (Ludwig-Müller *et al.*, 2009). *P.brassicae* has been proposed to synthesise and release CK into the host during clubroot disease development. *P.brassicae* has been shown to have the ability to synthesise CK by incorporating radioactive labelled Adenine in a compound identified as trans-zeatin (Muller & Hilgenberg, 1986). Identification of two isopentenyl transferase genes (IPT) which are key enzymes in the biosynthesis of CK but may also have other cellular functions beyond CK synthesis further supports the ability of *P.brassicae* to synthesise CK (Rolfe *et al.*, 2016). However, there are contradicting reports concerning the concentration and composition of CK species in *P.brassicae* infected plants. Concentration of CK species vary between plants species and cytokinin signalling is sensitive to variations in CK composition and concentrations of CK species , and even minor imbalances can

significantly alter the CK landscape (Gajdosová *et al.*, 2011; Kollmer *et al.*, 2014). Since IPT genes, which fall into two classes influence CK composition, this study focused on the experimental characterisation of *P.brassicae* IPTs. Genetic studies in yeast and *Arabidopsis thaliana* showed that *P.brassicae* IPT genes belong to the tRNA-IPT class. Restoration of root and shoot phenotypes of atipt2,9 double mutant, a phenotype that is not restored by exogenous application of cis-zeatin but only by the reintroduction of IPT9 (Antoniadi *et al.*, 2022) indicate that PbIPT1 plays a role in tRNA modifications that maybe required in different cellular processes. The subcellular localisation of PbIPT2 may play a major role for *P.brassicae* stress adaptations (Hinsch *et al.*, 2016). Because *P.brassicae* IPTs belong to tRNA-IPT class which synthesise cZ, it should be expected that the concentration of cZ in *P.brassicae* infected plants should increase. However, it has been reported that *P.brassicae* infection causes no changes or decreases cZ contents of infected plants (Malinowski *et al.*, 2016; Blicharz *et al.*, 2025) whereas some isopentyl-type have reported to increase in the early stages of infection (Devos *et al.*, 2006; Blicharz *et al.*, 2025). Decrease in cZ can be attributed to an increase in selective degradation of cZ (Kollmer *et al.*, 2014). Additionally, *P.brassicae* maybe releasing small amounts of cZ whose contribution to the total CK content of the host is insignificant. Nevertheless, the phenotype of *P.brassicae* infected Arabidopsis is typical for the shoot phenotypes in response to increased CK. The observation that *P.brassicae* infection of Arabidopsis ipt 1;3;5;7 mutant did activate CK responsive genes but did not rescue the mutant phenotype (Malinowski *et al.*, 2016) can be interpreted as *P.brassicae* produced CK species in this mutant may not be important in Arabidopsis but may be important in other Brassica plants. Based on the findings of this study a model explaining how *P.brassicae* highjack host CK signalling to reprogramme host metabolism and development in favour for its own growth and development.

The findings of this study offer a significant contribution to the body of knowledge as far as clubroot research is concerned. The gene model errors identified underscores the importance of comparing *P.brassicae* gene models from different annotations before any functional studies of *P.brassicae* genes. In addition, a web-based community based *P.brassicae* genome annotation platform was created. The resource is hosted at the university of Sheffield website. Once accessible to all, this resource would provide the clubroot research community with a standard reference genome with high quality and well curated gene models and allow centralised and continuous gene annotation updates as more data from RNAseq and long read/ full length mRNA sequencing projects become available. Transporters plays a major role in drug

resistance and are targets for drugs themselves (Martin, 2020). Transporters identified in this study especially the transporter families which are not shared between the host plants and *P.brassicae* could be targeted for clubroot disease control. Engineering crop plants (Gewin, 2003) or the *P.brassicae* biocontrol agents to produce antimicrobial compounds that interfere with the transport activity of unique *P.brassicae* transporters could reduce the economic loss from clubroot disease. The use of engineered microbes as biofertilizers and biopesticides in sustainable agriculture is gaining traction (Sudheer *et al.*, 2020). Microbial strains can be engineered to augment the production of biopesticides or metabolic pathways can be transferred to a strain to produce the biopesticide. For example, *Escherichia coli* was engineered to produce cinnamaldehyde a nematicide which produced comparable effectiveness as commercial cinnamaldehyde (Panda & Zhou, 2023). Bacteria strains *Bacillus velezensis* and *Bacillus amyloliquefaciens* have been isolated from the rhizosphere of *P.brassicae* infected *B.napus.* These strains were reported to produce antimicrobial compounds that inhibit root hair infection and clubroot disease development (Zhu *et al.*, 2019). Since these strains are already present in rhizosphere of *B. napus*, they can be engineered to produce antimicrobials that target the unique *P.brassicae* transporters. Development of drugs that specifically inhibit transport functions of *P.brassicae* transporters expressed in germinating spores could also be used to control clubroot disease. The characterisation of *P.brassicae* IPT genes has paved the way for future studies to investigate if *P.brassicae* produces other CK compounds that may induce host CK signalling cascades like *R.fascians* (Pertry *et al.*, 2009). Further research in I^6A_{37} tRNA modification landscape in both host plants and *P.brassicae* is needed to understand how I^6A_{37} tRNA modification may impact disease development independent of CK biosynthesis.

References

Akhtar SS, Mekureyaw MF, Pandey C, Roitsch T. 2019. Role of Cytokinins for Interactions of Plants With Microbial Pathogens and Pest Insects. *Front Plant Sci* **10**: 1777.

Al-Khami AA, Rodriguez PC, Ochoa AC. 2017. Energy metabolic pathways control the fate and function of myeloid immune cells. *Journal of Leukocyte Biology* **102**(2): 369-380.

Alves CS, Nogueira FTS. 2021. Plant Small RNA World Growing Bigger: tRNA-Derived Fragments, Longstanding Players in Regulatory Processes. *Frontiers in Molecular Biosciences* **8**.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* **11**(10): R106.

Antoniadi I, Mateo-Bonmati E, Pernisova M, Brunoni F, Antoniadi M, Villalonga MG, Ament A, Karady M, Turnbull C, Dolezal K, et al. 2022. IPT9, a cis-zeatin cytokinin biosynthesis gene, promotes root growth. *Front Plant Sci* **13**: 932008.

Antoniadi I, Plackova L, Simonovik B, Dolezal K, Turnbull C, Ljung K, Novak O. 2015. Cell-Type-Specific Cytokinin Distribution within the Arabidopsis Primary Root Apex. *Plant Cell* **27**(7): 1955-1967.

Asano T, Kageyama K. 2006. Growth and movement of secondary plasmodia of in turnip suspension-culture cells. *Plant Pathology* **55**(1): 145-151.

Auer S, Zamani-Noor N, Mahfoud Y, Ludwig-Müller J. 2024. Exploring the influence of rapeseed cultivar and pathogen isolate on Acremonium alternatum's efficacy in clubroot disease control. *European Journal of Plant Pathology*.

Balibrea-Lara ME, Gonzalez Garcia MC, Fatima T, Ehness R, Lee TK, Proels R, Tanner W, Roitsch T. 2004. Extracellular invertase is an essential component of cytokinin-mediated delay of senescence. *Plant Cell* **16**(5): 1276-1287.

Bass D, Chao EEY, Nikolaev S, Yabuki A, Ishida KI, Berney C, Pakzad U, Wylezich C, Cavalier-Smith T. 2009. Phylogeny of Novel Naked Filose and Reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea Revised. *Protist* **160**(1): 75-109.

Baxter L, Tripathy S, Ishaque N, Boot N, Cabral A, Kemen E, Thines M, Ah-Fong A, Anderson R, Badejoko W, et al. 2010. Signatures of Adaptation to Obligate Biotrophy in the Genome. *Science* **330**(6010): 1549-1551.

