Investigating phytopathogen putative effector proteins that target starch granule initiation

Sally Louise Jones

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

John Innes Centre

DECEMBER 2023

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

i) Abstract

The chloroplast is a key target of pathogen effectors. Two putative effector proteins, ChEC153 and HaRxL94b from *Colletotrichum higginsianum* and *Hyaloperonospora arabidopsidis*, respectively, localise to discrete, chloroplastic puncta when expressed in *N. benthamiana*. These punctate sub-chloroplastic localisations are remarkably similar to those of several proteins involved in the initiation of starch granules. I thereby sought to evaluate the putative effectors in the context of starch granule initiation, observing both to co-localise with key starch granule initiation protein MYOSIN-RESEMBLING CHLOROPLAST PROTEIN (MRC). Further, *ChEC153* expression *in planta* induces the formation of small starch granules.

I evaluated the role of *ChEC153* in infection, finding that expression of the putative effector increases host susceptibility to *C. higginsianum*. Additionally, I found that *mrc* mutants have increased resistance to infection, pointing to the potential for starch granule initiation to play a role in host susceptibility. I identified CRBIC (*AT1G53120*) as a potential host interactor of ChEC153 through immunoprecipitation-mass spectrometry, and found it was recruited to ChEC153-puncta during co-expression, supporting their interaction.

Following publication of an alternative gene model for *ChEC153*, I experimentally determined the true gene model: *ChEC153.3.* I also employed dual-transcriptomic analysis of Arabidopsis infected with *C. higginsianum* strains with and without *ChEC153.3* present to probe processes targeted by the putative effector in either the host or pathogen, identifying a number of genes differentially expressed between the two infections. These genes point to the importance of *ChEC153.3* in infection, and support a carbohydrate-targeting role for the putative effector.

In addition to investigating the two putative effector proteins, this thesis highlights new links between starch granule initiation and infection, as well as suggesting associations between the sites of starch granule initiation and plastidial transcription.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

ii) List of Contents

i	Abstract	2
ii	List of Contents	3
iii	List of Figures	10
iv	List of Tables	13
v	List of Appendices	13
vi	Abbreviations	15
vii	Acknowledgements	19
1	INTRODUCTION	21
1.1	Plant-pathogen interactions	22
1.1.1	An introduction to phytopathogens	22
1.1.2	The plant immune system is multi-layered and complex	23
1.1.2.1	Pattern-triggered immunity - PTI	23
1.1.2.2	Effector-triggered susceptibility - ETS	24
1.1.2.3	Effector-triggered immunity - ETI	25
1.1.3	Pathogen effectors	26
1.2	Fungal plant pathogens	28
1.2.1	Fungi as phytopathogens	28
1.2.2	Fungal effectors	28
1.3	*Colletotrichum higginsianum*	31
1.3.1	*C. higginsianum* lifecycle and infection strategy	31
1.3.2	*C. higginsianum* nutrient uptake and metabolism	34
1.3.3	*C. higginsianum* strains, genetic resources and genetic manipulation	37
1.3.4	The effector repertoire of *C. higginsianum*	37
1.4	Oomycete plant pathogens	38
1.4.1	Oomycetes as phytopathogens	38
1.4.2	Oomycetes produce an array of effectors, including RxLR effectors	39
1.5	*Hyaloperonospora arabidopsidis*	40
1.5.1	*H. arabidopsidis* lifecycle and infection strategy	41
1.5.2	*H. arabidopsidis* nutrient uptake and metabolism	42
1.5.3	*H. arabidopsidis* genetic resources	43
1.5.4	*H. arabidopsidis* effectors	43
1.6	The chloroplast during infection: a target of pathogen effectors	44
1.6.1	Photosynthesis generally appears suppressed during infection	44
1.6.2	The chloroplast is a key site for production of defence signals	45

1.6.3	Chloroplast repositioning in response to infection	46
1.6.4	A number of effector proteins localise to the chloroplast	47
1.7	Carbohydrates are stored as transitory starch in the chloroplasts of leaves	47
1.7.1	Starch synthesis and degradation	49
1.7.2	Starch granule initiation	50
1.8	Carbon availability during infection: implications for host and pathogen	52
1.8.1	Host carbohydrates as a resource for pathogens	52
1.8.2	Starch hyperaccumulates during infection in some pathosystems	54
1.9	Putative effectors ChEC153 and HaRxL94b localise to chloroplastic puncta	55
1.9.1	*C. higginsianum* putative effector ChEC153	55
1.9.2	*H. arabidopsidis* putative effector HaRxL94b	58
1.10	Proteins that form discrete, chloroplastic puncta	60
1.10.1	Nucleoid-associated proteins and plastidial transcription machinery	60
1.10.2	Punctate proteins implicated in infection or immunity	61
1.10.3	Plastid division machinery	61
1.10.4	Other chloroplastic punctate plant proteins	62
1.10.5	Punctate chloroplastic proteins in *Chlamydomonas*	62
1.10.6	Summary	62
1.11	Project aims	63
2	MATERIALS & METHODS	64
2.1	Arabidopsis mutant and transgenic lines	64
2.2	Plant growth	65
2.3	Bacterial strains and growth conditions	65
2.4	Bacterial transformation	66
2.5	Cloning	66
2.5.1	Amplification reactions	66
2.5.2	Golden Gate reactions	67
2.5.3	Construct verification	67
2.6	*Agrobacterium*-mediated transient transformation of *N. benthamiana*	67
2.7	Confocal microscopy	68
2.8	Protein purification and Western blotting	69
2.8.1	Western blotting solutions and SDS polyacrylamide gels	70
2.9	Statistics, data visualisation, and figure preparation	71
3	RESULTS 1: Probing the cell biology of putative effectors ChEC153 and HaRxL94b	72
3.1	INTRODUCTION	73
3.1.1	Protein translocation into the chloroplast: a route for pathogen effectors?	73
3.1.1.1	Overview of the general import pathway	73

3.1.1.2	Regulation of chloroplast protein import	74
3.1.1.3	Effector translocation into the chloroplast	74
3.1.2	Starch granules in Arabidopsis: wild-type and granule initiation mutant lines	75
3.1.3	Starch-related proteins do not necessarily contain starch-binding or catalytic domains	76
3.1.4	Chapter aims	76
3.2	MATERIALS & METHODS	77
3.2.1	Particle bombardment	77
3.2.2	Starch and sugar quantification from *N. benthamiana* and Arabidopsis	79
3.2.2.1	*N. benthamiana*	79
3.2.2.2	Arabidopsis	79
3.2.2.3	Starch and sugar extraction and quantification	79
3.2.3	Starch purification from Arabidopsis rosettes	80
3.2.4	Determination of starch granule particle size distributions	80
3.2.5	Scanning electron microscopy (SEM)	82
3.3	RESULTS	82
3.3.1	ChEC153 and HaRxL94b localise to the same discrete, chloroplastic puncta	82
3.3.2	*In silico* characterisation of the putative effectors ChEC153 and HaRxL94b	84
3.3.2.1	ChEC153 is predicted to be secreted from *C. higginsianum*	84
3.3.2.2	ChEC153 is a putative cytochrome P450 protein	84
3.3.2.2.1	Structure predictions of ChEC153	85
3.3.2.3	ChEC153 lacks a canonical chloroplast transit peptide	86
3.3.2.4	HaRxL94b lacks clear predicted characterised domains and structural motifs	86
3.3.2.4.1	Structure prediction of HaRxL94b	86
3.3.2.5	HaRxL94b has a canonical N-terminal chloroplast transit peptide	88
3.3.2.6	Further inferences about HaRxL94b from the literature	88
3.3.3	The chloroplastic localisation of HaRxL94b is dependent on its N-terminal sequence	89
3.3.4	Do ChEC153 and HaRxL94b target sites of starch granule initiation?	93
3.3.5	ChEC153 and HaRxL94b co-localise with TaCSP41a	95
3.3.6	The localisation of ChEC153 and HaRxL94b is partly associated with nucleoids	97
3.3.7	Transient expression of *ChEC153* induces formation of small starch granules	100
3.3.8	ChEC153-eGFP does not appear to be strongly associated with starch granules	104
3.3.9	The localisations of ChEC153 and HaRxL94b do not require MRC, SS4, or starch	105
3.3.10	Expression of *ChEC153* in Arabidopsis and *N. benthamiana* does not have a clear impact on the total starch or sugar content of leaves	106
3.3.10.1	Quantification of total starch content	107

3.3.10.2	Quantification of soluble sugars	110
3.3.11	The starch granules of Arabidopsis leaves expressing *ChEC153-eGFP* are slightly larger than those of the wild-type	114
3.3.11.1	Particle size analysis of purified starch	114
3.3.11.2	Analysis of purified starch using scanning electron microscopy	117
3.4	DISCUSSION	119
3.4.1	What is the nature of the punctate localisation of ChEC153 and HaRxL94b?	119
3.4.2	Does ChEC153 impact total starch content or starch granule size?	121
3.4.3	ChEC153-induced formation of GBSS puncta: formation of starch granules of mislocalisation of GBSS?	122
3.4.4	Chapter conclusions	125
4	RESULTS 2: EXPLORING THE ROLE OF ChEC153 IN INFECTION	126
4.1	INTRODUCTION	127
4.1.1	Infection assays	127
4.1.2	Genetic transformation of *C. higginsianum*	127
4.1.2.1	Targeted *C. higginsianum* transformation	128
4.1.2.2	Untargeted *C. higginsianum* transformation	129
4.1.2.3	Selection of *C. higginsianum* transformants	129
4.1.3	Chapter aims	130
4.2	MATERIALS & METHODS	130
4.2.1	Culturing of *C. higginsianum*	130
4.2.2	*C. higginsianum* infection assays	130
4.2.3	Generation of *C. higginsianum* gene knockout strains	133
4.2.3.1	Cloning of fungal gene knockout plasmids	133
4.2.3.2	*C. higginsianum* transformation	135
4.2.3.3	Extraction of *C. higginsianum* genomic DNA	136
4.2.3.4	Screening of *C. higginsianum* transformant colonies	137
4.2.4	Lugol staining of infected Arabidopsis leaves	137
4.2.5	Immunoprecipitation-mass spectrometry (IP-MS)	137
4.2.6	Arabidopsis genotyping: genomic DNA extraction and genotyping PCRs	139
4.3	RESULTS	140
4.3.1	Host expression of *ChEC153* increases susceptibility to *C. higginsianum*	140
4.3.2	*ΔChEC153 C. higginsianum* strains do not have altered virulence	142
4.3.2.1	Generation of *ChEC153*-knockout *C. higginsianum* strains	142
4.3.2.2	*In vitro* phenotypes of *ΔChEC153* strains	145
4.3.2.3	*ΔChEC153* strains have no clear virulence phenotype on Arabidopsis, nor confer a macroscopic starch phenotype during infection	148

4.3.3	*mrc* plants show increased resistance *C. higginsianum*	153
4.3.4	IP-MS reveals potential host interactors of ChEC153 in the Arabidopsis host	155
4.3.5	Candidate ChEC153-interactor CPNB3 localises to the chloroplast, while EIF2-A2 localises to the cytoplasm	156
4.3.6	Candidate ChEC153-interactor CRBIC is recruited to ChEC153 puncta	159
4.3.7	CRBIC localises to mitochondria, as well as chloroplasts	162
4.3.8	*crbic* Arabidopsis mutants do not have altered susceptibility to *C. higginsianum*	164
4.3.9	Extraction of ChEC153-eGFP from *N. benthamiana*	167
4.4	DISCUSSION	170
4.4.1	ChEC153 contributes to host susceptibility	170
4.4.2	CRBIC interacts with ChEC153 at chloroplastic puncta, but loss of *CRBIC* from the host does not alter susceptibility to infection by *C. higginsianum*	173
4.4.3	ChEC153 may target CPNB3, while EIF2-A2 represents an unlikely interactor	173
4.4.4	Candidate interactors of ChEC153 include PEP complex proteins	174
4.4.5	The *mrc* starch granule initiation mutant has enhanced resistance to infection	175
4.4.6	Chapter conclusions	176
5	RESULTS 3: ELUCIDATING THE *ChEC153* GENE MODEL AND EXPLORATORY TRANSCRIPTOME ANALYSIS	177
5.1	INTRODUCTION	178
5.1.1	*C. higginsianum* genetic resources	178
5.1.1.1	Genome annotations can present bottlenecks for the study of non-model organisms	178
5.1.1.2	*In silico* gene annotation pipelines and the annotation of *C. higginsianum* genome assemblies	179
5.1.1.3	Gene model validation and 5′ rapid amplification of cDNA ends	180
5.1.2	Dual transcriptome analysis of pathosystems	183
5.1.3	Chapter aims	184
5.2	MATERIALS & METHODS	184
5.2.1	Reverse transcription PCR (RT-PCR) for gene model validation	184
5.2.1.1	Initial RT-PCR	184
5.2.1.2	Additional RT-PCR	185
5.2.2	5′ RACE	185
5.2.3	RNA-Sequencing	185
5.2.3.1	RNA-Seq sample preparation	186
5.2.3.2	RNA sample validation	186
5.2.3.2.1	Reverse-transcription quantitative PCR (RT-qPCR)	186

5.2.3.2.2	Reverse-transcription PCR (RT-PCR)	187
5.2.3.3	RNA-Seq and data pre-processing	187
5.2.3.4	Data analysis	188
5.3	RESULTS	188
5.3.1	The *ChEC153* gene model differs between published genome assemblies	188
5.3.2	Elucidating the true gene model of the candidate effector gene *ChEC153*	191
5.3.2.1	RT-PCR analyses of *C. higginsianum* cDNA	191
5.3.2.2	Sequencing of cloned regions of *ChEC153* from cDNA	193
5.3.2.3	Localisation of ChEC153.2a in *N. benthamiana*	195
5.3.2.4	Determining the start site of *ChEC153*	195
5.3.2.4.1	Using 5' RACE as a strategy to determine the *ChEC153* transcription start site	195
5.3.2.4.2	Further RT-PCR-based analysis of *C. higginsianum* cDNA	197
5.3.2.5	Summary of the gene model for *ChEC153: ChEC153.3*	201
5.3.2.6	Protein ChEC153.3 is encoded by the new gene model *ChEC153.3*	202
5.3.3	Dual transcriptome analysis of Arabidopsis infected with *C. higginsianum* strains *ΔChEC153-2* and *ΔChKu80*	205
5.3.3.1	RNA-Seq sample preparation and validation	205
5.3.3.2	RNA-Seq data preliminary analyses	207
5.3.3.3	Principal component analysis reveals little variation between samples based on Arabidopsis reads	211
5.3.3.4	Principal component analysis reveals little variation between *ΔChKu80* and *ΔChEC153-2* samples in terms of fungal reads	213
5.3.3.5	A number of Arabidopsis and *C. higginsianum* genes are differentially expressed between *ΔChKu80* and *ΔChEC153-2* infections	215
5.3.3.5.1	Arabidopsis genes differentially expressed during infection with *ΔChEC153-2* versus *ΔChKu80* strains	215
5.3.3.5.2	*C. higginsianum* genes differentially expressed between *ΔChEC153-2* and *ΔChKu80* strains during infection	217
5.3.3.6	RNA-Seq reads from strain *ΔChEC111* map to *ChEC153.3*	218
5.4	DISCUSSION	220
5.4.1	The true gene model for *ChEC153*, *ChEC153.3*, differs from both published gene models	220
5.4.1.1	Is ChEC153.3 likely to be secreted from the fungus?	220
5.4.2	What are the effects of *ChEC153.3* in the host or fungus?	221

5.4.2.1	Host gene expression differences in *ΔChEC153-2* and *ΔChKu80* infections	221
5.4.2.2	Pathogen gene expression differences in *ΔChEC153-2* versus *ΔChKu80*	224
5.4.2.3	Dual RNA-Seq enables pathosystem analysis, but challenges posed by variable infection rates may confer limitations to the data	225
5.4.3	Chapter conclusions	226
6	**GENERAL DISCUSSION**	227
6.1	A role for starch granule initiation in pathogenicity	227
6.1.1	ChEC153.1 perturbs starch granule initiation and may make host carbon more accessible to the pathogen	228
6.1.2	The involvement of starch granule initiation in disease susceptibility is highlighted by granule initiation mutants	228
6.1.3	Starch is likely to act as a carbon source for *C. higginsianum*	230
6.1.4	Granule size may have implications in pathogen targeting of storage starches	230
6.1.5	Potential impacts of starch targeting on host defence	231
6.2	Verification of ChEC153.3 and HaRxL94b as effectors	232
6.2.1	Experimental confirmation of putative effector secretion and translocation	232
6.2.2	Can ChEC153.1 characterisation be used to infer roles of ChEC153.3?	233
6.2.3	Dual transcriptomic analysis supports the involvement of *ChEC153.3* in infection	234
6.2.4	Putative effector HaRxL94b: similarities and differences to ChEC153.1	235
6.2.4.1	How conserved is the targeting of punctate chloroplastic processes by effectors?	235
6.2.4.2	ChEC153.1 and HaRxL94b localisations and biological functions may differ	236
6.3	Links between starch granule initials and proteins relating to RNA-processing and plastid transcription	237
6.4	Importance of gene model validation	238
6.5	Final conclusions	239
7	**Bibliography**	240
viii	Appendices	269

iii) List of Figures

Figure 1.1	Schematic depicting the infection lifecycle of *C. higginsianum*
Figure 1.2	Schematic depicting the infection lifecycle of *H. arabidopsidis*
Figure 1.3	Schematic depicting the structure of starch
Figure 1.4	Simplified schematic of starch granule initiation and synthesis in Arabidopsis
Figure 1.5	ChEC153 localises to a number of discrete puncta within chloroplasts
Figure 1.6	HaRxL94b localises to a number of discrete puncta within chloroplasts, as well as to the nucleus
Figure 2.1	Schematic of *Agrobacterium*-mediated transient expression in *N. benthamiana*
Figure 3.1	Schematic depicting particle bombardment of Arabidopsis leaves
Figure 3.2	Schematic to illustrate starch granule particle size analysis
Figure 3.3	ChEC153 and HaRxL94b co-localise with one another
Figure 3.4	Predicted structure of the mature ChEC153 protein
Figure 3.5	Predicted structure of the mature HaRxL94b protein
Figure 3.6	HaRxL94b and HaRxL94 have considerable sequence identity at the amino acid level
Figure 3.7	The N-terminus of HaRxL94b dictates its chloroplastic localisation
Figure 3.8	ChEC153 and HaRxL94b co-localise with MRC
Figure 3.9	ChEC153 and HaRxL94b co-localise with TaCSP41a
Figure 3.10	MRC co-localises with TaCSP41a
Figure 3.11	ChEC153 and HaRxL94b do not consistently co-localise with $PEND_{1-88}$
Figure 3.12	Localisation of ChEC153 and HaRxL94b alongside starch marker GBSS
Figure 3.13	ChEC153 induces the formation of GBSS-RFP puncta
Figure 3.14	ChEC153-eGFP is not detected in starch isolated from putative effector-expressing Arabidopsis rosettes
Figure 3.15	Localisations of ChEC153 and HaRxL94b in Arabidopsis lines Col-0, *mrc*, *ss4,* and *pgm*
Figure 3.16	Starch content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines
Figure 3.17	End of day total starch content of *N. benthamiana* leaf tissue transiently expressing *p19* only or *35S::ChEC153-eGFP + p19*

Figure 3.18	Soluble sugar content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines
Figure 3.19	Glucose content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines grown with a 12-hour photoperiod
Figure 3.20	End of day soluble sugar content of *N. benthamiana* leaf tissue transiently expressing *p19* only or *35S::ChEC153-eGFP + p19*
Figure 3.21	Size distributions showing the diameters of starch granules purified from Col-0, *ChEC153-eGFP*-expressing lines, and *mrc*
Figure 3.22	GBSS-RFP puncta diameter (µm) during co-expression of *35S::GBSS-RFP* and *35S::ChEC153-eGFP* in *N. benthamiana*
Figure 3.23	Scanning electron micrographs of starch purified from Col-0, *ChEC153-eGFP* expressing lines 2-5 and 2-9, and *mrc*
Figure 4.1	Schematic illustrating *C. higginsianum* infection assays
Figure 4.2	Schematic for Golden Gate cloning of fungal gene-knockout constructs
Figure 4.3	*35S::ChEC153-eGFP* Arabidopsis lines show increased susceptibility to *C. higginsianum*
Figure 4.4	Vectors used for generation of fungal putative effector-knockout strains
Figure 4.5	Schematic depicting the generation of *C. higginsianum ChEC153*-knockout strains
Figure 4.6	PCR-based verification of *ChEC153*-knockout strains from extracted genomic DNA
Figure 4.7	The *in vitro* growth rate of strain *ΔChEC153-2* is no different to that of the background strain *ΔChKu80,* but *ΔChEC153-1* shows a slower rate of growth
Figure 4.8	The sporulation rate of *ΔChEC153-2* is no different to that of the background strain *ΔChKu80*
Figure 4.9	*ΔChEC153* strains show no clear virulence phenotype *in planta* relative to *ΔChKu80* that cannot be accounted for by *in vitro* growth defects
Figure 4.10	Leaves of Arabidopsis lines expressing *ChEC153* do not have a clear macroscopic starch phenotype during infection compared to Col-0
Figure 4.11	Arabidopsis leaves infected with *ΔChEC153* strains have no clear macroscopic starch phenotype compared to *ΔChKu80* infection based on Lugol staining
Figure 4.12	*mrc* Arabidopsis mutants are more resistant to infection by *C. higginsianum* than Col-0
Figure 4.13	AtCPNB3 localises to the chloroplast
Figure 4.14	AtEIF2-A2 localises to the cytoplasm
Figure 4.15	AtCRBIC localises to the chloroplast and is recruited to puncta by ChEC153
Figure 4.16	CRBIC is recruited to chloroplastic puncta during *CRBIC* co-expression with *MRC* or *TaCSP41a*

Figure 4.17	CRBIC localises to mitochondria, in addition to chloroplasts
Figure 4.18	ChEC153 localises to mitochondria at a low level
Figure 4.19	Confirming T-DNA insertion in *CRBIC (AT1G53120)* in SALK_099429C plants
Figure 4.20	*crbic* Arabidopsis mutants do not have altered susceptibility to infection by *C. higginsianum* at four- or five-days post inoculation
Figure 4.21	Western blot analyses of *N. benthamiana* protein extracts do not show detection of full length ChEC153-eGFP
Figure 4.22	Schematic for theoretical KASP-based competition assay
Figure 5.1	Schematic to illustrate the process of 5′ RACE
Figure 5.2	Differences in gene models corresponding to *ChEC153* in different genome assemblies, and their encoded proteins
Figure 5.3	RT-PCR amplification of *ChEC6* and a portion of *ChEC153*
Figure 5.4	RT-PCR amplification of different portions of *ChEC153*
Figure 5.5	Sanger sequencing of cloned regions of gDNA and cDNA for *ChEC153*
Figure 5.6	5′ RACE did not produce clear amplicons
Figure 5.7	RT-PCR amplification of an upstream region of *ChEC153*
Figure 5.8	RT-PCR of *ChEC153.3*
Figure 5.9	Summary of *ChEC153* gene models
Figure 5.10	Proteins encoded by *ChEC153* gene models: alignments of sequences and predicted structures
Figure 5.11	Validation of RNA: reverse transcription quantitative PCR amplification of *AtUBQ10* from cDNA samples, confirming Arabidopsis gene expression
Figure 5.12	Validation of RNA: RT-PCR to amplify *ChEC6* from RNA samples, confirming *C. higginsianum* gene expression
Figure 5.13	Total normalised read counts for three biological replicates of Arabidopsis treated with *ΔChEC153-2*, *ΔChKu80*, or mock (water) at 24- or 36-hours post inoculation
Figure 5.14	Dispersion plots of RNA-Seq reads for the *C. higginsianum*-Arabidopsis pathosystem
Figure 5.15	PCA of RNA-Seq reads mapping to the Arabidopsis genome
Figure 5.16	PCA of RNA-Seq reads mapping to the *C. higginsianum* genome
Figure 5.17	RNA-Seq reads from *C. higginsianum* strain *ΔChEC111* map to *ChEC153.3* in the Tsushima *et al.* genome (ASM492035v1)

iv) List of Tables

Table 1.1:	Expression of *ChEC153* during different stages of infection
Table 1.2:	Expression of *HaRxL94b* during infection
Table 2.1:	Arabidopsis lines used
Table 2.2:	Concentrations of antibiotics used for bacterial selection
Table 2.3:	Composition of 10% polyacrylamide gels used for SDS-PAGE
Table 2.4:	Composition of buffers used for Western blotting
Table 3.1:	Subcellular localisation predictions for HaRxL94b and HaRxL94 with various N-terminal truncations using LOCALIZER, and experimentally determined localisations
Table 3.2:	Average starch granule diameters of starch purified from Arabidopsis lines Col-0, *ChEC153*-expressing lines, and *mrc*
Table 4.1:	Candidate interactors of ChEC153 identified by IP-MS
Table 5.1:	Prediction of signal peptides and effector prediction using *in silico* tools for the proteins encoded by the various *ChEC153* potential gene models
Table 5.2:	Differentially expressed Arabidopsis genes at 24- or 36-hours post inoculation with *ΔChEC153-2* relative to *ΔChKu80*
Table 5.3:	*C. higginsianum* genes differentially expressed between strains *ΔChEC153-2* and *ΔChKu80* at 24- or 36-hours post Arabidopsis inoculation

v) List of Appendices

Appendix 1: Lists of plasmids and primers used in this work.

A1 Tab. A	List of Level 0 Golden Gate parts used in this work
A1 Tab. B	List of Golden Gate acceptor plasmids used in this work
A1 Tab. C	List of Level 1 Golden Gate plasmids used in this work
A1 Tab. D	List of Gateway plasmids used in this work
A1 Tab. E	List of primers used in this work
A1 Tab. F	List of gBlocks synthesised during this work

Appendix 2: Supplementary infection assay data and statistical analyses

A2 Fig. A	*35S::ChEC153-eGFP* lines show no difference in lesion size relative to Col-0 following infection with *C. higginsianum* at six days post inoculation
A2 Tab. A	Summary of statistical test results for *C. higginsianum* infection assay comparing Col-0 and *35S::ChEC153-eGFP* Arabidopsis lines
A2 Tab. B	Summary of statistical test results for *C. higginsianum* strain growth assay
A2 Fig. B	*ΔChEC153* knockout strains show no clear virulence phenotype *in planta* relative to *ΔChKu80* that cannot be accounted for by *in vitro* growth defects
A2 Tab. C	Summary of statistical test results for *C. higginsianum* infection assay comparing *ΔChKu80* and *ΔChEC153* strains
A2 Fig. C	*mrc* Arabidopsis mutants are more resistant to *C. higginsianum* infection than Col-0 at 4 dpi
A2 Tab. D	Summary of statistical test results for *C. higginsianum* infection assay comparing *mrc* and Col-0 susceptibility
A2 Fig. D	*crbic* Arabidopsis mutants do not have altered susceptibility to infection by *C. higginsianum* at 4 or 5 dpi
A2 Tab. E	Summary of statistical test results for *C. higginsianum* infection assay comparing *crbic* and Col-0 susceptibility

Appendix 3: Supplementary RNA-Seq analysis

A3 Fig. A	Total read counts (without normalisation) mapping to the *C. higginsianum* genome
A3 Tab. A	Differentially expressed Arabidopsis genes at 24- or 36-hours post inoculation with *ΔChKu80* relative to mock-inoculation

vi) ABBREVIATIONS

ACN	Acetonitrile
ADPG	ADP-glucose
AGPase	ADP-glucose pyrophosphorylase
APS	Ammonium persulfate
ATG8	AUTOPHAGY-RELATED GENE
ATMT	*Agrobacterium tumefaciens*-mediated transformation
ATR	ARABIDOPSIS THALIANA RECOGNISED
BAK1	BRI1-ASSOCIATED KINASE 1
BFDR	Bayesian false discovery rate
BH	Biotrophic hyphae
BIC	Biotrophic interfacial complex
BIK1	BOTRYTIS-INDUCED KINASE 1
BGC	Biosynthetic gene cluster
BP	Biotrophic phase
CAS	Chloroplast-localised Ca^{2+}-sensing
CAZyme	Carbohydrate-active enzyme
CBM	Carbohydrate-binding module
CCR	Carbon catabolite repression
CDS	Coding sequence
CERK1	CHITIN ELICITOR RECEPTOR KINASE 1
Ch	*Colletotrichum higginsianum*
ChEC153	*C. higginsianum* effector candidate 153
CHLORAD	Chloroplast-associated protein degradation
CME	Clathrin-mediated endocytosis
CNL	Coiled-coil NLR
CoxIV	Cytochrome c oxidase complex IV
CPK	Calcium-dependent protein kinase
CPN60	CHAPERONIN-60
CPNB3	CHAPERONIN-60BETA3
CPO	COPROPORPHYRINOGEN III OXIDASE
CRBIC	CHLOROPLAST RNA-BINDING INTERACTOR OF ChEC153
CRN	Crinkler
CSP41	CHLOROPLAST STEM-LOOP BINDING PROTEIN OF 41 KDA
CTP	CHLOROPLAST-TARGETED PROTEIN
CTE	CHLOROPLAST TARGETING EFFECTOR
cTP	Chloroplast transit peptide
DAMP	Damage-associated molecular pattern
DAPI	4′,6-diamidino-2-phenylindole
DBE	De-branching enzyme
dpi	Days post inoculation
EHM	Extrahaustorial membrane
EHMX	Extrahaustorial matrix
EIF2-A2	EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT

EIHM	Extra-invasive hyphal membranes
ESTs	Expressed sequence tags
ET	Ethylene
ETI	Effector-triggered immunity
ETS	Effector triggered susceptibility
EV	Extracellular vesicle
F6P	Fructose-6-phosphate
FLN	FRUCTOKINASE-LIKE
FLS2	FLAGELLIN SENSITIVE 2
FSD	IRON SUPEROXIDE DISMUTASE
fv	Fungal vector
G3P	Glycerol-3-phosphate
GBSS	GRANULE BOUND STARCH SYNTHASE
GCN	GENERAL CONTROL NON-REPRESSED
GFP	Green Fluorescent Protein
GI	Glycerol induction
GS	Gene-specific
GSP	Gene-specific primer
HA	Hemagglutinin
HaRxL94b	*H. arabidopsidis* RxLR effector candidate 94b
HLB	Huanglongbing
HopI1	HRP OUTER PROTEIN I1
Hpa	*Hyaloperonospora arabidopsidis*
HPH	HYGROMYCIN B PHOSPHOTRANSFERASE
hpi	Hours post inoculation
HR	Homologous recombination
HygR	Hygromycin resistance gene (*HPH*)
HXT	HEXOSE TRANSPORTER
ICS	Isochorismate
IGV	Integrative Genomics Viewer
IP	Immunoprecipitation
IP-MS	Immunoprecipitation-mass spectrometry
ISE2	INCREASED SIZE EXCLUSION LIMIT 2
ISR	Induced systemic resistance
ITL	INTEGRIN-LIKE
JA	Jasmonic acid
KASP	Competitive allele-specific PCR
KO	Knockout
LB	Lysogeny Broth
MAPK	Mitogen-activated protein kinase
MAX	*M. oryzae* avirulence and ToxB like
MFP1	MAR BINDING FILAMENT-LIKE PROTEIN 1
MOS	Maltooligosaccharides
MRC	MYOSIN-RESEMBLING CHLOROPLAST PROTEIN
MS	Murashige and Skoog
mTERF	MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR
NECG	Nuclear-encoded chloroplast gene