Bednarova A, Hanna M, Durham I, VanCleave T, England A, Chaudhuri A, Krishnan N. 2017. Lost in Translation: Defects in Transfer RNA Modifications and Neurological Disorders. *Front Mol Neurosci* **10**: 135.

Bharathi V, Girdhar A, Patel BK. 2017. A Protocol of Using White/Red Color Assay to Measure Amyloid-induced Oxidative Stress in. *Bio-Protocol* **7**(15).

Bishopp A, Lehesranta S, Vaten A, Help H, El-Showk S, Scheres B, Helariutta K, Mahonen AP, Sakakibara H, Helariutta Y. 2011. Phloem-transported cytokinin regulates polar auxin transport and maintains vascular pattern in the root meristem. *Curr Biol* **21**(11): 927-932.

Blicharz S, Stefanowicz K, Truman W, Basinska-Barczak A, Singh D, Kasprzewska A, de Diego N, Vrobel O, Zeljkovic SC, Tarkowski P, et al. 2025. Laser dissection-assisted phloem transcriptomics highlights the metabolic and physiological changes accompanying clubroot disease progression in oilseed rape. *Plant Journal* **121**(1).

Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, et al. 2021. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**(D1): D344-D354.

Boguta M, Hunter LA, Shen WC, Gillman EC, Martin NC, Hopper AK. 1994. Subcellular Locations of Mod5 Proteins - Mapping of Sequences Sufficient for Targeting to

Mitochondria and Demonstration That Mitochondrial and Nuclear Isoforms Commingle in the Cytosol. *Molecular and Cellular Biology* **14**(4): 2298-2306.

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

Borg J, Kiær LP, Lecarpentier C, Goldringer I, Gauffreteau A, Saint-Jean S, Barot S, Enjalbert J. 2018. Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps. *Field Crops Research* **221**: 298-313.

Bosch M, Franklin-Tong V. 2024. Regulating programmed cell death in plant cells: Intracellular acidification plays a pivotal role together with calcium signaling. *Plant Cell* **36**(11): 4692-4702.

Botero-Ramirez A, Kirk B, Strelkov SE. 2024. Optimizing Clubroot Management and the Role of Canola Cultivar Mixtures. *Pathogens* **13**(8).

Bouchard R, Bailly A, Blakeslee JJ, Oehring SC, Vincenzetti V, Lee OR, Paponov I, Palme K, Mancuso S, Murphy AS, et al. 2006. Immunophilin-like TWISTED DWARF1 modulates auxin efflux activities of P-glycoproteins. *Journal of Biological Chemistry* **281**(41): 30603-30612.

Brodmann D, Schuller A, Ludwig-Müller J, Aeschbacher AR, Wiemken A, Boller T, Wingler A. 2002. Induction of Trehalase in Arabidopsis Plants Infected With the Trehalose-Producing Pathogen Plasmodiophora brassicae. *MPMI* **15**(7): 8.

Bryant AS, Akimori D, Stoltzfus JDC, Hallem EA. 2023. A standard workflow for community-driven manual curation of genome annotations. *Philosophical Transactions of the Royal Society B-Biological Sciences* **379**(1894).

Burki F, Kudryavtsev A, Matz MV, Aglyamova GV, Bulman S, Fiers M, Keeling PJ, Pawlowski J. 2010. Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. *Bmc Evolutionary Biology* **10**.

Cao T, Tewari J, Strelkov SE. 2007. Molecular Detection of Plasmodiophora brassicae, Causal Agent of Clubroot of Crucifers, in Plant and Soil. *Plant Disease* **91**(1): 80-87.

Cardenas-Aquino MD, Sarria-Guzman Y, Martinez-Antonio A. 2022. Review: Isoprenoid and aromatic cytokinins in shoot branching. *Plant Sci* **319**: 111240.

Cavalier-Smith T, Chao EE, Lewis R. 2018. Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. *Protoplasma* **255**(5): 1517-1574.

Cerna-Vargas JP, Sanchez-Romera B, Matilla MA, Ortega A, Krell T. 2023. Sensing preferences for prokaryotic solute binding protein families. *Microb Biotechnol* **16**(9): 1823-1833.

Chanclud E, Kisiala A, Emery NR, Chalvon V, Ducasse A, Romiti-Michel C, Gravot A, Kroj T, Morel JB. 2016. Cytokinin Production by the Rice Blast Fungus Is a Pivotal Requirement for Full Virulence. *PLoS Pathog* **12**(2): e1005457.

Chang Q, Lin X, Yao M, Liu P, Guo J, Huang L, Voegele RT, Kang Z, Liu J. 2020. Hexose transporter PsHXT1-mediated sugar uptake is required for pathogenicity of wheat stripe rust. *Plant Biotechnol J* **18**(12): 2367-2369.

Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ, Guo WJ, Kim JG, Underwood W, Chaudhuri B, et al. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **468**(7323): 527-532.

Chen LQ, Qu XQ, Hou BH, Sosso D, Osorio S, Fernie AR, Frommer WB. 2012. Sucrose efflux mediated by SWEET proteins as a key step for phloem transport. *Science* **335**(6065): 207-211.

Chen Y, Bai B, Yan H, Wen F, Qin D, Jander G, Xia Q, Wang G. 2019. Systemic disruption of the homeostasis of transfer RNA isopentenyltransferase causes growth and development abnormalities in. *Insect Molecular Biology* **28**(3): 380-391.

Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. 2017. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J* **89**(4): 789-804.

Cheng JY, Luo T, Wu MD, Zhang J, Yang L, Chen WD, Li GQ. 2024. Biological Control of Rapeseed Clubroot (Plasmodiophora brassicae) Using the Endophytic Fungus Didymella macrostoma P2. *Plant Disease* **108**(8): 2399-2409.

Chou HJ, Donnard E, Gustafsson HT, Garber M, Rando OJ. 2017. Transcriptome-wide Analysis of Roles for tRNA Modifications in Translational Regulation. *Molecular Cell* **68**(5): 978-+.

Chu MG, Song T, Falk KC, Zhang XG, Liu XJ, Chang A, Lahlali R, McGregor L, Gossen BD, Peng G, et al. 2014. Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by Plasmodiophora brassicae. *BMC Genomics* **15**.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal* **16**(6): 735-743.

David R, Byrt CS, Tyerman SD, Gilliham M, Wege S. 2019. Roles of membrane transporters: connecting the dots from sequence to phenotype. *Annals of Botany* **124**(2): 201-207.

de Souza W. 2024. Contribution of microscopy to a better understanding of the anatomy of pathogenic protists. *Proceedings of the National Academy of Sciences of the United States of America* **121**(17).

Dekhuijzen HM. 1981. The Occurrence of Free and Bound Cytokinins in Plasmodia of Plasmodiophora brassicae Isolated from Tissue Cultures of Clubroots. *Plant Cell Reports* **1**: 3.

Dekhuijzen HM, Overeem JC. 1971. The Role of Cytokinins in Clubroot Formation. *Physiol.Pl. Path.* **1**: 10.

Del Bianco M, Giustini L, Sabatini S. 2013. Spatiotemporal changes in the role of cytokinin during root development. *New Phytol* **199**(2): 324-338.

Depuydt S, Dolezal K, Van Lijsebettens M, Moritz T, Holsters M, Vereecke D. 2008. Modulation of the Hormone Setting by Rhodococcus fascians Results in Ectopic KNOX Activation in Arabidopsis. *Plant Physiology* **146**(3): 1267-1281.

Desai SA, Bezrukov SM, Zimmerberg J. 2000. A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite. *Nature* **406**(6799): 1001-1005.

Devos S, Laukens K, Deckers P, Van Der Straeten D, Beeckman T, Inzé D, Van Onckelen H, Witters E, Prinsen E. 2006. A Hormone and Proteome Approach to Picturing the Initial Metabolic Events During Plasmodiophora brassicae Infection on Arabidopsis. *MPMI* **19**(12): 12.

Devos S, Vissenberg K, Verbelen JP, Prinsen E. 2005. Infection of Chinese cabbage by Plasmodiophora brassicae leads to a stimulation of plant growth: impacts on cell wall metabolism and hormone balance. *New Phytol* **166**(1): 241-250.

Dixon G, R. 2009. The Occurrence and Economic Impact of Plasmodiophora brassicae and Clubroot Disease. *J Plant Growth Regul* **28**(3): 9.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**(1): 15-21.

Doidy J, Grace E, Kuhn C, Simon-Plas F, Casieri L, Wipf D. 2012. Sugar transporters in plants and in their interactions with fungi. *Trends Plant Sci* **17**(7): 413-422.

Donald EC, Lawrence JM, Porter IJ. 2002. Evaluation of a fluorescent staining technique as an indicator of pathogenicity of resting spores of. *Australasian Plant Pathology* **31**(4): 373-379.

Donizelli M, Djite MA, Le Novère N. 2006. LGICdb:: a manually curated sequence database after the genomes. *Nucleic Acids Research* **34**: D267-D269.

Duan H, Jones AW, Hewitt T, Mackenzie A, Hu Y, Sharp A, Lewis D, Mago R, Upadhyaya NM, Rathjen JP, et al. 2022. Physical separation of haplotypes in dikaryons allows benchmarking of phasing accuracy in Nanopore and HiFi assemblies with Hi-C data. *Genome Biol* **23**(1): 84.

Duan XY, Pan SY, Fan MY, Chu BY, Ma ZH, Gao F, Zhao ZH. 2022. Cultivar Mixture Enhances Crop Yield by Decreasing Aphids. *Agronomy-Basel* **12**(2).

Dumoux M, Hayward RD. 2016. Membrane contact sites between pathogen-containing compartments and host organelles. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1861**(8): 895-899.

Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, et al. 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci U S A* **108**(22): 9166-9171.

Egan DF, Shackelford DB, Mihaylova MM, Gelino S, Kohnz RA, Mair W, Vasquez DS, Joshi A, Gwinn DM, Taylor R, et al. 2011. Phosphorylation of ULK1 (hATG1) by AMP-Activated Protein Kinase Connects Energy Sensing to Mitophagy. *Science* **331**(6016): 456-461.

Eisermann I, Motyka V, Kümmel S, Dobrev PI, Hübner K, Deising HB, Wirsel SGR. 2020. CgIPT1 is required for synthesis of cis-zeatin cytokinins and contributes to stress tolerance and virulence in Colletotrichum graminicola. *Fungal Genetics and Biology* **143**.

Elbourne LD, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res* **45**(D1): D320-D324.

Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Research* **45**(D1): D320-D324.

Ernst TW, Kher S, Stanton D, Rennie DC, Hwang SF, Strelkov SE. 2019. Plasmodiophora brassicae resting spore dynamics in clubroot resistant canola (Brassica napus) cropping systems. *Plant Pathology* **68**(2): 399-408.

Fatima U, Senthil-Kumar M. 2015. Plant and pathogen nutrient acquisition strategies. *Frontiers in Plant Science* **6**.

Feng J, Hwang SF, Strelkov SE. 2013a. Genetic transformation of the obligate parasite Plasmodiophora brassicae. *Phytopathology* **103**(10): 1052-1057.

Feng J, Hwang SF, Strelkov SE. 2013b. Studies into primary and secondary infection processes by Plasmodiophora brassicae on canola. *Plant Pathology* **62**(1): 177-183.

Feng Y, Yang X, Cai G, Wang S, Liu P, Li Y, Chen W, Li W. 2024. Identification and Characterization of High-Molecular-Weight Proteins Secreted by Plasmodiophora brassicae That Suppress Plant Immunity. *J Fungi (Basel)* **10**(7).

Fox NM, Hwang SF, Manolii VP, Turnbull G, Strelkov SE. 2022. Evaluation of lime products for clubroot (Plasmodiophora brassicae) management in canola (Brassica napus) cropping systems. *Canadian Journal of Plant Pathology* **44**(1): 21-38.

Frebort I, Kowalska M, Hluska T, Frebortova J, Galuszka P. 2011. Evolution of cytokinin biosynthesis and degradation. *J Exp Bot* **62**(8): 2431-2452.

Gajdosová S, Spíchal L, Kamínek M, Hoyerová K, Novák O, Dobrev PI, Galuszka P, Klíma P, Gaudinova A, Zizková E, et al. 2011. Distribution, biological activities, metabolism, and the conceivable function of cis-zeatin-type cytokinins in plants. *Journal of Experimental Botany* **62**(8): 2827-2840.

Galis I, Bilyeu K, Wood G, Jameson PE. 2005. Rhodococcus fascians: Shoot Proliferation without Elevated Cytokinins? *Plant Growth Regulation* **46**(2): 109-115.

Gancedo C, Flores CL. 2004. The importance of a functional trehalose biosynthetic pathway for the life of yeasts and fungi. *FEMS Yeast Res* **4**(4-5): 351-359.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, et al. 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**: 14.

Garnica DP, Nemri A, Upadhyaya NM, Rathjen JP, Dodds PN. 2014. The ins and outs of rust haustoria. *PLoS Pathog* **10**(9): e1004329.

Garnica DP, Upadhyaya NM, Dodds PN, Rathjen JP. 2013. Strategies for Wheat Stripe Rust Pathogenicity Identified by Transcriptome Sequencing. *PLoS One* **8**(6): e67150.

Garvetto A, Murua P, Kirchmair M, Salvenmoser W, Hittorf M, Ciaghi S, Harikrishnan SL, Gachon CMM, Burns JA, Neuhauser S. 2023. Phagocytosis underpins the biotrophic lifestyle of intracellular parasites in the class Phytomyxea (Rhizaria). *New Phytol* **238**(5): 2130-2143.

Gerstein MB, Rozowsky J, Yan KK, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Li JJ, et al. 2014. Comparative analysis of the transcriptome across distant species. *Nature* **512**(7515): 445-448.

Gewin V. 2003. Genetically modified corn--environmental benefits and risks. *PLoS Biol* **1**(1): E8.

Gietz RD, Woods RA. 2006. Yeast transformation by the LiAc/SS Carrier DNA/PEG method. *Methods Mol Biol* **313**: 107-120.

Gillman EC, Slusher LB, Martin NC, Hopper AK. 1991. Mod5 Translation Initiation Sites Determine N6- Isopentenyladenosine Modification of Mitochondrial and Cytoplasmic Transfer-Rna. *Molecular and Cellular Biology* **11**(5): 2382-2390.

Goethals K, Vereecke D, Jaziri M, Van Montagu M, Holsters M. 2001. Leafy gall formation by Rhodococcus fascians. *Annu Rev Phytopathol* **39**: 27-52.

Golovko A, Sitbon F, Tillberg E, Nicander B. 2002. Identification of a tRNA isopentenyltransferase gene from Arabidopsis thaliana. *Plant Molecular Biology* **49**(2): 161-169.

Gossen BD, Adhikari KKC, Mcdonald MR. 2012. Effect of seeding date on development of clubroot in short-season Brassica crops. *Canadian Journal of Plant Pathology* **34**(4): 516-523.

Gossen BD, Deora A, Peng G, Hwang SF, McDonald MR. 2014. Special Issue: Effect of environmental parameters on clubroot development and the risk of pathogen spread. *Canadian Journal of Plant Pathology* **36**: 37-48.

Goulet KM, Saville BJ. 2017. Carbon acquisition and metabolism changes during fungal biotrophic plant pathogenesis: insights fromUstilago maydis. *Canadian Journal of Plant Pathology* **39**(3): 247-266.