NHEJ	Non-homologous end joining
NLR	Nucleotide-binding leucine-rich repeat receptor
NP	Necrotrophic phase
$OD_{600\ nm}$	Optical density measured at 600_{nm}
PA	*in planta* appressoria
PAGE	Polyacrylamide gel electrophoresis
PAL	Phenylalanine ammonia-lyase
PAMP	Pathogen-associated molecular pattern
PAP	PEP-associated protein
PARC6	PARALOG OF ARC6
PBL2	PBS1-LIKE PROTEIN 2
PCA	Principal component analysis
PDA	Potato dextrose agar
PDB	Potato dextrose broth
PEND	PLASTID ENVELOPE DNA-BINDING PROTEIN
PEP	Plastid-encoded RNA polymerase
PGI	Phosphoglucose isomerase
PGM	Phosphoglucomutase
PMA	PLASMA MEMBRANE H^+-ATPASE
PRR	Pattern recognition receptor
PSR	PHYTOPHTHORA SUPPRESSOR OF RNA SILENCING
Pst	*Puccina striiformis* f. sp. *tritici*
pTAC	PLASTID TRANSCRIPTIONALLY ACTIVE CHROMOSOME
PTI	Pattern-triggered immunity
PTST	PROTEIN TARGETING TO STARCH
RACE	Rapid amplification of cDNA ends
RBOH	RESPIRATORY BURST OXIDASE HOMOLOG
RFP	Red Fluorescent Protein
RKS1	RESISTANCE-RELATED KINASE 1
RLCK	Receptor-like cytoplasmic kinase
RLK	Receptor-like kinase
RLP	Receptor-like proteins
RNA-Seq	RNA-sequencing
ROS	Reactive oxygen species
RPH1	RESISTANCE TO PHYTOPHTHORA 1
RPP	RECOGNITION OF PERONOSPORA
RT-PCR	Reverse transcription polymerase chain reaction
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SA	Salicylic acid
SBE	Starch branching enzyme
SDI1	SUCCINATE DEHYDROGENASE 1
SDS	Sodium dodecyl sulphate
SEX	STARCH EXCESS
SID	SALICYLIC ACID INDUCTION DEFICIENT
SM	Secondary metabolite
SNP	Single nucleotide polymorphism
SP	Signal peptide

sRNA	Small non-coding RNA
SS	Starch synthase
SWEET	SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTER
TAIR	The Arabidopsis Information Resource
TALE	Transcription Activator-Like Effector
T-DNA	Transfer DNA
TEAB	Triethylammonium bicarbonate buffer
TEMED	Tetramethylethylenediamine
TF	Transcription factor
TIC	TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS
TIR	Toll/interleukin-1 receptor
TNL	Toll/interleukin-1 receptor (TIR) NLRs
TOC	TRANSLOCON AT THE OUTER ENVELOPE MEMBRANE OF CHLOROPLASTS
TOR	Target of rapamycin
TPS1	TREHALOSE-6-PHOSPHATE SYNTHASE 1
UTR	Untranslated region
VA	*in vitro* appressoria
WT	Wild-type
ZAR1	HOPZ-ACTIVATED RESISTANCE 1

vii) ACKNOWLEDGEMENTS

First and foremost, I am indebted to my supervisors: Christine and David. As my primary supervisor, Christine supported me immensely and I have learnt so much from her as a scientist and as a person. I am grateful for her patience, understanding and guidance. I have also greatly valued David's approachability and enthusiasm throughout my project. To me, the kindness and support shown by both of my supervisors was as valuable as their extensive expertise and scientific rigour. Similarly, I am grateful to Claire for her mentorship, particularly in my Rotation year – for the sound advice and kind encouragement that I needed to find my feet.

I am very grateful to all the members, past and present, of both the Faulkner and Seung labs for their tuition, advice, feedback, encouragement, and friendship. In particular, this project would not have been possible without preliminary work carried out by Mina, Jo, Xiaokun, and Andy. To Jo and Mina for their help early in my project as I settled into the lab. I am further grateful to Andy for his role on my supervisory committee, and his experimental assistance particularly in my final year. Moreover, for his reassuring and positive attitude, and the continual support he has shown me and gives the entire Faulkner lab. Hannah Rae's input on my gene model troubleshooting, her RNA-Seq analysis help, and her humour, were also essential for my final results chapter.

I am grateful to Jorunn and Jake for constructive discussion and examination of this thesis.

I also owe thanks to Mark and Diane for allowing me to be part of their lab groups for my rotations, and for teaching me a lot about plant-pathogen interactions from different perspectives. I am grateful for the members of both of their labs: particularly Juan Carlos, Adam, and Pilar for their day-to-day supervision through my rotation projects.

To all of the technical and support staff at the John Innes Centre, particularly horticultural services, the proteomics platform, the bioimaging department, and the lab support teams. Further, to Mark Youles and TSL SynBio. Particular thanks to Roy.

To Volkan for inspiring me.

To Karen for her help and perspective.

To the John Innes Foundation for funding this work.

Many thanks to my Rotation cohort: Andy, Anson, Josh, and Sam - especially to Sam for the lunches, chats, and empathy. Also, to Andy, Basti, Jo, Surbhi, and Joe.

To Jiawen, Lara, Rose, Thomas, and Qi Yang,

Hannah Rae and Leo,

Emma and Emma,

Estee and Chris.

Jamie and Daniel, Matt, and Louise and Fred, for their friendship throughout.

To Anna and Ryanne. No words can do justice but perhaps these will come closer than any others: *Nunu tus immeasurable amold yee kindness! Ka ern vus mupa mucha, yee ka sama adtut be jignow grateful nunu mi obe yambrr.*

Finally, to my family: to Hannah and Daisy, and especially to my parents.

DEDICATED TO

Sonia Le Feuvre & Pat Jones

Chapter 1

Introduction

1.1 Plant-pathogen interactions

1.1.1 An introduction to phytopathogens

Plants are exposed to pathogenic attack by manifold microorganisms including fungi, oomycetes, and bacteria, as well as viruses and pests. As such, phytopathogens represent an ongoing threat to global food security. Large-scale plant disease epidemics, such as the 1840s outbreak of late potato blight caused by *Phytophthora infestans* and resulting in the Irish potato famine, highlight the clear capacity of phytopathogens to have devastating consequences, and their potential to be exacerbated by socio-political issues (Turner, 2005). It is estimated that annually pathogens and pests result in 20-30% yield losses for crops globally (Savary et al., 2019). Despite this, plants are generally resistant to the majority of pathogens to which they are exposed.

As the world's population continues to grow increasing the demand on crop production, climate change expedites emergence of pathogens causing debilitating crop losses (Delgado-Baquerizo et al., 2020) particularly in the cases of fungal and oomycete pathogens, as reviewed by Fones *et al.* (Fones et al., 2020). Further, a global effort to support food and nutrient access may require improved, equitable distribution of food, itself presenting a risk in terms of disease transmission. Advances in pathogen surveillance (Salcedo et al., 2021, Buja et al., 2021), identification of immunity-related host genes as targets for breeding (Deng et al., 2020), and genetic engineering (Dong and Ronald, 2019) offer hope in tackling this issue as we aim to move away from excess application of agrochemicals to control disease. It is therefore vital that we increase our understanding of plant pathogens, pathogenesis and host immunity as part of efforts to instigate and sustainably maintain true food security.

Pathogens are often classified by their mode of nutrient acquisition and lifestyle: as biotrophic, hemibiotrophic or necrotrophic. Biotrophic pathogens derive their energy from live host cells, requiring them to suppress the host defence responses while maintaining host viability. Some biotrophs colonise the host apoplast, from which they gain nutrients (Stotz et al., 2014), while others form intracellular biotrophic interactions with the host. In such parasitic relationships, specialised pathogen feeding structures form close association with host cells, and the pathogen may form a host nutrient sink at the site of infection, disadvantaging the plant and allowing long-term infection during which the pathogen can complete its lifecycle. Hemi-biotrophy refers to a lifestyle in which pathogens that initially form biotrophic interactions with their hosts, including the formation of intracellular biotrophic hyphae, later switch to a destructive, necrotrophic stage. Necrotrophic pathogens obtain energy from dead and dying host cells – which they themselves kill. In some cases, this may be preceded by a cryptic biotrophic-like phase, or quiescent phase in the case of necrotrophic pathogens causing post-harvest diseases (highlighted in Rajarammohan,

2021). These diverse lifestyles necessitate host plants to mount specific, appropriate defence responses to different invading pathogens. For example, a hypersensitive response, during which the host triggers localised cell death, may halt the spread of a biotrophic pathogen, but do little to impede infection by a necrotroph (Glazebrook, 2005). Exemplifying the difference in immune responses to pathogens of different lifestyles, hormone responses typifying defence against biotrophs generally involve salicylic acid (SA) while the jasmonic acid (JA)/ethylene (ET) pathway is generally thought to be more important in responses to necrotrophic pathogens (Glazebrook, 2005). Pathogens also alter host primary metabolism even in incompatible interactions where disease symptoms are not apparent (Duan et al., 2013).

1.1.2 The plant immune response is multi-layered and complex

Our knowledge of the mechanisms and complexities underpinning plant immunity has developed drastically since the gene-for-gene hypothesis was presented by Harold Flor, who observed a requirement for a host resistance gene and cognate pathogen "avirulence" gene for disease resistance (Flor, 1971). How plants mount an immune response to stave off infection and how pathogens seek to evade host immunity are concepts of extensive research, interlinked by the evolutionary drivers on either side. Strong selective pressures on both host and pathogen lead to rapid diversification of both plant defence and pathogen virulence factors.

The execution of an immune response is energetically demanding, leading to a trade-off between plant defence and growth in terms of resource allocation (Huot et al., 2014). To maximise fitness, the plant innate immune response is consequently inducible rather than constitutive, requiring host molecular surveillance mechanisms to ensure defence responses are initiated appropriately. Plant innate immunity encompasses localised and systemic, and transient and prolonged, responses to pathogen perception.

1.1.2.1 Pattern-triggered immunity - PTI

For the plant to initiate an immune response the presence of a pathogen must first be recognised. At the plant cell surface, pattern recognition receptors (PRRs) detect conserved molecular signatures in the apoplast. These signatures include pathogen-associated molecular patterns (PAMPs), as well as self-molecules resulting from physical damage of the host such as those derived by the degradative action of pathogen enzymes, referred to as damage-associated molecular patterns (DAMPs).

PRRs may be either receptor-like kinases (RLKs) or receptor-like proteins (RLPs), and comprise of extracellular ligand-binding domains, transmembrane domains, and either intracellular kinase signalling domains (in the case of RLKs) or short intracellular domains (in the case of RLPs) which

upon activation may interact with other kinases to enable signal transduction. In general terms, upon cell-surface perception of their cognate PAMP ligands, PRRs associate with co-receptors (such as BRI1-ASSOCIATED KINASE 1, BAK1) to form receptor complexes which launch a largely conserved series of defence responses. Receptor complex formation leads to PRR phosphorylation among numerous other trans- and auto-phosphorylation events. Receptor-like cytoplasmic kinases (RLCKs, such as BOTRYTIS-INDUCED KINASE 1, BIK1) are phosphorylated by activated receptor complexes, triggering various downstream signalling pathways via phosphorylation. Activation of calcium channels leads to Ca^{2+} influx into the cell, which in turn activates calcium-dependent protein kinases (CPKs; Negi et al., 2023). One of the early responses to pathogen-perception is the rapid production of reactive oxygen species (ROS) in the apoplast, a process mediated by RESPIRATORY BURST OXIDASE HOMOLOGS (RBOHs), particularly RBOHD (Torres et al., 2002). This oxidative burst may be critical in inducing cross-linking of structural cell wall proteins to increase the physical strength of the cell wall as a barrier to infection (Bradley et al., 1992), and even in directly damaging the pathogen itself (Lamb and Dixon, 1997). Further, downstream mitogen-activated protein kinase (MAPK) signalling pathways are induced by PAMP detection, and transcription is reprogrammed to upregulate defence-related host genes and enhance production of defence hormones such as SA, JA, and ET (Lal et al., 2018, Macho and Zipfel, 2014) and antimicrobial phytoalexins (Umemoto et al., 1997). Additionally, PAMP recognition is associated with deposition of callose at the cell wall, and plasmodesmal as well as stomatal closure (Faulkner et al., 2013, Melotto et al., 2006). As an example, an array of bacterial PAMPs exist and are perceived in turn by a range of host PRRs, initiating immune responses to bacteria. One of the best-characterised of these is the perception of the bacterial flagellin-derived peptide flg22 by the *Arabidopsis thaliana* (hereafter: Arabidopsis) PRR FLAGELLIN SENSITIVE 2 (FLS2; Gomez-Gomez and Boller, 2000, Chinchilla et al., 2006), and the induced interaction of FLS2 with the BAK1 co-receptor which triggers FLS2 endocytosis (Robatzek et al., 2006) and downstream defence signalling.

In this way, host detection of PAMPs leads to pattern-triggered immunity (PTI). Molecules detected as PAMPs are often derived from key facets of a pathogen's physicality, and may be critical for pathogen fitness, making mutation of these features to avoid detection impracticable. Hence, in order to evade or suppress plant surveillance, pathogens have adapted methods to interfere with the perception of PAMPs through the action of effectors.

1.1.2.2 Effector-triggered susceptibility – ETS

Pathogens produce pathogenicity-related factors, termed effectors, to suppress the host immune response and support their own virulence in a process of immune evasion called effector triggered susceptibility (ETS). Effectors may act to sequester or alter the pathogen's own PAMPs such that

PTI can be evaded. For example, the perception of the fungal PAMP chitin is dampened by the action of a *Cladosporium fulvum* effector, EXTRACELLULAR PROTEIN 6, that appears to compete with the host PRR in binding chitin (de Jonge et al., 2010). Similarly, effectors are implicated in targeting and blocking PRRs and their associated kinases, an example of which is the interference of conserved fungal effector NECROSIS-INDUCING SECRETED PROTEIN 1 in BAK1 and BIK1 signalling (Irieda et al., 2019).

Aside from suppressing cell-surface immune recognition, other effectors target core host processes and host gene expression, manipulating the plant to support pathogen colonisation and virulence. Some nuclear-localised effectors are able to modulate host gene expression by manipulating chromatin configuration (Kong et al., 2017, Arbibe et al., 2007), by directly or indirectly targeting host transcription factors (TFs) to perturb endogenous gene regulation (Qi et al., 2019, Zhang et al., 2021a, Tanaka et al., 2014), or by mimicking TFs and directly binding host gene promoters. In *Xanthomonas* spp., Transcription Activator-Like Effectors (TALEs) exemplify the capacity for pathogens to manipulate the host via effector proteins by transcriptional reprogramming, inducing expression of host susceptibility genes to promote successful colonisation (Romer et al., 2010). In some pathosystems plant defence genes are seen to be directly upregulated by effectors, suggesting host evolution to exploit and nullify the action of TALEs (Romer et al., 2007). The downstream impacts of nuclear gene targeting by pathogen effectors can be wide-ranging, including but not limited to manipulating host primary metabolism to increase availability of nutrients to the pathogen (Xing et al., 2021, Gupta et al., 2021), the perturbation of hormone signalling (Gimenez-Ibanez et al., 2014, Jiang et al., 2013, Caillaud et al., 2013) and mis-regulation of host secondary metabolite production (Tanaka et al., 2014).

1.1.2.3 Effector-triggered immunity – ETI

The effectors secreted by pathogens can be recognised directly or indirectly by host intracellular nucleotide-binding leucine-rich repeat receptors (NLRs), activating a second form of defence: effector-triggered immunity (ETI). NLRs typically comprise a variable N-terminal domain, a conserved central nucleotide-binding and oligomerisation domain, and a C-terminal LRR domain. NLRs are frequently categorised based on the variable N-terminal domain, for example into coiled-coil NLRs (CNLs) and Toll/interleukin-1 receptor (TIR) NLRs (TNLs). Some NLRs also contain integrated domains which may act as decoys by mimicking effector targets (Kroj et al., 2016).

In their role perceiving effector proteins, NLRs may act alone (as singleton NLRs), in pairs (paired NLRs), and as networks (Adachi et al., 2019). Singleton NLRs possess the ability to both recognise effectors and to activate the required immune response. In contrast, paired NLRs comprise a sensor NLR for effector recognition and a helper NLR for immune signalling. Activation of singleton

NLRs is associated with oligomerisation, forming NLR protein complexes termed resistosomes (Wang et al., 2019, Forderer et al., 2022). One example of this is the formation of the ZAR1 resistosome (Wang et al., 2019). The Arabidopsis HOPZ-ACTIVATED RESISTANCE 1 (ZAR1) is a CNL involved in ETI in perception of a number of bacterial effectors in combination with different RLCKs. Prior to activation, ZAR1 exists in a complex with the RLCK RESISTANCE-RELATED KINASE 1 (RKS1). The ZAR1 resistosome complex that forms upon recognition of the effector AvrAC incorporates an AvrAC-activated BIK1 paralog, PBS1-LIKE PROTEIN 2 (PBL2) kinase, $PBL2^{UMP}$ (Wang et al., 2015). It is this ZAR1-RKS1-$PBL2^{UMP}$ complex that pentamerises to form the funnel-shaped ZAR1 resistosome, undergoing large conformational changes during activation. The activated ZAR1 resistosome is associated with plasma membranes and forms a Ca^{2+}-permeable channel, leading to Ca^{2+} influx and other aspects of immune signalling (Bi et al., 2021). In contrast to singleton NLRs, our understanding of how paired NLRs are activated and carry out signalling at a molecular level is relatively limited, but may involve effector-activated sensor NLRs triggering their cognate helper NLRs to form oligomeric resistosomes from which the sensor NLR is excluded (Ahn et al., 2023, Contreras et al., 2023).

The downstream processes initiated in ETI are similar to those seen in PTI, with Ca^{2+} influx, ROS accumulation, transcriptional reprogramming for induction of defence genes, and production of phytohormones. In ETI, these responses often culminate in localised death of a small number of host cells at the infection site (hypersensitive response), preventing further colonisation of healthy host tissues by the pathogen. There is apparent cross-talk between PTI and ETI, with neither form of immunity being sufficient to provide resistance to *Pseudomonas syringae* in isolation (Ngou et al., 2021).

Responses to infection are not limited to localised PTI and ETI, as signalling cascades can also trigger a systemic immune response known as systemic acquired resistance. Among the proposed signals mediating systemic acquired resistance are SA, SA-derivatives, and ROS (Gao et al., 2015). Further, induced systemic resistance (ISR) is a form of enhanced resistance initiated by the recognition of non-pathogenic microorganisms and involves JA and ET responses (Choudhary et al., 2007). Induced resistance responses can be long-lasting and prime plants to respond to future infection. In these ways, successful immune responses, sometimes amplified by priming and induced resistance, lead to host disease resistance.

1.1.3 Pathogen effectors

As detailed in **Section 1.1.2.2**, effectors can target various host processes to suppress immune responses and create favourable conditions for pathogen virulence. The adaption and evolution of the effector repertoire of a pathogen is critical to its viability. Effectors that are recognised by the

host immune system and trigger ETI may be deleted, silenced, or else mutated such that they are no longer perceived – into so-called "stealthy" effectors which maintain their functional activity. The effector repertoire may also evolve to target a new host protein or allow infection of a new host.

The mechanisms of effector secretion employed by filamentous pathogens are not yet well understood. Effector candidates can be predicted based on a number of criteria. From genomic data, prediction of signal peptides targeting proteins for secretion from the pathogen and similarity to known effector proteins facilitate the identification of candidate effector proteins. However, limitations exist in our understanding of effector secretion, and some characterised secreted effector proteins of filamentous pathogens appear to lack canonical signal peptides (Liu et al., 2014). Transcriptomic data are therefore powerful in complementing these predictions, with the identification of genes expressed preferentially at specific infection stages pointing to putative roles in virulence. Pathogens demonstrate temporal regulation of effector expression through the course of the infection process. Indeed, in hemibiotrophs, effector expression may be important in mediating the transition from biotrophy to necrotrophy. Effectors of a given hemibiotroph may act antagonistically, for example the SUPPRESSOR OF NECROSIS 1 and NECROSIS-INDUCING PHYTOPHTHORA PROTEIN 1.1 effectors of *P. infestans* which respectively suppress and induce necrosis and are produced in the biotrophic and necrotrophic phases of infection (Lee and Rose, 2010, Kelley et al., 2010, Kanneganti et al., 2006). Thus, from genomic and transcriptomic data, suites of putative effectors can be identified for a given pathogen for further validation. Alongside these methods to discern candidate effectors, identification of loss-of-virulence phenotypes in forward genetics experiments can also reveal genes associated with pathogen effectors (Plaumann et al., 2018).

Effectors are categorised generally, based on the site of their action, as being apoplastic or cytoplasmic. While some apoplastic effectors may be recognised by the plant analogously to PAMPs, others may function in PAMP and PAMP-derived molecule scavenging to dampen PRR recognition (de Jonge et al., 2010), targeting of plant-derived hydrolytic enzymes to prevent PAMP production (Marshall et al., 2011), and direct inhibition of PRRs to suppress PTI (Xiang et al., 2008). Pathogens, especially those associated with necrotrophy, may employ apoplastic effectors to degrade or weaken cell walls to facilitate invasion. Another mechanism by which effectors are implicated in manipulating host physical defences is at stomata, which for some pathogens present a key point of entry into plants. For example, effectors have been characterised that oppose the PAMP-triggered response of stomatal closure, seeking to maintain access to the host by preventing stomatal closure or inducing stomatal opening (Zhou et al., 2015, Raffeiner et al., 2022).

Cytoplasmic effectors have been seen to localise to an array of subcellular compartments in plants, reflective of the diversity of their host targets. Further, some effector proteins appear able to mis-localise the endogenous host proteins that they target, perturbing their function in defence (Zhang et al., 2015, Qi et al., 2019). Characterisation of effector functions can highlight critical immune processes. However, studying the role of individual effectors can often be compounded by functional redundancy and small but additive impacts of effectors on pathogen virulence. In this way, deletion of a single effector gene from the pathogen may be insufficient to measurably impede infection. Further, some pathogens appear to have evolved decoy effectors to subvert ETI. Such effectors, while they may lack their own functional activity, are able to support infection by diminishing the capacity of host immune proteins to target the non-decoy forms of effectors (Ma et al., 2017).

1.2 Fungal plant pathogens

1.2.1 Fungi as phytopathogens

Phytopathogenic fungi represent a dominant cause of plant diseases (Dean et al., 2012). Reflective of their own diversity, fungi use a range of strategies to colonise their hosts. The kingdom comprises biotrophs including numerous smuts such as *U. maydis* and rusts such as *Puccinia* spp., hemibiotrophs like *M. oryzae*, and necrotrophs such as *B. cinerea*. Filamentous fungi develop vegetative multicellular structures including hyphae, which facilitate colonisation by penetrating plant cell walls and allowing the pathogen to rapidly spread through infected tissues.

A number of fungal PAMPs are associated with initiation of PTI in response to fungal attack. One such fungal PAMP, chitin, is an essential and structural component of fungal cell walls composed of β(1,4)-linked *N*-acetylglucosamine units. In Arabidopsis, chitin is perceived by PRRs LYSM-CONTAINING RECEPTOR-LIKE KINASE 4 and 5 and CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1) (Cao et al., 2014, Miya et al., 2007). Phosphoactivated CERK1 activates the RLCK PBS1-LIKE 27 to trigger downstream PTI responses (Yamada et al., 2016, Shinya et al., 2014). Plasmodesmal closure in response to chitin, which limits symplastic connectivity as part of the defence response, is reliant on an independent PRR, LYSM-CONTAINING GPI-ANCHORED PROTEIN 2 which is enriched at plasmodesmata (Faulkner et al., 2013).

1.2.2 Fungal effectors

The number of specific fungal effectors that have been characterised is limited. Fungal phytopathogens may employ effectors to evade PTI. Effectors associated with fungal necrotrophy

are often characterised as inducing cell death, including a capacity to subvert ETI, harnessing the plant's ability to trigger the hypersensitive response for the gain of the pathogen (Shao et al., 2021).

The mechanisms by which fungal effector proteins translocate from pathogen to host are not well understood. For effectors of biotrophic or hemibiotrophic fungal pathogens to reach the cytoplasm, they must exit the pathogen appressoria, haustoria or invasive hyphae and translocate across plant-derived membranes which separate these specialised infection structures from the host cell. Typically, phytopathogenic filamentous fungal effectors are believed to have N-terminal sequences for export from the fungus and import into the plant in series, with diverse C-termini specialised for their functions within the plant (Petre and Kamoun, 2014). As consensus cell-entry motifs have not been widely characterised for fungi, prediction of fungal effectors from genomic resources represents a challenge, and largely relies on the prediction of N-terminal signal peptides for secretion and an absence of predicted transmembrane domains. One of the few fungal effector motifs to be identified is the Y/F/WxC motif seen to be enriched at the N-terminus of powdery mildew fungus *B. graminis* effectors and some rust fungal effectors (Godfrey et al., 2010). Further, a highly degenerate RxLR-like N-terminal motif has been described in some fungal effectors (Kale et al., 2010), somewhat analogous to the RxLR motif more commonly observed in oomycete effectors. A number of *M. oryzae* effectors lacking sequence homology share structural features and are classified together as *M. oryzae* avirulence and ToxB like (MAX) effectors. MAX effectors appear to be especially important in the biotrophic phase of infection by *M. oryzae*, where they represent a large portion of the known effector repertoire (de Guillen et al., 2015). Advances such as the development and high-throughput application of *in silico* protein structure prediction tools such as AlphaFold (Jumper et al., 2021) may reveal further structural features common to fungal effectors and aid in identifying additional effector families and candidates. Generally, fungal effectors are often small and cysteine-rich, criteria which are also employed for effector prediction (Sperschneider et al., 2015a), but which may preclude identification of less conventional effectors. Therefore, in the identification of fungal effectors, which for many species lack conserved motifs and effector superfamilies, multi-omics approaches are particularly powerful, combining genomic, transcriptomic, and/or proteomic data (Gonzalez-Fernandez and Jorrin-Novo, 2012).

In *M. oryzae*, two pathways for effector secretion are present. Apoplastic effectors are released to the extracellular space by the conventional ER-Golgi secretory pathway (Giraldo et al., 2013). In contrast, there is increasing evidence that cytoplasmic effectors are secreted by a non-conventional pathway associated with extracellular vesicles (EVs). These effectors accumulate at the pathogen biotrophic interfacial complex (BIC), a specialised structure that develops at the tip of invasive hyphae, from which they appear to translocate into the host cytoplasm (Khang et al., 2010). Their secretion appears to occur as exocytosis in an exocyst component dependent manner

(Giraldo et al., 2013). Recently, plant clathrin-mediated endocytosis (CME) has been implicated in the translocation of cytoplasmic *M. oryzae* effectors from the fungal BIC (Oliveira-Garcia et al., 2023). Rather than being a fungal-specific mechanism of effector uptake, this may reflect a conserved effector translocation route for filamentous pathogens (Wang et al., 2023a), with similar observations in the uptake of oomycete *P. infestans* RxLR effectors (Wang et al., 2023b). In *U. maydis* a potential secretion system for effector proteins was recently described in which five effector proteins and two membrane proteins form a protein complex in the fungal membrane which is seen to extend into plant cells and is required for virulence. This complex may facilitate the secretion of effectors or effector-containing vesicles from pathogen to host (Ludwig et al., 2021). Thus, from our limited understanding of fungal effector protein secretion, it is possible that proteins hitherto disregarded when identifying putative effectors based on a lack of signal peptide prediction may function as effector proteins with non-typical methods of secretion/translocation.

While much of effector research focuses on proteinaceous effectors, the effector molecules produced by phytopathogens during infection are not limited to secreted proteins. Non-proteinaceous effectors associated with phytopathogenic fungi include small non-coding RNA (sRNA) effectors and secondary metabolite (SM) chemical effectors. sRNA effectors can function to suppress host defence gene expression. Exemplifying this, Weiberg *et al*. demonstrated that *B. cinerea* sRNAs can be transferred to Arabidopsis or tomato host cells and silence immunity-related genes by hijacking the hosts' endogenous RNAi machinery (Weiberg et al., 2013). Since, similar sRNA effectors have been observed in other phytopathogenic fungi, including *Puccinia striiformis* f. sp. *tritici* (Wang et al., 2017). Transfer of sRNAs is bidirectional, not exclusively in the direction of pathogen to host; several plant sRNAs have been suggested to traffic into fungal pathogens via extracellular vesicles in order to suppress pathogen virulence genes (Cai et al., 2018, Zhang et al., 2016). Fungi produce a variety of SMs, some which may function as effectors in infection. SMs are typically encoded in biosynthetic gene clusters, which can be expressed selectively at specific stages of infection (Dallery et al., 2017). While characterisation of specific fungal secondary metabolite effectors is limited, they have the capacity to target a number of host processes such as protein synthesis and trafficking, and calmodulin signalling (Collemare et al., 2019). Further, fungi are capable of producing, mimicking or modifying phytohormones, highlighting numerous ways that fungal secondary metabolites or proteinaceous effectors may mitigate host defence responses (Patkar et al., 2015, Chanclud et al., 2016). Some fungal secondary metabolites function to ward off pests and herbivores, and others have anti-bacterial or anti-fungal capacities indicative of the competition between various phytopathogens to colonise the host. During necrotrophy, fungal pathogens may secrete secondary metabolite phytotoxins.

1.3 *Colletotrichum higginsianum*

A hemibiotrophic, filamentous fungus of the phylum Ascomycota, *C. higginsianum (Ch)* is the causative agent of anthracnose disease on numerous cruciferous plants. The *Colletotrichum* genus comprises around 250 species, which are principally grouped into 14 species complexes. The majority of *Colletotrichum* spp., including *C. higginsianum*, are phytopathogenic but other species are endophytic, saprophytic or even entomopathogenic (Jayawardena et al., 2016, Marcelino et al., 2009). *Colletotrichum* spp. are able to infect staple food crops including cassava, sugar cane, bananas, maize and sorghum, fruits and vegetables, and ornamentals (as reviewed by Cannon et al., 2012). Emphasizing their impact, *Colletotrichum* spp. were ranked within the top ten fungal pathogens in terms of their scientific/economic importance in 2012 (Dean et al., 2012). *C. higginsianum* itself, part of the destructivum species complex, infects a range of *Brassicaceae* crops including *Brassicas* and *Raphani* (Damm et al., 2014). Additionally, *C. higginsianum* can infect many ecotypes the model plant Arabidopsis (O'Connell et al., 2004) aiding its study and providing a model pathosystem from which to investigate plant-fungal pathogen interactions. While Col-0 is susceptible to infection by *C. higginsianum*, some Arabidopsis ecotypes have resistance to this pathogen (Narusaka et al., 2009). In contrast, model *Nicotiana* species appear resistant to *C. higginsianum* infection (O'Connell et al., 2004).