Gravot A, Richard G, Lime T, Lemarie S, Jubault M, Lariagon C, Lemoine J, Vicente J, Robert-Seilaniantz A, Holdsworth MJ, et al. 2016. Hypoxia response in Arabidopsis roots infected by Plasmodiophora brassicae supports the development of clubroot. *BMC Plant Biol* **16**(1): 251.

Gu S, Roderick HL, Camacho P, Jiang JX. 2000. Identification and characterization of an amino acid transporter expressed differentially in liver. *Proc Natl Acad Sci U S A* **97**(7): 3230-3235.

Guerillot P, Salamov A, Louet C, Morin E, Frey P, Grigoriev IV, Duplessis S. 2023. A Remarkable Expansion of Oligopeptide Transporter Genes in Rust Fungi (Pucciniales) Suggests a Specialization in Nutrient Acquisition for Obligate Biotrophy. *Phytopathology* **113**(2): 252-264.

Guivarc'h A, Rembur J, Goetz M, Roitsch T, Noin M, Schmülling T, Chriqui D. 2002. Local expression of the gene in transgenic tobacco (L. cv. SR1) axillary buds establishes a role for cytokinins in tuberization and sink formation. *Journal of Experimental Botany* **53**(369): 621-629.

Gulzar S, Liaquat F, Hameed MK, Rehman A, Zahid MS, Manzoor MA, Shah IH, Zhang YD. 2024. Clubroot (Plasmodiophora brassicae) Suppression Under Biocontrol Agents in Pak choi with Variations in Physiological, Biochemical, and Bacterial Diversity. *Journal of Plant Growth Regulation* **43**(5): 1331-1349.

Hahn M, Neef U, Struck C, Göttfert M, Mendgen K. 1997. A Putative Amino Acid Transporter Is Specifically Expressed in Haustoria of the Rust Fungus Uromyces fabae. *MPMI* **10**(4): 8.

Harrison SJ, Mott EK, Parsley K, Aspinall S, Gray JC, Cottage A. 2006. A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods* **2**: 19.

Hasan M, Hossain M, Jiang DH. 2023. New endophytic strains of Trichoderma promote growth and reduce clubroot severity of rapeseed (Brassica napus). *PLoS One* **18**(10).

Hatakeyama K, Suwabe K, Tomita RN, Kato T, Nunome T, Fukuoka H, Matsumoto S. 2013. Identification and Characterization of Crr1a, a Gene for Resistance to Clubroot Disease (Plasmodiophora brassicae Woronin) in Brassica rapa L. *PLoS One* **8**(1).

Hediger MA, Clémençon B, Burrier RE, Bruford EA. 2013. The ABCs of membrane transporters in health and disease (SLC series): Introduction. *Molecular Aspects of Medicine* **34**(2-3): 95-107.

Hennig BC, Hwang SF, Manolii VP, Turnbull G, Robinson SVJ, Strelkov SE. 2022. Evaluation of Host Resistance, Hydrated Lime, and Weed Control to Manage Clubroot in Canola. *Horticulturae* **8**(3).

Henshall IG, Spielmann T. 2023. Critical interdependencies between nutrient flux and drugs. *Trends in Parasitology* **39**(11): 936-944.

Hinsch J, Galuszka P, Tudzynski P. 2016. Functional characterization of the first filamentous fungal tRNA-isopentenyltransferase and its role in the virulence of Claviceps purpurea. *New Phytol* **211**(3): 980-992.

Hirose N, Takei K, Kuroha T, Kamada-Nobusada T, Hayashi H, Sakakibara H. 2008. Regulation of cytokinin biosynthesis, compartmentalization and translocation. *Journal of Experimental Botany* **59**(1): 75-83.

Hittorf M, Letsch-Praxmarer S, Windegger A, Bass D, Kirchmair M, Neuhauser S. 2020. Revised Taxonomy and Expanded Biodiversity of the Phytomyxea (Rhizaria, Endomyxa). *J Eukaryot Microbiol* **67**(6): 648-659.

Hollman KB, Hwang SF, Manolii VP, Strelkov SE. 2021. Pathotypes of Plasmodiophora brassicae collected from clubroot resistant canola (Brassica napus L.) cultivars in western Canada in 2017-2018. *Canadian Journal of Plant Pathology* **43**(4): 622-630.

Houten SM, Violante S, Ventura FV, Wanders RJ. 2016. The Biochemistry and Physiology of Mitochondrial Fatty Acid beta-Oxidation and Its Genetic Disorders. *Annu Rev Physiol* **78**: 23-44.

Howard RJ, Strelkov SE, Harding MW. 2010. Clubroot of cruciferous crops - new perspectives on an old disease. *Canadian Journal of Plant Pathology* **32**(1): 43-57.

Hu YJ, Shani E. 2023. Cytokinin activity - transport and homeostasis at the whole plant, cell, and subcellular levels. *New Phytologist* **239**(5): 1603-1608.

Humann JL, Lee T, Ficklin S, Main D. 2019. Structural and Functional Annotation of Eukaryotic Genomes with GenSAS. *Methods Mol Biol* **1962**: 29-51.

Hwang SF, Ahmed HU, Strelkov SE, Zhou Q, Gossen BD, Peng G, Turnbull GD, Fu H. 2018. Effects of Rate and Application Method on the Efficacy of Metam Sodium to Reduce Clubroot (Plasmodiophora brassicae) of Canola. *European Journal of Plant Pathology* **150**(2): 341-349.

Hwang SF, Ahmed HU, Zhou Q, Fu H, Turnbull GD, Fredua-Agyeman R, Strelkov SE, Gossen BD, Peng G. 2019. Influence of resistant cultivars and crop intervals on clubroot of canola. *Canadian Journal of Plant Science* **99**(6): 862-872.

Hwang SF, Ahmed HU, Zhou Q, Strelkov SE, Gossen BD, Peng G, Turnbull GD. 2011. Influence of cultivar resistance and inoculum density on root hair infection of canola (Brassica napus) by Plasmodiophora brassicae. *Plant Pathology* **60**(5): 820-829.

Hwang SF, Ahmed HU, Zhou Q, Turnbull GD, Strelkov SE, Gossen BD, Peng G. 2015. Effect of host and non-host crops on resting spore concentrations and clubroot of canola. *Plant Pathology* **64**(5): 1198-1206.

Hwang SF, Strelkov SE, Feng J, Gossen BD, Howard RJ. 2012. Plasmodiophora brassicae: a review of an emerging pathogen of the Canadian canola (Brassica napus) crop. *Molecular Plant Pathology* **13**(2): 105-113.

Ingram DS. 1969a. Abnormal Growth of Tissues Infected with Plasmodiophora brassicae. *J. geiz. Microbiol* **56**: 15.

Ingram DS. 1969b. Growth of Plasmodiophora brassicae in Host Callus. *J. gen. Microbiol* **55**: 12.

Inzé D, De Veylder L. 2006. Cell cycle regulation in plant development. *Annual Review of Genetics* **40**: 77-105.

Javed MA, Mukhopadhyay S, Normandeau E, Brochu AS, Perez-Lopez E. 2024. Telomere-to-telomere Genome Assembly of the Clubroot Pathogen Plasmodiophora Brassicae. *Genome Biol Evol* **16**(6).

Javed MA, Schwelm A, Zamani-Noor N, Salih R, Vañó MS, Wu JX, García MG, Heick TM, Luo CY, Prakash P, et al. 2023. The clubroot pathogen: A profile update. *Molecular Plant Pathology* **24**(2): 89-106.

Jimenez V, Mesones S. 2022. Down the membrane hole: Ion channels in protozoan parasites. *PLoS Pathog* **18**(12): e1011004.

Judge A, Dodd MS. 2020. Metabolism. *Understanding Biochemistry 6* **64**(4): 607-647.

Jump DB. 2009. Mammalian fatty acid elongases. *Methods Mol Biol* **579**: 375-389.

Kageyama K, Asano T. 2009. Life Cycle of Plasmodiophora brassicae. *Journal of Plant Growth Regulation* **28**(3): 203.