1.3.1 *C. higginsianum* lifecycle and infection strategy

The infection process of *C. higginsianum* is characterised as hemibiotrophic: in which an initial phase of biotrophy precedes necrotrophy (O'Connell et al., 2004). Some of our understanding of *C. higginsianum* infection has been inferred from knowledge of other, similar, *Colletotrichum* species or *Magnaporthe oryzae* – the well-studied, hemibiotrophic, filamentous fungus that causes rice blast disease. While *C. higginsianum* can infect leaves, stems and fruits of some of its hosts, most research centres on the infection process in leaves and this will form the focus here. The *C. higginsianum* infection process is illustrated in **Figure 1.1**.

Infection is initiated as the unicellular, stadium-shaped conidia of *C. higginsianum* attach onto the leaf surface. Conidia become septate just prior to germination. A single germ tube extends from each germinating conidium, at the apex of which an ovate appressorium then differentiates (O'Connell et al., 2004). In this pre-penetration stage, energy is derived from the conidium itself as host resources remain inaccessible to the fungus. The *C. higginsianum* appressorium is a melanised, dome-shaped structure at the base of which an appressorial pore forms. The appressorial pore, even prior to penetration, is a site of effector secretion (Kleemann et al., 2012). Unlike some fungi, that invade their hosts via appressoria at stomatal pores, *C. higginsianum* penetrates epidermal cells, necessitating crossing of the cuticle and cell wall. *C. higginsianum*

appressoria therefore use considerable mechanical force from turgor pressure, alongside enzymatic degradation, to breach the leaf surface via a penetration peg which extends from the base of the appressorium (Perfect et al., 1999). Host penetration by *C. higginsianum* is expedited in Arabidopsis starch-related mutants, possibly due to their altered cell wall composition (Engelsdorf et al., 2017).

The roles of specific proteins in appressoria formation and host penetration in *C. higginsianum* are not fully characterised, but several genes have been implicated in these processes and prove important for successful infection. Melanisation of the appressorium appears critical, with mutants showing an appressorial melanin deficiency being incapable of penetrating host tissues (Liu et al., 2013). It is clear from work by Korn and colleagues that a fungal plasma membrane proton pump, PLASMA MEMBRANE H^+-ATPASE (ChPMA2), is required for host penetration (Korn et al., 2015). Despite the ability of Δ*ChPMA2* mutants to produce *in vitro* appressoria with normal turgor pressure, *in vivo* appressoria formation and even infection of wounded tissue appears not to be possible. The impact of losing *ChPMA2* on pathogenicity may be via a reduction in the capacity of the fungus for proton-coupled transport at this early stage of infection (Korn et al., 2015). The importance of proper cell cycle control in these early infection stages is also evident and has been studied more thoroughly in *M. oryzae* (Saunders et al., 2010). In *C. higginsianum*, the significance of G_1/S phase progression in appressoria development is highlighted by *BUDDING UNINHIBITED BY BENZIMIDAZOLE 2* mutants in which early transition to S phase results in premature nuclear division prior to septum formation. In this strain, septum formation appears impeded and multinucleated appressoria form with melanisation defects and reduced host penetration (Fukada et al., 2019).

Following host penetration, biotrophy is established through the formation of primary, biotrophic hyphae (BH). In the case of *C. higginsianum*, biotrophy is restricted to the initially infected cell (O'Connell et al., 2004). Biotrophic hyphae remain physically separated from the host cytoplasm by plant-derived, extra-invasive hyphal membranes (EIHM). The bulbous, biotrophic hyphae of *C. higginsianum* resemble the haustoria of biotrophic pathogens and form a multi-lobed, multi-septate structure (Yan et al., 2018). Biotrophic interfacial bodies, which decorate the biotrophic hyphae, further secrete effectors into the host (Kleemann et al., 2012). The biotrophic phase of infection is expected to involve pathogen acquisition of nutrients from the host. Limited evidence of specific transcriptional reprogramming to increase expression of nutrient transporters in this stage is observed for *C. higginsianum*. Induced expression during biotrophy is associated primarily with genes involved in secondary metabolism and putative effector genes (O'Connell et al., 2012).

Ultimately, the pathogen transitions to necrotrophy, during which narrower secondary necrotrophic hyphae develop and ramify into surrounding cells, causing cell death. Thus, shortly

after the switch from biotrophy to necrotrophy, macroscopic lesions become visible on the leaf. These water-soaked, necrotic lesions are often accompanied by chlorosis (O'Connell et al., 2004), with chlorotic halos surrounding expanding necrotic lesions. During necrotrophy, expression of lytic enzymes appears to be induced, including proteases and carbohydrate-active enzymes (CAZymes) which may act upon the plant cell wall. Alongside this, induced expression of transporters which may be involved in the uptake of sugars, amino acids and oligopeptides from the host is seen (O'Connell et al., 2012). Alkalinisation of local host tissues also occurs during necrotrophy, likely due to ammonia secretion from the fungus (O'Connell et al., 2012). The mechanisms governing the switch from biotrophy to necrotrophy are not well understood. Successful completion of the infection cycle appears dependent on the presence of pathogenicity-related genes encoded on an otherwise seemingly dispensable minichromosome. Loss of minichromosome 11 has been shown to result in infections which arrest in biotrophy, possibly due to a diminished suppression of the host immune response (Plaumann et al., 2018).

Localised hemibiotrophy, the restriction of biotrophy to the initially infected cell as seen in *C. higginsianum*, is not conserved across hemibiotrophic *Colletotrichum* species. Other hemibiotrophic species, such as *Colletotrichum graminicola,* establish a biotrophic phase which extends across numerous cells, with biotrophic hyphae invading neighbouring mesophyll cells before switching to necrotrophy (Mims and Vaillancourt, 2002). Additionally, in *C. graminicola*, at the edges of expanding lesions further biotrophic invasion of adjoining cells occurs while at the central portions of the lesion the pathogen transitions to necrotrophy. This is potentially facilitated by the action of diffusible factors inducing a susceptibility to biotrophic invasion (Torres et al., 2014).

The final stage of the *C. higginsianum* lifecycle is conidiation. *C. higginsianum* has no known sexual morph (Damm et al., 2014). Asexual conidia are mitotically produced by sporulating acervuli which form on necrotic tissues at the end of the infection. Within the acervulus, conidia are formed in specialised hyphae called conidiophores, beyond which narrow setae may extend (O'Connell et al., 2004). These conidia are likely disseminated by rain splash, wind, and, potentially, insects (Gasparoto et al., 2017). Unlike obligate pathogens, *C. higginsianum* can be cultured axenically, facilitating its study in a laboratory setting.

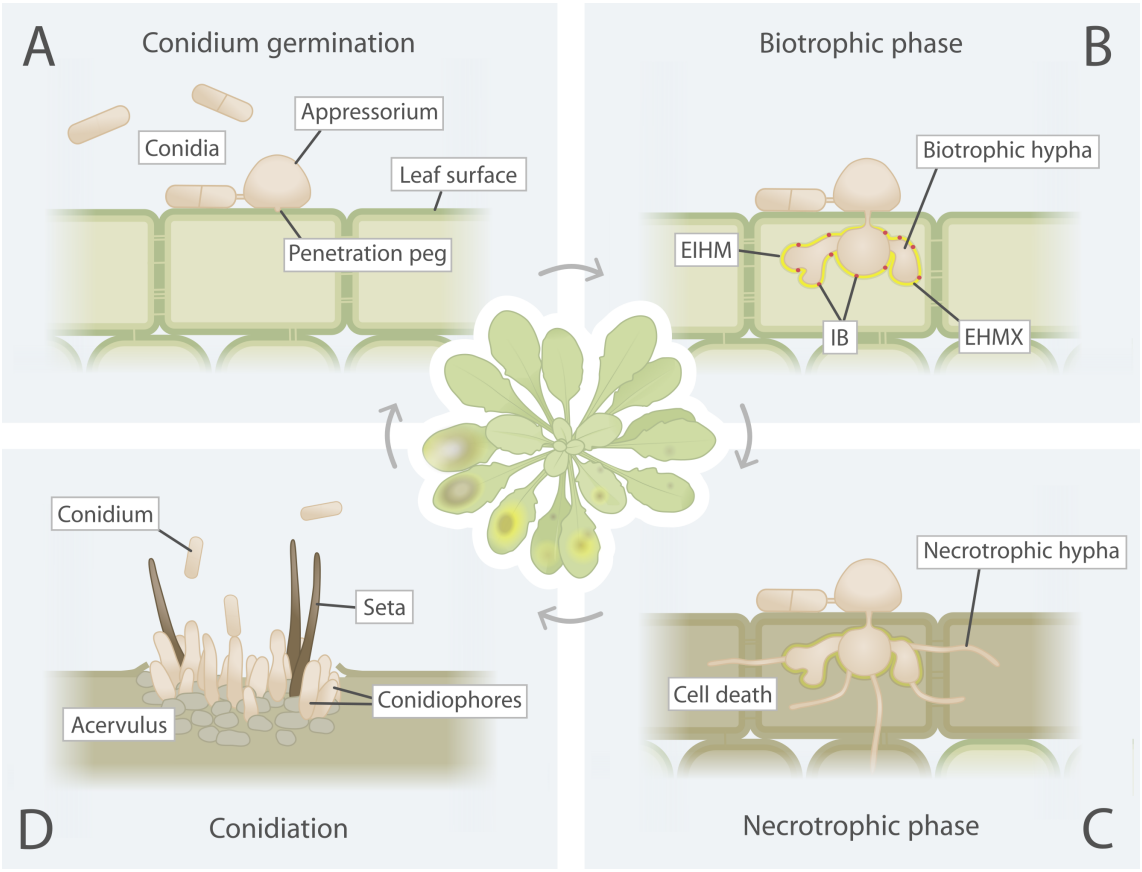


Figure 1.1: **Schematic depicting the infection lifecycle of *C. higginsianum*.** **(A)** Asexual conidia germinate and appressoria form to penetrate host cells via a penetration peg. **(B)** In the biotrophic phase of infection, biotrophic hyphae develop within the primary infected cell. Effector molecules are secreted from biotrophic interfacial bodies (IB, red) and translocate from pathogen to host. Primary, biotrophic hyphae are separated from the host cytosol by the extra-invasive hyphal membrane (EIHM), a plant-derived membrane continuous with the plant plasma membrane. The extrahaustorial matrix (EHMX) is coloured yellow for illustrative purposes but is continuous with the apoplast. **(C)** The fungus enters the necrotrophic phase as secondary, necrotrophic hyphae expand into adjacent cells, killing the host tissue. At this stage, macroscopic disease lesions become visible. **(D)** At the end of the infection cycle conidiation occurs, resulting in the production of numerous asexual spores which can go on to infect further tissues/hosts.

1.3.2 *C. higginsianum* nutrient uptake and metabolism

To facilitate fungal virulence, *C. higginsianum* genes involved in nutrient uptake and metabolism are selectively expressed at specific stages of infection. As mentioned in **Section 1.3.1**, the nature of *C. higginsianum* nutrient uptake during the biotrophic phase of infection is unclear. Beyond mobilisation of lipids and carbohydrates from the conidium for germination and host penetration the fungus is expected, as a heterotroph, to become quickly reliant on its host for nutrient acquisition. CAZymes secreted early in infection may act to weaken the host cell wall to aid

penetration, with appressorially upregulated CAZyme genes targeting cellulose, hemicellulose, cutin and pectin (O'Connell et al., 2012). Further, degradation of the host cell wall may provide a carbon source for pathogen uptake.

Yuan and colleagues characterised *C. higginsianum* homologues of known *C. graminicola* hexose transporters (HXTs), termed ChHXTs 1-5, and an additional hexose transporter, ChHXT6, in terms of their ability to transport various hexoses (Yuan et al., 2021, Lingner et al., 2011). *ChHXT4* and *ChHXT6* have been implicated in infection due to strains lacking these genes showing reduced virulence. While *ΔChHXT4* showed reduced establishment of both biotrophic and necrotrophic hyphae, *ΔChHXT6* showed only a reduction in necrotrophic hyphae (Yuan et al., 2021). Thus, it is possible that the ability of the fungus to transport sugars is essential for full virulence prior to necrotrophy.

While O'Connell and colleagues report an absence of specific upregulation of nutrient transporters during biotrophy, and postulate that the biotrophic phase of *C. higginsianum* infection may be more critical for effector secretion and host defence suppression than nutrient acquisition, a number of pathogen genes annotated as being related to sugar transport appear to be expressed early in infection or through the duration of infection into necrotrophy, with 13 *C. higginsianum* genes for sugar transporters appearing to be expressed only during biotrophy (O'Connell et al., 2012). However, Engelsdorf and colleagues reason that host carbon availability becomes important for infection predominantly in the necrotrophic phase, with a perceived lower significance for carbohydrate availability in the biotrophic phase based on observations of normal infection rates in plants with induced carbon-starvation early in the infection process. Indeed, the authors conclude that during biotrophy *C. higginsianum* does not require carbohydrate supply from the host (Engelsdorf et al., 2013). Thus, the extent to which *C. higginsianum* relies on nutrient acquisition from the host during its biotrophic phase remains unclear.

Fungi can utilise a range of carbon sources, with most phytopathogenic fungi believed to uptake hexoses as a carbon source. In the biotrophic infection stage of some hemibiotrophic pathogens, such as *M. oryzae*, expression of invertases that cleave sucrose into hexoses, as well as transporters that allow hexose uptake, could indicate the importance of plant hexoses for carbon acquisition during biotrophy (Lindsay et al., 2016). Some fungal pathogens, such as the biotroph *U. maydis*, can directly uptake sucrose from the apoplast without first hydrolysing it to hexoses (Wahl et al., 2010). Some fungi, such as arbuscular mycorrhizal fungi have been seen to uptake carbon from plants both as sugars and lipids (Keymer et al., 2017). *In vitro*, *C. higginsianum* is able to grow on various carbon sources such as glucose, sucrose, glycerol, acetate, ethanol, and acetaldehyde (Gu et al., 2019, Chanda et al., 2008).

As *C. higginsianum* transitions into necrotrophy, expression of genes encoding CAZymes and other lytic enzymes, and plasma membrane transporters drastically increases, reflecting the destruction of host tissues and concomitant uptake by the pathogen of available nutrients (O'Connell et al., 2012). Starch content has been shown to decrease in leaves during necrotrophic growth of *C. higginsianum* compared to mock-inoculated leaves, with a concomitant increase in soluble sugar content likely due to a reduced conversion of sugars to transitory starch during this phase of infection (Engelsdorf et al., 2013).

Glycerol-3-phosphate (G3P) has been implicated in Arabidopsis defence to *C. higginsianum*, with increased G3P in the host following inoculation. Mutation or overexpression of plastidial genes altering host G3P levels revealed a positive correlation between host G3P content and *C. higginsianum* resistance (Chanda et al., 2008). Amino acid biosynthesis and homeostasis also appear to be important contributors to *C. higginsianum* pathogenicity, and may suggest uptake of amino acids by the fungus, potentially as a source of nitrogen. For example, abolishment of the amino acid transporter LYSINE HISTIDINE TRANSPORTER 1 in Arabidopsis confers enhanced resistance to a number of pathogens, including *C. higginsianum* (Liu et al., 2010). Also, reduced host penetration in *C. higginsianum* arginine-auxotrophic mutants has been demonstrated despite normal conidia germination and appressoria formation (Takahara et al., 2012).

Key fungal metabolic processes are modulated during infection, including in response to nutrient availability. Carbon catabolite repression (CCR) is the process by which fungal genes involved in carbon metabolism are regulated to ensure the more efficient use of preferred carbon sources over other alternative carbon sources present. In this process, genes involved in metabolism of alternative carbon sources are repressed by the presence of the preferred carbon source. A CATABOLITE RESPONSIVE ELEMENT TF, CreA, and homologues thereof, has been implicated in CCR in a number of filamentous fungal species (Hong et al., 2021). In *M. oryzae*, TREHALOSE-6-PHOSPHATE SYNTHASE 1 (TPS1), a sensor of glucose, is critical in CCR alongside CreA with direct links to pathogenicity (Fernandez et al., 2012, Hong et al., 2021). It is possible that similar systems of glucose-sensing and CCR exist in *C. higginsianum*. The target of rapamycin (TOR) pathway also functions in nutrient sensing in fungi (Loewith and Hall, 2011), promoting growth in response to nutrient perception, and is suggested to be involved in initiation of biotrophic growth following host penetration in *M. oryzae* (Fernandez et al., 2014). Analogously to CCR, nitrogen catabolite repression also occurs in fungi and has been suggested to regulate *C. higginsianum* gene expression (Takahara et al., 2012).

1.3.3 *C. higginsianum* strains, genetic resources, and genetic manipulation

The genomes of phytopathogens face continual pressure to adapt due to the co-evolutionary arms-race they are locked in against their hosts. Two *C. higginsianum* strains are particularly prevalent within laboratory studies of this fungus: IMI 349063, isolated from *Brassica campestris* subsp. *chinensis* in Trinidad (O'Connell et al., 2004), and MAFF 305635, isolated from *Brassica rapa* var. *perviridis* in Japan (Horie et al., 1988). There are three published genomes of *C. higginsianum*, with two for strain IMI 349063 (O'Connell et al., 2012, Zampounis et al., 2016) and one for MAFF 305635 (Tsushima et al., 2019a). From these genome assemblies, the fungus is thought to possess around 14,000 protein-coding genes. The genetic resources available for *C. higginsianum*, including these genome assemblies, are introduced in more detail in Section 5.1.1 due to their particular relevance to the work in Chapter 5.

C. higginsianum is amenable to genetic transformation, including by *Agrobacterium*-mediated transformation. This is exemplified by work by O'Connell and colleagues (O'Connell et al., 2004) and is described more thoroughly in Chapter 4 Section 4.1.2. The study of *C. higginsianum* is greatly benefited by its genetic tractability, rendering the *C. higginsianum*-Arabidopsis pathosystem one in which both host and pathogen can be genetically manipulated. This model pathosystem thereby has great scope for considering plant-pathogen interactions more generally, in addition to investigating *C. higginsianum* pathogenicity specifically.

1.3.4 The effector repertoire of *C. higginsianum*

As a phytopathogen, *C. higginsianum* is known to secrete an array of effectors which act to facilitate infection. Relative to *C. graminicola*, *C. higginsianum* has an expanded number of genes encoding candidate effectors. This may be reflective of the wider range of host plants that can be infected by *C. higginsianum* (O'Connell et al., 2012).

Effectors may act in the host apoplast or translocate into the cytoplasm, from where they may target various organelles. As detailed in Section 1.2.2, the mechanisms of secretion and translocation of fungal effectors are not yet fully understood. Early effector secretion from *C. higginsianum* occurs at the appressorial penetration pore (Kleemann et al., 2012). The surface of *C. higginsianum* biotrophic hyphae are studded with numerous interfacial bodies at which effectors appear to accumulate. These interfacial bodies somewhat resemble the *M. oryzae* BIC and may similarly be a key site of effector delivery to the host (Kleemann et al., 2012). The passage of effectors from pathogen to host may therefore rely at least in part on the action of EVs. Recently,

the presence of *C. higginsianum* EVs has been confirmed (Rutter et al., 2022), but their relevance in effector secretion remains to be determined.

Transcriptomics experiments spanning infection stages have charted the expression of putative *C. higginsianum* effector genes during infection (Kleemann et al., 2012, O'Connell et al., 2012). From these data, predicted effectors with distinct expression profiles have been identified, and several waves of putative effector expression characterised, corresponding to different stages of infection (Kleemann et al., 2012). The transition from biotrophy to necrotrophy appears to involve the differential expression antagonistic effectors suppressing and promoting cell death (Kleemann et al., 2012). Transient expression of effector candidates in *N. benthamiana* revealed targeting of a number of subcellular locations including nuclei, peroxisomes, microtubules and Golgi bodies (Robin et al., 2018). These transcriptomic data (O'Connell et al., 2012, Dallery et al., 2017) were further used by Ohtsu and colleagues to identify additional putative effector proteins targeted for secretion and expressed predominantly during early infection (Ohtsu et al., 2023). In addition to identifying nucleocytoplasmic cell-to-cell mobile effectors which may be travelling via and indirectly targeting plasmodesmata to manipulate symplastic connectivity (Ohtsu et al., 2023), this localisation screen identified putative effectors localising to novel host subcellular locations including plasmodesmata and the chloroplast as well as previously seen localisations such as to the nucleus and microtubules (Jennings, 2021).

Specific *C. higginsianum* effectors have been implicated in suppression of PTI (Takahara et al., 2016). Some *C. higginsianum* effectors trigger cell death in *N. benthamiana* (Takahara et al., 2021, Tsushima et al., 2021, Kleemann et al., 2012), while others are seen to suppress cell death (Kleemann et al., 2012). The temporal control of these antagonistic effectors may be coordinated such that they facilitate the transition from biotrophy to necrotrophy. Further to the abovementioned proteinaceous effectors of *C. higginsianum*, the fungus also encodes 14 biosynthetic gene clusters (BGCs) expressed during host penetration and biotrophic growth in Arabidopsis, highlighting the potential for as yet uncharacterised secondary metabolite effectors in this species (Dallery et al., 2017).

1.4 Oomycete plant pathogens

1.4.1 Oomycetes as phytopathogens

Despite resembling some fungi in terms of their reproduction via spores and their filamentous growth, oomycetes are evolutionarily distinct from fungi. Unlike the chitinaceous cell walls of fungi, oomycetes have cell walls largely comprising of β-glucans and cellulose, and little or no chitin. The major β-glucans making up oomycete cell walls include β-1,4-glucans, β-1,3-glucans and β-1,6-

glucans in varying fractions depending on the oomycete species (Melida et al., 2013). β-glucans therefore feature prominently in the study of oomycete PAMP-recognition, being implicated in the perception and response to oomycetes such as *P. infestans* and *P. sojae* (Robinson and Bostock, 2015). Perception of β-glucans is not restricted to host responses to oomycetes, with β-glucans also being found in cell walls of many filamentous phytopathogens including fungi, though the PRRs guiding responses to β-glucans are not yet well characterised (Fesel and Zuccaro, 2016, Wanke et al., 2020). Some oomycetes, including *Phytophthora* spp. produce and secrete small elicitin proteins which can be perceived by plants as PAMPs and elicit an immune response (Ricci et al., 1989, Du et al., 2015).

Among oomycetes, the order Peronosporales has received considerable focus in research, containing both hemibiotrophic *Phytophthora* species and biotrophic downy mildew species. In biotrophy, oomycetes form specialised hyphae termed haustoria which breach host cell walls, being separated from the plant cell by the plant-derived extrahaustorial membrane (EHM).

1.4.2 Oomycetes produce an array of effectors, including RxLR effectors

In contrast to fungi, oomycetes encode a number of effectors with relatively well-characterised motifs at their N-termini, which may be involved in effector translocation. However, the mechanisms of translocation of intracellular oomycete effectors from pathogen to host are not yet fully understood. A large proportion of oomycete effectors that translocate into the host cytoplasm possess an N-terminal motif (downstream of the N-terminal signal peptide for secretion) consisting of the consensus amino acid sequence RxLR (Arg, any amino acid, Leu, Arg). This sequence is commonly followed by a sequence of acidic amino acids, forming an RxLR-dEER motif (Rehmany et al., 2005). The role of the RxLR motif in effector secretion or translocation remains somewhat unclear. While it has long been believed that the RxLR motif was involved in effector translocation into the host (Whisson et al., 2007), some effectors are cleaved of their RxLR motif within the pathogen, prior to secretion (Wawra et al., 2012). This would suggest that translocation is independent of the RxLR motif, which may instead play a role in secretion. It has recently been suggested that RxLR effector uptake may involve CME (Wang et al., 2023b). RxLR effectors are frequently found in relatively gene-sparse, repeat-rich regions of the genome, which may aid plasticity of the effector repertoire through gain or loss of RxLR genes (Haas et al., 2009).

Another large group of oomycete effector proteins are the Crinkler (CRN) effectors (Torto et al., 2003), typified by an LxLFLAK motif following the signal peptide at their N-termini, which has been implicated in translocation into the host (Schornack et al., 2010). Similarly, a number of CHxC (Cys, His, any amino acid, Cys) oomycete effectors have been identified (Links et al., 2011, Kemen et al., 2011). This motif has recently been redefined as CxxCxxxxxG (simplified to “CCG”) and appears to

be enriched in *Albugo* relative to *Phytophthora* or *Hyaloperonospora* species (Furzer et al., 2022). This motif, like the RxLR and LxLFLAK motifs is suggested to play a role in effector translocation by hitherto uncharacterised mechanisms (Kemen et al., 2011). These conserved N-terminal hallmarks of oomycete effectors facilitate their identification from genomic data.

The C-termini of oomycete effectors, like fungal effectors, can be highly diverse as the portion of the protein expected to confer virulence and/or avirulence functions (Bos et al., 2006). Within the C-terminal region of oomycete RxLR effectors, W, Y, and L motifs have been identified and found to frequently occur in tandem repeat modules (Jiang et al., 2008). These motifs appear to form α-helical WY domains in which structure is conserved despite variable sequences, potentially facilitating effector evolution (Boutemy et al., 2011).

1.5 *Hyaloperonospora arabidopsidis*

Hyaloperonospora arabidopsidis (Hpa), previously *Peronospora parasitica* and *Hyaloperonospora parasitica*, is an oomycete capable of infecting Arabidopsis, wherein it causes downy mildew disease (Koch and Slusarenko, 1990). *Hpa*-related downy mildew species infect numerous species of the *Cruciferae* family, including agronomically important crops such as oilseed rape, broccoli, and cabbage, marking the importance of *Hpa* as a model for studying downy mildew crop pathogenesis (McDowell, 2014). *Hpa* belongs to the *Peronosporaceae* family, members of which are obligate biotrophic plant pathogens and therefore cannot be cultured axenically. While some *Peronosporaceae* genera contain species which can be genetically manipulated, such as *P. sojae* (Fang and Tyler, 2016), genetic transformation of *Hpa* has not yet been demonstrated.

N. benthamiana is considered a nonhost for *H. arabidopsidis*, but many *H. arabidopsidis* isolates are virulent on various Arabidopsis ecotypes (McLellan et al., 2022). Genotype specificity is seen in interactions between *Hpa* isolates and Arabidopsis accessions. Within this pathosystem, a variety of degrees of compatibility can therefore be seen (Holub et al., 1994, Krasileva et al., 2011), lending strength to the use of this pathosystem for studying host-pathogen co-evolution. *H. arabidopsidis* effectors termed ARABIDOPSIS THALIANA RECOGNISED (ATR) effectors appear to be recognised by cognate host RECOGNITION OF PERONOSPORA PARASITICA (RPP) receptors, conferring resistance through the hypersensitive response (Holub et al., 1994).

Transcriptome profiling of the *Hpa*-Arabidopsis pathosystem by Asai and colleagues has revealed a number of genes involved in pathogen virulence and host responses (Asai et al., 2014). In this study, host and pathogen expression data were acquired following challenge of Arabidopsis Col-0 with both virulent (Waco9, compatible interaction) and avirulent (Emoy2, incompatible interaction) *Hpa* isolates. This work revealed that ATR1, an effector of *Hpa* recognised by RPP1 (Rehmany et al., 2005), is expressed by Emoy2 but not Waco9 (Asai et al., 2014). Similarly, Emoy2

is recognised by the host NLR RPP4 (van der Biezen et al., 2002) at least in part due to expression of the effector HaRxL103 (Asai et al., 2018). Further, investigation of this pathosystem at a cellular level is increasingly possible, with development of Arabidopsis hormone reporter lines (Ghareeb et al., 2020) and haustoriated cell-specific transcriptomics (Asai et al., 2023) being demonstrated in this system.

1.5.1 *H. arabidopsidis* lifecycle and infection strategy

As an obligate biotroph, the full lifecycle of *Hpa* relies on a live host. The infection process of *H. arabidopsidis* is illustrated in **Figure 1.2**. Asexual *Hpa* conidiospores germinate on the leaf surface and form appressoria. Unlike *C. higginsianum*, the initial penetration of the host by *Hpa* occurs between anticlinal cells of the epidermis. The initial penetrative hyphae that form are intercellular, existing between epidermal cells (Coates and Beynon, 2010). Often, a small number of haustoria expand into the adjacent epidermal cells prior to the branching of hyphae which extend into the intercellular space and form numerous haustoria in mesophyll cells. Surrounding haustoria is the extracellular extrahaustorial matrix (EHMX) and EHM (Mims et al., 2004), across which *Hpa* secretes effectors into, and takes up nutrients and water from, haustoriated host cells. As infection progresses through the cell layers of the leaf, hyphae reaching substomatal cavities develop conidiophores. Developing conidiophores extend through stomata, elongating and ramifying into dendriform structures at which asexual conidiospores develop (Soylu and Soylu, 2003). Unlike *C. higginsianum*, *H. arabidopsidis* is additionally able to produce sexual oospores – an avenue by which genetic diversity can be introduced. In incompatible interactions in which the host accession is resistant to the *Hpa* strain it faces, the hypersensitive response is triggered shortly after host penetration, when early haustoria form and secrete effectors recognised by the host, arresting infection.

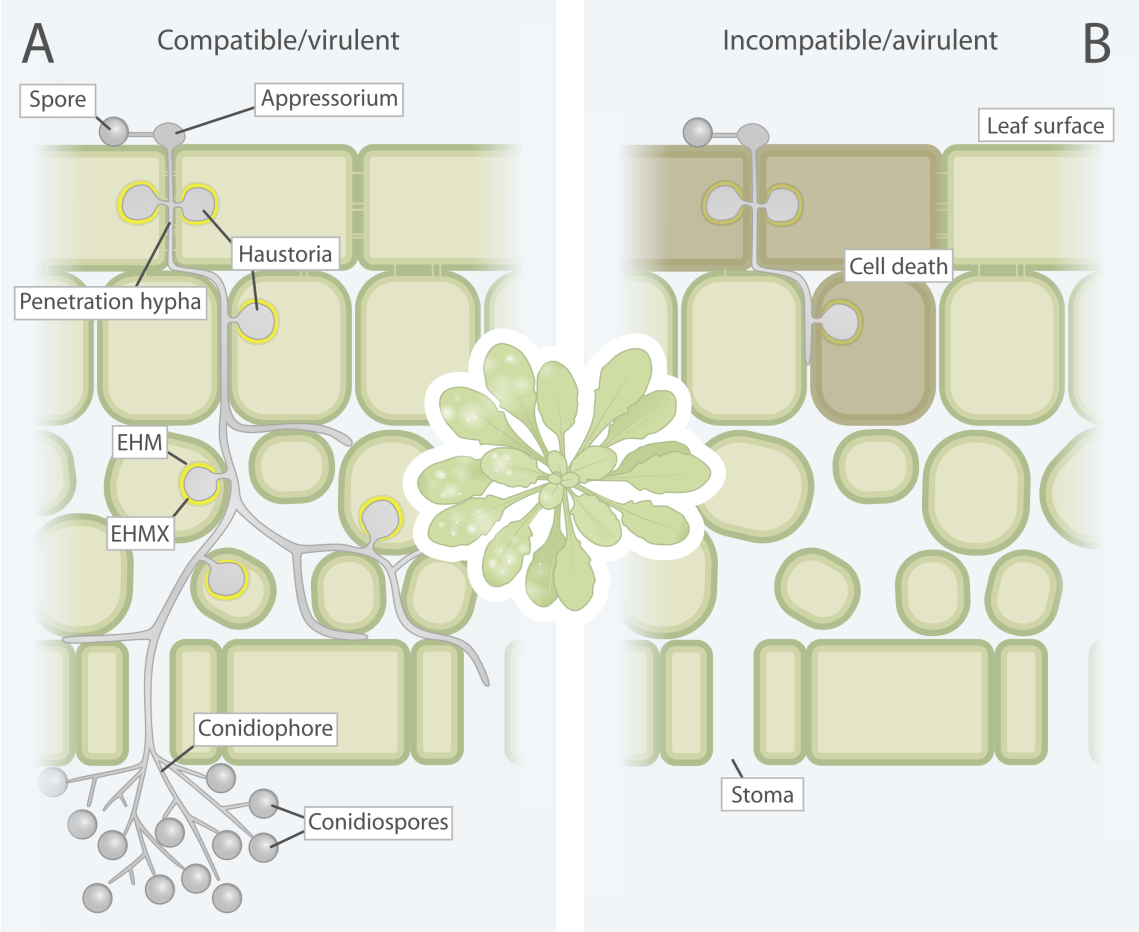


Figure 1.2: **Schematic depicting the infection lifecycle of *H. arabidopsidis*. (A)** During a compatible interaction, germinating conidiospores land on the host tissue and penetrate the epidermis via an appressorium, and penetrative hyphae develop. Haustoria develop from these hyphae, predominantly in the mesophyll cell layer, invaginating the host cells and establishing a key site of interaction between the pathogen and host at which effectors are secreted and nutrients taken up by the pathogen. The pathogen is separated from the host by the plant-derived extrahaustorial membrane (EHM) and the extrahaustorial matrix (EHMX). Further progression of hyphae leads to development of conidiophores extending through stomata, and production of conidiospores. **(B)** During an incompatible interaction, localised cell death results in rapid arrest of the infection prior to extensive host colonisation. This is triggered by effector-perception by host immune receptors. This schematic is based on a figure from Coates and Beynon, 2010.