Kämper J, Kahmann R, Bölker M, Ma LJ, Brefort T, Saville BJ, Banuett F, Kronstad JW, Gold SE, Müller O, et al. 2006. Insights from the genome of the biotrophic fungal plant pathogen. *Nature* **444**(7115): 97-101.

Karim MM, Dakouri A, Zhang Y, Chen Q, Peng G, Strelkov SE, Gossen BD, Yu F. 2020. Two Clubroot-Resistance Genes, Rcr3 and Rcr9(wa), Mapped in Brassica rapa Using Bulk Segregant RNA Sequencing. *Int J Mol Sci* **21**(14).

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

Keen NT, Williams PH. 1969. Translocation of Sugars into Infected Cabbage Tissues During Clubroot Development. *Plant Physiol* **44**: 7.

Kettle AJ, Carere J, Batley J, Benfield AH, Manners JM, Kazan K, Gardiner DM. 2015. A γ-lactamase from Cereal infecting Fusarium spp. catalyses the first step in the degradation of the benzoxazolinone class of phytoalexins. *Fungal Genetics and Biology* **83**: 1-9.

Khalique A, Mattijssen S, Haddad AF, Chaudhry S, Maraia RJ. 2020. Targeting mitochondrial and cytosolic substrates of TRIT1 isopentenyltransferase: Specificity determinants and tRNA-i6A37 profiles. *PLoS genetics* **16**(4): e1008330-e1008330.

Kieber JJ, Schaller GE. 2018. Cytokinin signaling in plant development. *Development* **145**(4).

Kolde R 2018 pheatmap: Pretty Heatmaps.

Kollmer I, Novak O, Strnad M, Schmulling T, Werner T. 2014. Overexpression of the cytosolic cytokinin oxidase/dehydrogenase (CKX7) from Arabidopsis causes specific changes in root growth and xylem differentiation. *Plant J* **78**(3): 359-371.

Kong L, Li X, Zhan Z, Piao Z. 2022. Sugar Transporters in Plasmodiophora brassicae: Genome-Wide Identification and Functional Verification. *Int J Mol Sci* **23**(9).

Larsen B, Xu DY, Halkier BA, Nour-Eldin HH. 2017. Advances in methods for identification and characterization of plant transporter function. *Journal of Experimental Botany* **68**(15): 4045-4056.

Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. 2013. Web Apollo: a web-based genomic annotation editing platform. *Genome biology* **14**(8).

Lee SH, Stephens JL, Paul KS, Englund PT. 2006. Fatty acid synthesis by elongases in trypanosomes. *Cell* **126**(4): 691-699.

Lemoine R, La Camera S, Atanassova R, Déedaldéechamp F, Allario T, Pourtau N, Bonnemain JL, Laloi M, Coutos-Thévenot P, Maurousset L, et al. 2013. Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science* **4**.

Letunic I, Bork P. 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**(W1): W78-W82.

Li CL, Meng D, Pineros MA, Mao YX, Dandekar AM, Cheng LL. 2020. A Sugar Transporter Takes Up both Hexose and Sucrose for Sorbitol-Modulated In Vitro Pollen Tube Growth in Apple. *Plant Cell* **32**(2): 449-469.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**(18): 3094-3100.

Li H, Li X, Xuan Y, Jiang J, Wei Y, Piao Z. 2018. Genome Wide Identification and Expression Profiling of SWEET Genes Family Reveals Its Role During Plasmodiophora brassicae-Induced Formation of Clubroot in Brassica rapa. *Front Plant Sci* **9**: 207.

Li HQ, Benedito VA, Udvardi MK, Zhao PX. 2009. TransportTP: A two-phase classification approach for membrane transporter prediction and characterization. *Bmc Bioinformatics* **10**.

Liu LJ, Qin L, Zhou ZQ, Hendriks WGHM, Liu SY, Wei YD. 2020. Refining the Life Cycle of Plasmodiophora brassicae. *Phytopathology* **110**(10): 1704-1712.

Lorca GL, Barabote RD, Zlotopolski V, Tran C, Winnen B, Hvorup RN, Stonestrom AJ, Nguyen E, Huang LW, Kim DS, et al. 2007. Transport capabilities of eleven gram-positive bacteria: comparative genomic analyses. *Biochim Biophys Acta* **1768**(6): 1342-1366.

Ludwig-Müller J, Jülke S, Geiss K, Richter F, Mithöfer A, Sola I, Rusak G, Keenan S, Bulman S. 2015. A novel methyltransferase from the intracellular pathogen methylates salicylic acid. *Molecular Plant Pathology* **16**(4): 349-364.

Ludwig-Müller J, Prinsen E, Rolfe SA, Scholes JD. 2009. Metabolism and Plant Hormone Action During Clubroot Disease. *Journal of Plant Growth Regulation* **28**(3): 229-244.

Malinowski R, Novák O, Borhan MH, Spíchal L, Strnad M, Rolfe SA. 2016. The role of cytokinins in clubroot disease. *European Journal of Plant Pathology* **145**(3): 543-557.

Malinowski R, Singh D, Kasprzewska A, Blicharz S, Basinska-Barczak A. 2024. Vascular tissue - boon or bane? How pathogens usurp long-distance transport in plants and the defence mechanisms deployed to counteract them. *New Phytologist* **243**(6): 2075-2092.

Malinowski R, Smith JA, Fleming AJ, Scholes JD, Rolfe SA. 2012. Gall formation in clubroot-infected Arabidopsis results from an increase in existing meristematic activities of the host but is not essential for the completion of the pathogen life cycle. *Plant J* **71**(2): 226-238.

Martin RE. 2020. The transportome of the malaria parasite. *Biol Rev Camb Philos Soc* **95**(2): 305-332.

Martin RE, Henry RI, Abbey JL, Clements JD, Kirk K. 2005. The 'permeome' of the malaria parasite: an overview of the membrane transport proteins of Plasmodium falciparum. *Genome biology* **6**(3): 22.

Matsumoto-Kitano M, Kusumoto T, Tarkowski P, Kinoshita-Tsujimura K, Václavíková K, Miyawaki K, Kakimoto T. 2008. Cytokinins are central regulators of cambial activity. *Proceedings of the National Academy of Sciences of the United States of America* **105**(50): 20027-20031.

McDonnell E, Strasser K, Tsang A. 2018. Manual Gene Curation and Functional Annotation. *Methods Mol Biol* **1775**: 185-208.

Mcgee SL, van Denderen BJW, Howlett KF, Mollica J, Schertzer JD, Kemp BE, Hargreaves M. 2008. AMP-activated protein kinase regulates GLUT4 transcription by phosphorylating histone deacetylase 5. *Diabetes* **57**(4): 860-867.

McGrann GRD, Gladders P, Smith JA, Burnett F. 2016. Control of clubroot (Plasmodiophora brassicae) in oilseed rape using varietal resistance and soil amendments. *Field Crops Research* **186**: 146-156.

Medeiros TC, Mehra C, Pernas L. 2021. Contact and competition between mitochondria and microbes. *Current Opinion in Microbiology* **63**: 189-194.

Menda N, Buels RM, Tecle I, Mueller LA. 2008. A community-based annotation framework for linking solanaceae genomes with phenomes. *Plant Physiol* **147**(4): 1788-1799.

Meves H. 2008. Arachidonic acid and ion channels: an update. *British Journal of Pharmacology* **155**(1): 4-16.

Meyer A, Eskandari S, Grallath S, Rentsch D. 2006. AtGAT1, a high affinity transporter for gamma-aminobutyric acid in Arabidopsis thaliana. *J Biol Chem* **281**(11): 7197-7204.

Mithen R, Magrath R. 1992. A Contribution to the Life-History of Plasmodiophora-Brassicae - Secondary Plasmodia Development in Root Galls of Arabidopsis-Thaliana. *Mycological Research* **96**: 877-885.