1.5.2 *H. arabidopsidis* nutrient uptake and metabolism

Due to the obligate biotrophic nature of *Hpa*, its reliance on the host for nutrients is absolute. Despite this, little seems to be directly known about nutrient uptake in this pathogen. It is assumed that haustoria are the site of nutrient uptake for the pathogen, though direct evidence for this is lacking. Characterisation of *Phytophthora infestans* metabolomic models suggests that oomycetes are able to metabolise plant-derived starch, hexoses, disaccharides and organic acids (Rodenburg

et al., 2018). *P. infestans* invertases have been seen to localise to haustoria and be expressed under infection-specific control (upregulated pre-infection and in biotrophy), indicating they may play a role in carbon acquisition for the pathogen (Kagda et al., 2020).

Reflective of its obligate biotrophy, *Hpa* lacks genes involved in nitrogen assimilation seen in other oomycetes (Baxter et al., 2010), with the preferred source of nitrogen in general believed to be amino acids for both fungi and oomycetes. Primary amino acid metabolism in the host appears to be important in the *Hpa-Arabidopsidis* pathosystem. Stuttmann and colleagues identified Arabidopsis mutants perturbed in amino acid homeostasis and over-accumulating certain amino acids that show a specific, increased resistance to *Hpa*, but not to obligate biotrophic fungus *Golovinomyces orontii* (Stuttmann et al., 2011). The authors attribute this reduced susceptibility to host threonine accumulation, but the mechanisms by which this impacts infection are not fully characterised (Stuttmann et al., 2011).

1.5.3 *H. arabidopsidis* genetic resources

The genome of *H. arabidopsidis* isolate Emoy2 was published in 2010 by Baxter and colleagues (Baxter et al., 2010). A combination of Sanger and Illumina sequencing of DNA from asexual spores was used to assemble a genome of approximately 100 Mb, containing 14,543 precited genes (Baxter et al., 2010). This genome assembly, compared to genomes of *Phytophthora* species, showed large reductions in the numbers of predicted RxLR effectors, reduction in genes for secreted degradative enzymes, and loss of genes for some metabolic pathways such as nitrogen and sulphur assimilation, all of which may reflect conserved adaptations for obligate biotrophy (Baxter et al., 2010). Recently, genomes of related *Hyaloperonospora* species have been published (You et al., 2021, Wu et al., 2023), which may facilitate elucidation of host-specific adaptation in these pathogens.

1.5.4 *H. arabidopsidis* effectors

Like other members of the Peronosporales order, *Hpa* possesses a number of candidate effector proteins with RxLR motifs. Based on the genome sequence, 134 high-confidence RxLR *Hpa* effector candidates were identified in 2010 (Baxter et al., 2010). This is a relatively small number of effectors compared to related hemibiotrophic *Phytophthora* species, perhaps reflective of the enduring biotrophic lifestyle of *Hpa* (Baxter et al., 2010). The subcellular localisation of a set of 49 of these effectors transiently expressed as fluorophore fusions in *N. benthamiana* was published in 2012, revealing effectors localising to the nucleus, cytoplasm and membranes, as well as one associated with the tonoplast (Caillaud et al., 2012). In this work, the majority of effector candidates were tagged with eGFP or RFP at their N-termini. None of the *Hpa* RxLR effectors

predicted by Wolf Psort to localise to the chloroplast did so in these experiments (Caillaud et al., 2012) – possibly due to N-terminal tags interfering with the chloroplast transit peptide at the N-terminus of the protein.

As mentioned in Section 1.5, *Hpa* RxLR effector ATR1 is recognised by the host NLR RPP1 (Rehmany et al., 2005), and HaRxL103 is recognised by RPP4 (Asai et al., 2018). Other ATR-RPP interactions have also been seen, such as the recognition of ATR13 by RPP13 (Allen et al., 2004). Some specific *Hpa* effectors have been characterised aside from identification of corresponding host NLRs. HaRxL23 acts to suppress host defences (Deb et al., 2018a) and is structurally similar to bacterial effector AvrE1 (Deb et al., 2018b), and HaRxLL470 appears to regulate host gene expression via interaction with endogenous transcription factor HY5, suppressing activation of defence genes (Chen et al., 2021b). HaRxL77 has been shown to increase host symplastic connectivity via increasing plasmodesmal permeability (Liu et al., 2022). The effector repertoire of *Hpa* is not restricted to RxLR effectors, with a number of *Hpa* effectors being described which lack the RxLR motif (Bailey et al., 2011). Dunker and colleagues explored the role of non-RxLR *Hpa* effector protein CYSTEINE-RICH PROTEIN 1, which appears to suppress the host hypersensitive response in order to facilitate infection (Dunker et al., 2021).

1.6 The chloroplast during infection: a target of pathogen effectors

The involvement of the chloroplast in infection and defence is critical and broad. Effectors targeting the chloroplast may apply a diverse range of mechanisms to suppress the host immune response or to promote pathogen virulence due to the abundance of host processes taking place within the plastid.

1.6.1 Photosynthesis generally appears suppressed during infection

Chloroplasts are best known as the site of photosynthesis. Plant carbohydrate biosynthesis is underpinned by photosynthesis – the process by which plants fix carbon dioxide to produce organic carbohydrates, with the conversion of light energy to chemical energy. Carbohydrates produced during photosynthesis form the basis for synthesis of amino acids, fatty acids and phytohormones.

A decrease in photosynthetic rate and down-regulation of photosynthesis-related genes is often observed following inoculation with either virulent or avirulent pathogens of diverse taxa and various lifestyles (Berger et al., 2007). This contrasts with an increase in photosynthetic capacity during plant interactions with beneficial, symbiotic microorganisms (Ye et al., 2019). It has been variously reasoned that downregulation of photosynthesis in response to pathogens may occur via action of pathogen effectors (Truman et al., 2006) or feedback regulation from sugar signalling

(Rolland et al., 2006). Lowered photosynthesis-related protein turnover may provide an opportunity for the plant to temporarily increase resources allocated for cost-intensive defence responses. Carbohydrate partitioning in the host is shifted during infection, and often infected source tissues transition to sink tissues in order to accommodate the increased demand for photosynthetic assimilates (Berger et al., 2007, Rojas et al., 2014).

Chloroplasts are also the host sites of crucial steps in the synthesis of amino acids, another aspect of primary metabolism implicated in plant-pathogen interactions (Liu et al., 2010). Further, a number of these amino acids are precursors for phytohormone biosynthesis (Fabregas and Fernie, 2022).

1.6.2 The chloroplast is a key site for production of defence signals

The chloroplast is a prime target for pathogen effectors as the site of synthesis of a number of immune signalling molecules, including phytohormones, ROS and nitric oxide (NO). The importance of these molecules in defence responses is evident, but the interplay between these signals is convoluted.

The chloroplast is a site of JA synthesis, a phytohormone particularly involved in defence responses against necrotrophic pathogens (Macioszek et al., 2023) and in ISR (Van der Ent et al., 2009). JA synthesis can be targeted by effectors, such as phytoplasma effector SAP11 which downregulates JA synthesis via destabilisation of host TFs (Sugio et al., 2011). While ethylene biosynthesis occurs in the cytoplasm, methionine, the precursor for ET biosynthesis is produced in the chloroplast. As well as being an important regulator of plant growth, ET is also implicated in responses to biotic, as well as abiotic, stresses. Indeed, ET has been implicated in stomatal closure (Desikan et al., 2006), and so may be important in physical defence against pathogen stomatal entry. Further, ET may dampen symptoms of necrotrophic infection (Thomma et al., 1999) while promoting cell death caused by biotrophic infection (Hoffman et al., 1999), pointing to a potential role in resistance. Salicylic acid is a key phytohormone involved in systemic acquired resistance and the hypersensitive response, and the synthesis of SA involves the chloroplast. Biosynthesis of SA occurs via the isochorismate (ICS) pathway and the phenylalanine ammonia-lyase (PAL) pathway. While the ICS pathway is suggested to be the predominant pathway (Wildermuth et al., 2001), both ICS and PAL pathways rely on chloroplast-produced precursor chorismate. Pathogen effectors have been seen to target SA production (Liu et al., 2014). For example, the *Ustilago maydis* chorismate mutase CMU1 appears to translocate into the host and interact with plant cytosolic chorismate mutases and lead to a repression in SA synthesis (Djamei et al., 2011). The importance of SA signalling in defence is also exemplified by the increased susceptibility of various SA-signalling-defective plant mutants including *SALICYLIC ACID INDUCTION DEFICIENT* mutants *sid1* and *sid2*

(Nawrath and Metraux, 1999, Nawrath et al., 2002, Wildermuth et al., 2001). NahG transgenic plants in which a bacterial SA hydroxylase is expressed are incapable of accumulating SA, resulting in perturbations in defence responses which often lead to enhanced susceptibility (Delaney et al., 1994). Further, SA appears to influence the production of other defence molecules, including ROS.

The primary oxidative burst seen in response to pathogen recognition occurs in the host apoplast, largely by the activation and action of plasma membrane RBOHs. However, chloroplasts, as well as mitochondria and peroxisomes, are also involved in plant production of ROS, which may be key defence signals (Kuzniak and Kopczewski, 2020). Observations of the light-dependency of the hypersensitive response have led to the implication of chloroplastic ROS production in the defence response (Liu et al., 2007). Targeting of the synthesis of phytohormones, phytohormone precursors, and other defence signals such as ROS by pathogen effectors can result in perturbations in immune responses.

1.6.3 Chloroplast repositioning in response to infection

During infection the positioning of chloroplasts appears altered, potentially indicating their involvement in infection-related processes. The accumulation of chloroplasts proximal to the nucleus, a process known as chloroplast perinuclear clustering, is widely observed in response to biotic stresses (Ding et al., 2019). Perinuclear clustering involves the formation of stromules - thin stroma-filled tubules which project from the chloroplast surface and can establish physical contact with the nucleus and other chloroplasts. ROS accumulation has been associated with inducing stromule formation (Brunkard et al., 2015, Caplan et al., 2015), which may reflect the mechanism of stromule formation in response to infection (Savage et al., 2021). The formation of plastid-nuclear complexes and stromules may facilitate signalling between the chloroplast and nucleus (Mullineaux et al., 2020), possibly via ROS transfer (Breeze and Mullineaux, 2022). Perinuclear clustering is also seen in response to viral infections such as with turnip mosaic virus. Here, in infection of *N. benthamiana*, perinuclear clustering has been correlated with resistance to infection and expression of the host chloroplast NADH dehydrogenase-like complex M subunit gene (Zhai et al., 2021). Chloroplast-nuclear retrograde signalling may allow chloroplast-targeted proteins to effect nuclear gene expression. A chloroplastic localisation of pathogen effectors could therefore reflect an extensive array of potential impacts for the host.

In addition to perinuclear clustering, relocation of chloroplasts to the haustorial interface during infection with *P. infestans* has also been observed, in a process which may involve pathogen-induced stromule formation (Savage et al., 2021).

1.6.4 A number of effector proteins localise to the chloroplast

Evidencing the importance of chloroplast-manipulation during infection, effectors from various pathogens have been reported to translocate into host chloroplasts (Petre et al., 2016). These effectors may suppress chloroplast function (Xu et al., 2019), potentially in order to interfere with chloroplastic production of defence signals, as is proposed for the *P. syringae* chloroplast-localised effector Hopi1 (Jelenska et al., 2007). Some pathogen effectors, such as fungal effector RsCRP1, are dually targeted to chloroplasts and mitochondria (Tzelepis et al., 2021). Further, viral proteins also appear able to target the chloroplast to suppress host defences (Alam et al., 2021).

Several specific host chloroplast proteins have been identified as targets of pathogen effectors, such as RESISTANCE TO PHYTOPHTHORA 1 (RPH1) which is shown in Arabidopsis to be critical for activation of defence responses to *Phytophthora brassicae* – with the pathogen believed to target RPH1 with an effector molecule (Belhaj et al., 2009). An INTEGRIN-LIKE (ITL) effector of the necrotrophic fungus *S. sclerotiorum,* localises to chloroplasts during transient expression in *N. benthamiana*, and interacts with the chloroplast-localised Ca^{2+}-sensing receptor (CAS) in Arabidopsis (Tang et al., 2020). CAS has been implicated in chloroplastic defence responses triggered by PAMP perception, being required for a transient chloroplastic Ca^{2+} signal, SA biosynthesis, and possibly having a role in both the suppression of chloroplast gene expression and the induced expression of nuclear-encoded defence genes (Nomura et al., 2012). SsITL targeting of CAS appears to perturb chitin-triggered SA signalling pathways, facilitating infection (Tang et al., 2020).

Additionally, indirect targeting of the chloroplast by pathogen effectors may be achieved by the manipulation of nuclear-encoded chloroplast gene (NECG) expression. In this way, nuclear-localised effectors may target chloroplastic processes at a transcriptional level. Demonstrating this capacity, PAMP perception has been correlated with a large-scale downregulation of NECGs (Zabala et al., 2015). Further, post-transcriptional perturbation of chloroplast function via nuclear-encoded chloroplast proteins could be achieved by disrupting mechanisms for chloroplast protein import.

1.7 Carbohydrates are stored as transitory starch in the chloroplasts of leaves

Starch is the major storage carbohydrate of plants. It may be possible for a pathogen to mobilise starch as a source of energy, or to impede the production of starch in order to avail itself of an increased pool of soluble sugars, in addition to targeting host sugar as it is produced by photosynthesis. In light of the suppression of photosynthesis during infection, starch may present

an important source of carbohydrates for the pathogen to exploit. Alternatively, manipulation of starch biosynthesis and metabolism by a pathogen may imply interference in sugar signalling or priming of the defence response. The proteins of starch synthesis, branching/debranching, degradation and granule initiation thereby represent potential novel effector targets that have not yet been thoroughly investigated in this context.