Miyawaki K, Tarkowski P, Matsumoto-Kitano M, Kato T, Sato S, Tarkowska D, Tabata S, Sandberg G, Kakimoto T. 2006. Roles of Arabidopsis ATP/ADP isopentenyltransferases and tRNA isopentenyltransferases in cytokinin biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **103**(44): 16598-16603.

Montel-Hagen A, Kinet S, Manel N, Mongellaz C, Prohaska R, Battini JL, Delaunay J, Sitbon M, Taylor N. 2008. Erythrocyte glut1 triggers dehydroascorbic acid uptake in mammals unable to synthesize vitamin C. *Cell* **132**(6): 1039-1048.

Mothes K, Engelbrecht L. 1961. Kinetin-induced Directed Transport of Substances in Exercised Leaves in the Dark. *Phytochemistry* **1**: 58-64.

Muller P, Hilgenberg W. 1986. Isomers of zeatin and zeatin riboside in clubroot tissue: evidence for trans-zeatin biosynthesis by Plasmodiophora brasskae. *Physiol.Plant* **66**: 5.

Nakagawa T, Suzuki T, Murata S, Nakamura S, Hino T, Maeo K, Tabata R, Kawai T, Tanaka K, Niwa Y, et al. 2007. Improved Gateway binary vectors: high-performance

vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci Biotechnol Biochem* **71**(8): 2095-2100.

Neuhauser S, Kirchmair M, Bulman S, Bass D. 2014. Cross-kingdom host shifts of phytomyxid parasites. *BMC Evol Biol* **14**(1): 33.

Nguitragool W, Bokhari AA, Pillai AD, Rayavara K, Sharma P, Turpin B, Aravind L, Desai SA. 2011. Malaria parasite clag3 genes determine channel-mediated nutrient uptake by infected red blood cells. *Cell* **145**(5): 665-677.

Nilsson O, Olsson O. 2006. Getting to the root: The role of the Agrobacterium rhizogenes rol genes in the formation of hairy roots. *Physiologia Plantarum* **100**(3): 463-473.

Nishii K, Wright F, Chen YY, Moller M. 2018. Tangled history of a multigene family: The evolution of ISOPENTENYLTRANSFERASE genes. *PLoS One* **13**(8): e0201198.

Olszak M, Truman W, Stefanowicz K, Sliwinska E, Ito M, Walerowski P, Rolfe S, Malinowski R. 2019. Transcriptional profiling identifies critical steps of cell cycle reprogramming necessary for Plasmodiophora brassicae-driven gall formation in Arabidopsis. *Plant J* **97**(4): 715-729.

Opitz MW, Daneshkhah R, Lorenz C, Ludwig R, Steinkellner S, Wieczorek K. 2021. Serendipita indica changes host sugar and defense status in Arabidopsis thaliana: cooperation or exploitation? *Planta* **253**(3).

Ortega A, Matilla MA, Krell T. 2022. The Repertoire of Solute-Binding Proteins of Model Bacteria Reveals Large Differences in Number, Type, and Ligand Range. *Microbiology Spectrum* **10**(5).

Panda S, Zhou K. 2023. Engineering microbes to overproduce natural products as agrochemicals. *Synthetic and Systems Biotechnology* **8**(1): 79-85.

Peng G, Lahlali R, Hwang SF, Pageau D, Hynes RK, Mcdonald MR, Gossen BD, Strelkov SE. 2014. Crop rotation, cultivar resistance, and fungicides/biofungicides for managing clubroot (Plasmodiophora brassicae) on canola. *Canadian Journal of Plant Pathology* **36**: 99-112.

Peng G, Pageau D, Strelkov SE, Gossen BD, Hwang SF, Lahlali R. 2015. A >2-Year Crop Rotation Reduces Resting Spores of Plasmodiophora brassicae in Soil and the Impact of Clubroot on Canola. *European Journal of Agronomy* **70**: 78-84.

Pertea G, Pertea M. 2020. GFF Utilities: GffRead and GffCompare. *F1000Res* **9**.

Pertry I, Václavíková K, Depuydt S, Galuszka P, Spíchal L, Temmerman W, Stes E, Schmülling T, Kakimoto T, Van Montagu MCE, et al. 2009. Identification of cytokinins and their modus operandi to reshape the plant. *Proceedings of the National Academy of Sciences of the United States of America* **106**(3): 929-934.

Picci G, Marchesan S, Caltagirone C. 2022. Ion Channels and Transporters as Therapeutic Agents: From Biomolecules to Supramolecular Medicinal Chemistry. *Biomedicines* **10**(4).

Podolsky IA, Seppala S, Xu H, Jin YS, O'Malley MA. 2021. A SWEET surprise: Anaerobic fungal sugar transporters and chimeras enhance sugar uptake in yeast. *Metab Eng* **66**: 137-147.

Pracana R, Priyam A, Levantis I, Nichols RA, Wurm Y. 2017. The fire ant social chromosome supergene variant Sb shows low diversity but high divergence from SB. *Mol Ecol* **26**(11): 2864-2879.

Pratt-Hyatt M, Pai DA, Haeusler RA, Wozniak GG, Good PD, Miller EL, McLeod IX, Yates JR, Hopper AK, Engelke DR. 2013. Mod5 protein binds to tRNA gene complexes and affects local transcriptional silencing. *Proceedings of the National Academy of Sciences of the United States of America* **110**(33): E3081-E3089.

Putnam RW 2012. Intracellular pH Regulation. *Cell Physiology Source Book*, 303-321.

Qin H, Yin W, Luo C, Liu L. 2024. The Identification, Characterization, and Functional Analysis of the Sugar Transporter Gene Family of the Rice False Smut Pathogen, Villosiclava virens. *Int J Mol Sci* **25**(1).

Quick M, Javitch JA. 2007. Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proceedings of the National Academy of Sciences of the United States of America* **104**(9): 3603-3608.

Rahman H. 2013. Review: Breeding spring canola (L) by the use of exotic germplasm. *Canadian Journal of Plant Science* **93**(3): 363-373.

Remize M, Brunel Y, Silva JL, Berthon JY, Filaire E. 2021. Microalgae n-3 PUFAs Production and Use in Food and Feed Industries. *Marine Drugs* **19**(2).

Rocherieux J, Glory P, Giboulot A, Boury S, Barbeyron G, Thomas G, Manzanares-Dauleux MJ. 2004. Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in Brassica oleracea. *Theoretical and Applied Genetics* **108**(8): 1555-1563.

Rolfe SA, Strelkov SE, Links MG, Clarke WE, Robinson SJ, Djavaheri M, Malinowski R, Haddadi P, Kagale S, Parkin IA, et al. 2016. The compact genome of the plant pathogen Plasmodiophora brassicae is adapted to intracellular interactions with host Brassica spp. *BMC Genomics* **17**: 272.

Rudsari MK, Zouhar M, Manasova M, Li TD. 2024. Biocontrol potential of cell-free supernatant of Paenibacillus chitinolyticus against Plasmodiophora brassicae in two important Brassica species. *European Journal of Plant Pathology* **170**(1): 119-130.

Russum S, Lam KJK, Wong NA, Iddamsetty V, Hendargo KJ, Wang J, Dubey A, Zhang Y, Medrano-Soto A, Saier MH, Jr. 2021. Comparative population genomic analyses of transporters within the Asgard archaeal superphylum. *PLoS One* **16**(3): e0247806.

Saier MH, Jr. 2000. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**(2): 354-411.

Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, Lam KJK, Tian N, Russum S, Wang J, et al. 2021. The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res* **49**(D1): D461-D467.

Saito M, Kubo N, Matsumoto S, Suwabe K, Tsukada M, Hirai M. 2006. Fine mapping of the clubroot resistance gene, Crr3, in Brassica rapa. *Theoretical and Applied Genetics* **114**(1): 81-91.

Sakakibara H. 2003. Nitrate-specific and cytokinin-mediated nitrogen signaling pathways in plants. *Journal of Plant Research* **116**(3): 253-257.

Sakakibara H. 2006. Cytokinins: activity, biosynthesis, and translocation. *Annual review of plant biology* **57**(1): 431-449.

Sakakibara H. 2006. Cytokinins: Activity, biosynthesis, and translocation. *Annual Review of Plant Biology* **57**: 431-449.

Sakakibara H, Kasahara H, Ueda N, Kojima M, Takei K, Hishiyama S, Asami T, Okada K, Kamiya Y, Yamaya T, et al. 2005. increases cytokinin production in plastids by modifying the biosynthetic pathway in the host plant. *Proceedings of the National Academy of Sciences of the United States of America* **102**(28): 9972-9977.

Sakakibara H, Takei K, Hirose N. 2006. Interactions between nitrogen and cytokinin in the regulation of metabolism and development. *Trends in Plant Science* **11**(9): 440-448.

Santos HJ, Nozaki T. 2021. Interorganellar communication and membrane contact sites in protozoan parasites. *Parasitology International* **83**: 102372.

Sauer N, Stolz J. 1994. SUC1 and SUC2: two sucrose transporters from Arabidopsis thaliana; expression and characterization in baker's yeast and identification of the histidine-tagged protein. *Plant J* **6**(1): 67-77.

Savchenko T, Walley JW, Chehab EW, Xiao YM, Kaspi R, Pye MF, Mohamed ME, Lazarus CM, Bostock RM, Dehesh K. 2010. Arachidonic Acid: An Evolutionarily Conserved Signaling Molecule Modulates Plant Stress Signaling Networks. *Plant Cell* **22**(10): 3193-3205.

Schaarschmidt S, Kopka J, Ludwig-Müller J, Hause B. 2007. Regulation of arbuscular mycorrhization by apoplastic invertases:: enhanced invertase activity in the leaf apoplast affects the symbiotic interaction. *Plant Journal* **51**(3): 390-405.

Schluepmann H, Pellny T, van Dijken A, Smeekens S, Paul M. 2003. Trehalose 6-phosphate is indispensable forcarbohydrate utilization and growth inArabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **100**(11): 6849-6854.

Schluepmann H, van Dijken A, Aghdasi M, Wobbes B, Paul M, Smeekens S. 2004. Trehalose mediated growth inhibition of Arabidopsis seedlings is due to trehalose-6-phosphate accumulation. *Plant Physiol* **135**(2): 879-890.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**(7): 671-675.

Schuler D, Wahl R, Wippel K, Vranes M, Munsterkotter M, Sauer N, Kamper J. 2015. Hxt1, a monosaccharide transporter and sensor required for virulence of the maize pathogen Ustilago maydis. *New Phytol* **206**(3): 1086-1100.

Schuller A, Kehr J, Ludwig-Muller J. 2014. Laser microdissection coupled to transcriptional profiling of Arabidopsis roots inoculated by Plasmodiophora brassicae indicates a role for brassinosteroids in clubroot formation. *Plant Cell Physiol* **55**(2): 392-411.

Schwelm A, Fogelqvist J, Knaust A, Julke S, Lilja T, Bonilla-Rosso G, Karlsson M, Shevchenko A, Dhandapani V, Choi SR, et al. 2015. The Plasmodiophora brassicae genome reveals insights in its life cycle and ancestry of chitin synthases. *Sci Rep* **5**: 11153.

Seif E, Hallberg BM. 2009. RNA-protein mutually induced fit: structure of Escherichia coli isopentenyl-tRNA transferase in complex with tRNA(Phe). *J Biol Chem* **284**(11): 6600-6604.

Seppala S, Gierke T, Schauer EE, Brown JL, O'Malley MA. 2023. Identification and expression of small multidrug resistance transporters in early-branching anaerobic fungi. *Protein Sci* **32**(9): e4730.

Seppala S, Solomon KV, Gilmore SP, Henske JK, O'Malley MA. 2016. Mapping the membrane proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and transporters. *Microb Cell Fact* **15**(1): 212.

Shahan R, Hsu CW, Nolan TM, Cole BJ, Taylor IW, Greenstreet L, Zhang S, Afanassiev A, Vlot AHC, Schiebinger G, et al. 2022. A single-cell root atlas reveals developmental trajectories in wild-type and cell identity mutants. *Developmental Cell* **57**(4): 543-+.

Sharma K, Gossen BD, McDonald MR. 2011. Effect of Temperature on Cortical Infection by Plasmodiophora brassicae and Clubroot Severity. *Phytopathology* **101**(12): 1424-1432.

Shen J, Zeng Y, Zhuang X, Sun L, Yao X, Pimpl P, Jiang L. 2013. Organelle pH in the Arabidopsis endomembrane system. *Mol Plant* **6**(5): 1419-1437.

Siemens J, Gonzalez MC, Wolf S, Hofmann C, Greiner S, Du Y, Rausch T, Roitsch T, Ludwig-Muller J. 2011. Extracellular invertase is involved in the regulation of clubroot disease in Arabidopsis thaliana. *Mol Plant Pathol* **12**(3): 247-262.

Siemens J, Keller I, Sarx J, Kunz S, Schuller A, Nagel W, Schmulling T, Parniske M, Ludwig-Muller J. 2006. Transcriptome Analysis of Arabidopsis Clubroots Indicate a Key Role for Cytokinins in Disease Development. *MPMI* **19**(5): 15.

Siemens J, Nagel M, Ludwig-Mu¨ ller J, Sacrista´n M, D. 2002. The Interaction of Plasmodiophora brassicae and Arabidopsis thaliana: Parameters for Disease Quantification and Screening of Mutant Lines. *J. Phytopathology* **150**: 13.

Skinner OS, Blanco-Fernandez J, Goodman RP, Kawakami A, Shen H, Kemeny LV, Joesch-Cohen L, Rees MG, Roth JA, Fisher DE, et al. 2023. Salvage of ribose from uridine or RNA supports glycolysis in nutrient-limited conditions. *Nat Metab* **5**(5): 765-776.

Smolko A, Jülke S, Benade F, Siroká J, Pencik A, Vuk T, Bauer N, Salopek-Sondi B, Ludwig-Müller J. 2024. The GH3 Protein from the Clubroot Pathogen Plasmodiophora brassicae Causes Hormone-Related Phenotypes in Arabidopsis thaliana. *Journal of Plant Growth Regulation*.

Soderberg T, Poulter CD. 2001. dimethylallyl diphosphate:tRNA dimethylallyltransferase:: Site-directed mutagenesis of highly conserved residues. *Biochemistry* **40**(6): 1734-1740.

Soprano AS, Smetana JHC, Benedetti CE. 2018. Regulation of tRNA biogenesis in plants and its link to plant growth and response to pathogens. *Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms* **1861**(4): 344-353.

Sosso D, van der Linde K, Bezrutczyk M, Schuler D, Schneider K, Kamper J, Walbot V. 2019. Sugar Partitioning between Ustilago maydis and Its Host Zea mays L during Infection. *Plant Physiol* **179**(4): 1373-1385.

Spallek T, Gan P, Kadota Y, Shirasu K. 2018. Same tune, different song-cytokinins as virulence factors in plant-pathogen interactions? *Current Opinion in Plant Biology* **44**: 82-87.

Spallek T, Gan P, Kadota Y, Shirasu K. 2018. Same tune, different song—cytokinins as virulence factors in plant–pathogen interactions? *Current Opinion in Plant Biology* **44**: 82-87.

Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, van Themaat EVL, Brown JKM, Butcher SA, Gurr SJ, et al. 2010. Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism. *Science* **330**(6010): 1543-1546.

Stillwell W 2016. Membrane transport. In: Stillwell W ed. *An Introduction to Biological Membranes*: Elsevier.

Stjelja S, Fogelqvist J, Tellgren-Roth C, Dixelius C. 2019. The architecture of the Plasmodiophora brassicae nuclear and mitochondrial genomes. *Sci Rep* **9**(1): 15753.

Strelkov SE, Hwang SF. 2014. Clubroot in the Canadian canola crop: 10 years into the outbreak. *Canadian Journal of Plant Pathology* **36**: 27-36.

Strelkov SE, Hwang SF, Manolii VP, Cao T, Feindel D. 2016. Emergence of New Virulence Phenotypes of Plasmodiophora brassicae on Canola (Brassica napus) in Alberta, Canada. *European Journal of Plant Pathology* **145**(3): 517-529.

Struck C. 2015. Amino acid uptake in rust fungi. *Front Plant Sci* **6**: 40.

Struck C, Ernst M, Hahn M. 2002. Characterization of a developmentally regulated amino acid transporter (AAT1p) of the rust fungus Uromyces fabae. *Mol Plant Pathol* **3**(1): 23-30.

Struck C, Mueller E, Martin H, Lohaus G. 2004. The Uromyces fabae UfAAT3 gene encodes a general amino acid permease that prefers uptake of in planta scarce amino acids. *Mol Plant Pathol* **5**(3): 183-189.

Sudheer S, Bai RG, Usmani Z, Sharma M. 2020. Insights on Engineered Microbes in Sustainable Agriculture: Biotechnological Developments and Future Prospects. *Current Genomics* **21**(5): 321-333.

Suzuki G, Shimazu N, Tanaka M. 2012. A Yeast Prion, Mod5, Promotes Acquired Drug Resistance and Cell Survival Under Environmental Stress. *Science* **336**(6079): 355-359.
Takei K, Sakakibara H, Taniguchi M, Sugiyama T. 2001. Nitrogen-dependent accumulation of cytokinins in root and the translocation to leaf: implication of cytokinin species that induces gene expression of maize response regulator. *Plant Cell Physiol* **42**(1): 85-93.
Tamura K, Stecher G, Kumar S. 2021. MEGA11 Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* **38**(7): 3022-3027.
Tarkowski P, Ge LY, Yong JWH, Tan SN. 2009. Analytical methods for cytokinins. *Trac-Trends in Analytical Chemistry* **28**(3): 323-335.
Thevelein JM, Voordeckers K. 2009. Functioning and Evolutionary Significance of Nutrient Transceptors. *Molecular Biology and Evolution* **26**(11): 2407-2414.
Tolerico LH, Benko AL, Aris JP, Stanford DR, Martin NC, Hopper AK. 1999. Mod5p-II contains sequences antagonistic for nuclear and cytosolic locations. *Genetics* **151**(1): 57-75.
Vañó MS, Nourimand M, MacLean A, Pérez-López E. 2023. Getting to the root of a club-Understanding developmental manipulation by the clubroot pathogen. *Seminars in Cell & Developmental Biology* **148**: 22-32.
Voegele TRa, Struck C, Hahn M, Mendgen K. 2001. The role of haustoria in sugar supply during infection of broad bean by the rust fungus Uromyces fabae. *Pnas* **98**(14): 6.
Voorrips RE, Jongerius MC, Kanne HJ. 1997. Mapping of two genes for resistance to clubroot (Plasmodiophora brassicae) in a population of doubled haploid lines of Brassica oleracea by means of RFLP and AFLP markers. *Theoretical and Applied Genetics* **94**(1): 75-82.
Vormittag S, Ende RJ, Derre I, Hilbi H. 2023. Pathogen vacuole membrane contact sites - close encounters of the fifth kind. *Microlife* **4**: uqad018.
Wahl R, Wippel K, Goos S, Kamper J, Sauer N. 2010. A novel high-affinity sucrose transporter is required for virulence of the plant pathogen Ustilago maydis. *PLoS Biol* **8**(2): e1000303.
Walerowski P, Gundel A, Yahaya N, Truman W, Sobczak M, Olszak M, Rolfe S, Borisjuk L, Malinowski R. 2018. Clubroot Disease Stimulates Early Steps of Phloem Differentiation and Recruits SWEET Sucrose Transporters within Developing Galls. *Plant Cell* **30**(12): 3058-3073.
Wallenhammar AC. 1996. Prevalence of Plasmodiophora brassicae in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathology* **45**(4): 710-719.
Walters DR, McRoberts N, Fitt BD. 2008. Are green islands red herrings? Significance of green islands in plant interactions with pathogens and pests. *Biol Rev Camb Philos Soc* **83**(1): 79-102.
Williams GW, Davis PA, Rogers AS, Bieri T, Ozersky P, Spieth J. 2011. Methods and strategies for gene structure curation in WormBase. *Database (Oxford)* **2011**: baq039.
Williams PH, Reddy N, Strandberg J. 1969. Growth of noninfected and PIasmodiophora brassicae infected cabbage callus in culture. *CANADIAN JOURNAL OF BOTANY* **47**(8): 7.
Wunderlich J. 2022. Updated List of Transport Proteins in Plasmodium falciparum. *Front Cell Infect Microbiol* **12**: 926541.
Ye AY, Liu QR, Li CY, Zhao M, Qu H. 2014. Human Transporter Database: Comprehensive Knowledge and Discovery Tools in the Human Transporter Genes (vol 9, e88883, 2014). *PLoS One* **9**(5).
Yonekura-Sakakibara K, Kojima M, Yamaya T, Sakakibara H. 2004. Molecular characterization of cytokinin-responsive histidine kinases in maize: Differential ligand

preferences, response to cis-zeatin and the subcellular localizations. *Plant and Cell Physiology* **45**: S41-S41.

Yu FQ, Zhang XG, Peng G, Falk KC, Strelkov SE, Gossen BD. 2017. Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of Plasmodiophora brassicae in Brassica rapa. *Scientific Reports* **7**.

Yuan Q, Yan Y, Sohail MA, Liu H, Huang J, Hsiang T, Zheng L. 2021. A Novel Hexose Transporter ChHxt6 Is Required for Hexose Uptake and Virulence in Colletotrichum higginsianum. *Int J Mol Sci* **22**(11).

Zamani-Noor N, Brand S, Söchting HP. 2022. Effect of pathogen virulence on pathogenicity, host range and reproduction of Plasmodiophora brassicae, the causal agent of clubroot disease. *Plant Disease* **106**(1): 57-64.

Zhang RY, Wang YR, Liu RL, Rhee SK, Zhao GP, Quan ZX. 2024. Metagenomic characterization of a novel non-ammonia-oxidizing Thaumarchaeota from hadal sediment. *Microbiome* **12**(1): 7.

Zhang T, Zhao Z, Zhang CY, Pang WX, Choi SR, Lim YP, Piao ZY. 2014. Fine genetic and physical mapping of the CRb gene conferring resistance to clubroot disease in Brassica rapa. *Molecular Breeding* **34**(3): 1173-1183.

Zhang W, Wang S, Yu F, Tang J, Yu L, Wang H, Li J. 2019. Genome-Wide Identification and Expression Profiling of Sugar Transporter Protein (STP) Family Genes in Cabbage (Brassica oleracea var. capitata L.) Reveals their Involvement in Clubroot Disease Responses. *Genes (Basel)* **10**(1).

Zhu M, He Y, Li Y, Ren T, Liu H, Huang J, Jiang D, Hsiang T, Zheng L. 2019. Two New Biocontrol Agents Against Clubroot Caused by Plasmodiophora brassicae. *Front Microbiol* **10**: 3099.

Zuzak KA, Strelkov SE, Turnbull GD, Manolii VP, Hwang SF. 2022. Soil fumigation with Vapam (metam sodium) to control clubroot (Plasmodiophora brassicae) of canola (Brassica napus). *Canadian Journal of Plant Science*.