Starch granules are composed of two α-glucans: amylopectin and amylose, which themselves are composed of glucose monomers. The major constituent of starch, amylopectin is a highly branched polymer consisting of linear chains of α-1,4-linked glucose monomers linked through α-1,6-linked branchpoints. Amylopectin appears to largely dictate the semi-crystalline structure of starch granules, with formation of helices from parallel adjacent linear chains which pack to give concentric ordered, crystalline regions interspaced by amorphous regions where branchpoints of the linear chains disrupt this order. In contrast, amylose consists of mostly unbranched chains of α-1,4-linked glucose molecules which are thought to predominantly exist in the amorphous regions of the granule (Smith and Zeeman, 2020). The structures of these α-glucans are illustrated in Figure 1.3.

In the leaves of plants, transitory starch is synthesised in the chloroplasts during the day using glucose 1-phosphate (G1P) derived from fructose-6-phosphate (F6P) produced by the Calvin-Benson cycle of photosynthesis, and functions as a source of energy to be used via starch degradation during the night when synthesis of sugars by photosynthesis is not possible (Smith and Zeeman, 2020). Remarkably, the near-linear synthesis and degradation of starch adapts to the photoperiod such that the carbon available for growth is maximised. During the day, leaves synthesise only sufficient starch to finish the night with a very low starch content, with degradation rates based on the time until the anticipated dawn such that growth is maximised while periods of starvation are avoided (Graf et al., 2010). Additionally, plants accumulate starch for longer term storage in specialised, non-photosynthetic plastids called amyloplasts which are found in roots, seeds/grains, and storage tissues (such as tubers). The semi-crystalline starch granules found in different plastids, tissues, and of different species vary considerably in their morphology. The granules of most relevance to this work are the transitory starch granules of Arabidopsis leaves, which are lenticular in shape and comprise around 5-10% amylose (Zeeman et al., 2002), with around five to seven starch granules forming within each chloroplast (Crumpton-Taylor et al., 2012). Mutations in starch-related genes often correspond with aberrant granule composition, morphology, size, and/or number of granules per chloroplast, though the mechanisms by which these genotypes dictate their corresponding phenotypes remain nebulous (Liu et al., 2021).

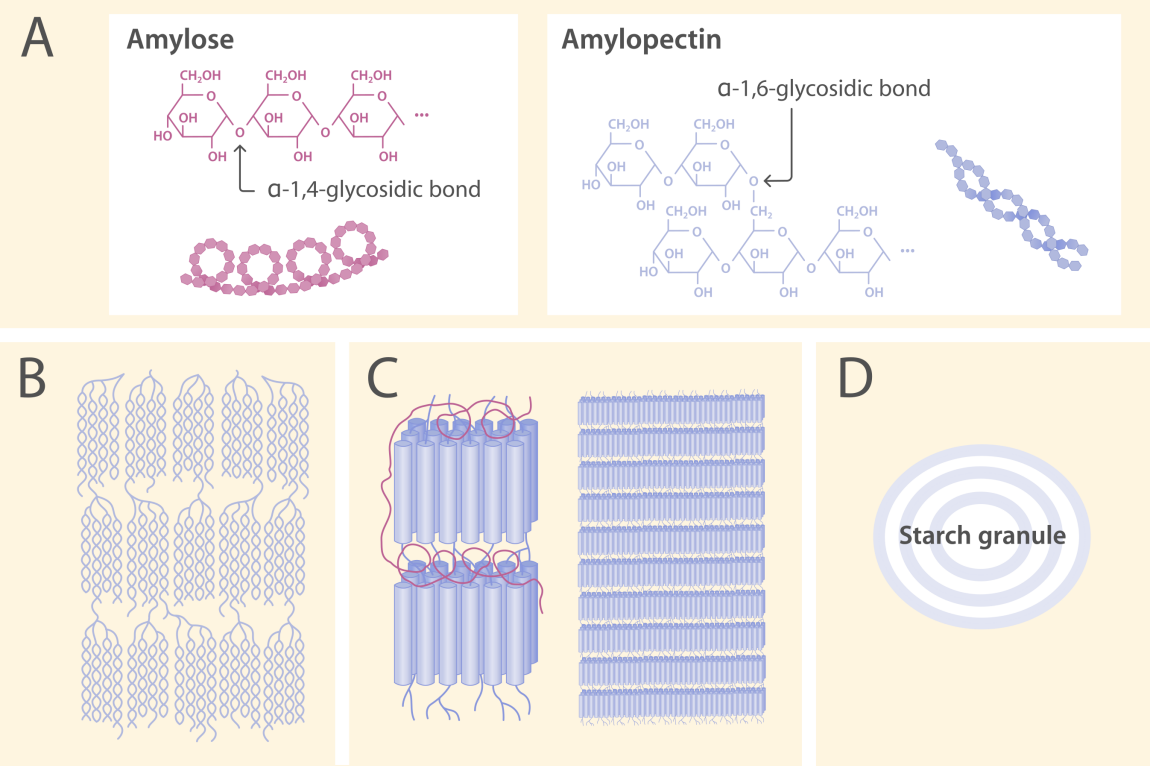


Figure 1.3: **Schematic depicting the structure of starch.** **(A)** Structure of amylose and amylopectin. **(B)** Expanded view of an amylopectin molecule in which linear regions form double helices. **(C)** The semi-crystalline structure of starch comprising crystalline regions formed from packing of numerous amylopectin helices (blue), and amorphous regions where amylopectin α-1,6 branchpoints and amylose molecules (pink) are present. **(D)** A starch granule. Radial arrangement of the semi-crystalline starch structure forming growth rings around a central hilum, shown by concentric white and blue rings. This schematic is based on a figure from Smith and Zeeman, 2020.

1.7.1 Starch synthesis and degradation

Starch synthases are responsible for the elongation of glucan chains using ADP-glucose (ADPG) as a glucosyl donor (Pfister and Zeeman, 2016). ADPG is generated from F6P, a product of the Calvin cycle, via the stepwise action of enzymes phosphoglucose isomerase (PGI), phosphoglucomutase (PGM), and ADP-glucose pyrophosphorylase (AGPase). As illustrated in **Figure 1.4 A**, these enzymes convert F6P to ADPG. ADPG acts as a glucosyl donor for the growth of pre-formed starch polymers, and the generation of short maltooligosaccharides (MOS) which form the substrate of starch granule initiation (Pfister and Zeeman, 2016). Synthesis of MOS in plastids in which starch granules have not already formed may require maltose as the initial glucosyl acceptor, but the mechanism for de novo MOS synthesis is not currently known (Seung and Smith, 2019).

Starch biosynthesis involves the concerted action of three classes of enzymes: starch synthases (SSs), starch branching enzymes (SBEs), and debranching enzymes (DBEs). Amylopectin biosynthesis involves SS1, SS2 and SS3 (Zhang et al., 2008), and, as amylopectin chains are

elongated by SSs, the branched structure of amylopectin is generated through the action of SBEs. DBEs including isoamylases cleave branch points to generate the normal amylopectin structure, as illustrated in **Figure 1.4 C** (Zeeman et al., 2010). GRANULE BOUND STARCH SYNTHASE (GBSS) is responsible for the processive elongation of amylose chains. PROTEIN TARGETING TO STARCH 1 (PTST1) is also required for amylose biosynthesis, and directly interacts with GBSS, functioning to recruit GBSS to starch, with which PTST1 interacts via a carbohydrate-binding module (CBM) (**Figure 1.4 D**, Seung et al., 2015). The vast majority of GBSS is granule-associated (Mu-Forster et al., 1996); as the starch granule grows, GBSS which was surface-bound becomes internalised into the granule matrix.

Degradation of transitory starch involves the phosphorylation of α-glucan chains by glucan, water dikinase and phosphoglucan, water dikinase (Ritte et al., 2006), followed by hydrolysis of the chains by α-amylases, β-amylases and DBEs, as well as glucan dephosphorylation by phosphoglucan phosphatases. Mutations in some of these degradation-associated genes result in starch excess phenotypes, for example the phosphoglucan phosphatase *STARCH EXCESS 4* (SEX4; Kotting et al., 2009). The ultimate products of starch degradation are primarily glucose and maltose, with MOS being released from starch degradation potentially priming the initiation of further starch granules (Streb and Zeeman, 2012, Zeeman et al., 2010).

1.7.2 Starch granule initiation

While the biochemical steps required for the synthesis of amylose and amylopectin described above are now fairly well characterised, the mechanism by which the starch granule is first initiated remains less well understood and comparatively nuanced. However, as recently reviewed by Seung and Smith, ongoing research in this field has identified a number of proteins and enzymes implicated in granule initiation (Seung and Smith, 2019). Loss of any gene with an integral role in starch granule initiation results in a reduction in the number of starch granules produced. One of the only starch granule initiation proteins with clear catalytic activity, STARCH SYNTHASE 4 (SS4) appears to be critical in the initiation of starch granules, with *ss4* mutants in Arabidopsis showing reduced starch accumulation in young leaves (Crumpton-Taylor et al., 2013) and a decreased number of starch granules per chloroplast (Roldan et al., 2007, Seung et al., 2016). SS4 shows a distinctive localisation within chloroplasts, localising to puncta-like structures, which has been suggested to indicate association with plastoglobules via interaction with fibrillins (Gamez-Arjona et al., 2014). However, a later publication was not able to confirm the association of SS4 with fibrillins (Lundquist et al., 2017). Plastoglobules and fibrillins are associated with thylakoids (Austin et al., 2006, Lundquist et al., 2012), and some aspects of starch granule initiation may be associated at least transiently with the thylakoid. PTST2 and PTST3 have also been implicated in granule

initiation, and to have patchy localisations within the chloroplast when expressed in *N. sylvestris* (Seung et al., 2017) or Arabidopsis, in the case of PTST2 (Seung et al., 2018). PTST2 has been shown to interact with SS4, and may be required for normal SS4 function (Seung et al., 2017).

MAR BINDING FILAMENT-LIKE PROTEIN 1 (MFP1) and MYOSIN-RESEMBLING CHLOROPLAST PROTEIN (MRC, also known as PROTEIN INVOLVED IN STARCH INITIATION, PII1; Vandromme et al., 2019) were also identified as essential for normal granule initiation (Seung et al., 2018). MFP1 is thought to be required to correctly locate PTST2 in association with the thylakoids and has a similar localisation to PTST2 in Arabidopsis. Similarly to SS4, MRC has a punctate localisation within the chloroplast, being localised to fewer, more discrete puncta than PTST2 (Seung et al., 2018, Vandromme et al., 2019). MRC is not thylakoid-associated (Seung et al., 2018) and has been shown to interact directly with SS4 (Vandromme et al., 2019). Further, it has been shown that MRC co-expression may result in a more discrete, less diffuse localisation of SS4 than is seen when SS4 is expressed alone (Chen, 2022). Arabidopsis *mrc* mutants have fewer, larger starch granules per chloroplast than the wild-type (Seung et al., 2018, Vandromme et al., 2019). Based on the observation that MRC does not have a CBM, or any other recognisable domain annotation, it has been proposed that MRC functions in the correct folding or associations of SS4 (Vandromme et al., 2019). Summarising current knowledge of starch granule initiation, Seung and Smith propose these identified granule initiation-related proteins to form a 'granule initial' which enables other starch biosynthesis enzymes to synthesise starch (Seung and Smith, 2019).

More recently an additional starch synthase, SS5, has been implicated in granule initiation by Abt and colleagues (Abt et al., 2020). It was observed that when expressed alone, SS5 frequently displayed a diffuse localisation likely indicative of localisation to the stroma. In contrast, co-expression with MRC resulted in a more frequently punctate localisation of SS5, which co-localised with MRC, supporting the notion of a direct interaction between the two proteins. This interaction of SS5 and MRC was also confirmed by co-immunoprecipitation (Abt et al., 2020). Therefore, SS5 can likely be considered part of the putative starch granule initial complex, with parallels to SS4 in terms of its relationship with MRC. The key proteins of the starch granule initial are illustrated in Figure 1.4 B.

From these granule initials, semicrystalline starch granules grow by the action of starch synthases, branching, and de-branching enzymes highlighted in Section 1.7.1. The growth of lenticular Arabidopsis leaf starch granules is known to be anisotropic, a phenomenon which appears to be dependent on the presence of starch granule initiation protein SS4, potentially indicating the importance of these proteins in the formation of starch granules beyond their initiation (Bürgy et al., 2021).

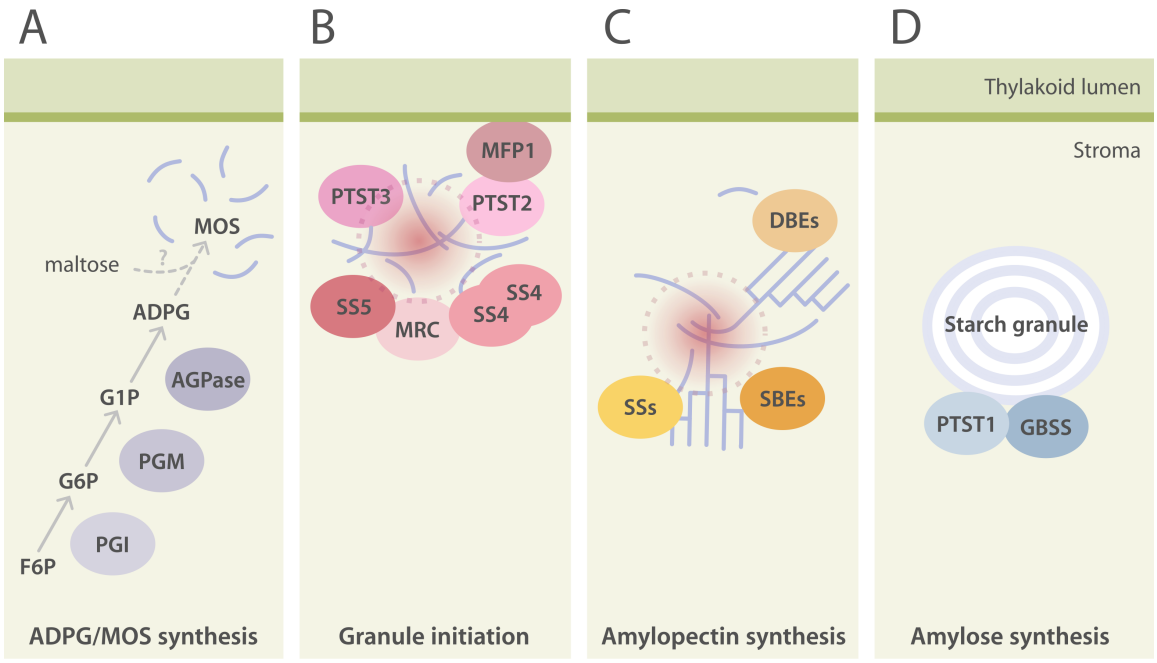


Figure 1.4: **Simplified schematic of starch granule initiation and synthesis in Arabidopsis. (A)** ADPG, the glucosyl donor for starch glucan elongation, is generated from F6P produced by the Calvin cycle. Conversion of F6P occurs by the stepwise action of enzymes PGI, PGM, and AGPase to produce ADPG. MOS are generated in the plastid using ADPG as a glucosyl donor, with maltose possibly being the initial glucosyl acceptor in de novo MOS synthesis (dashed arrow). **(B)** The starch granule initial (red dotted circle) comprises a number of proteins including MRC, SS4, SS5, PTST2, PTST3, and MFP1 (which targets the granule initial to the thylakoid membrane). The action of these proteins allows elongation of MOS. The products of starch granule initiation are glucan chains of sufficient length to be acted on by other enzymes of starch synthesis. **(C)** Amylopectin synthesis involves the function of SSs, SBEs and DBEs. **(D)** Amylose is synthesised by GBSS, and a semi-crystalline starch granule is formed. This schematic is based on a figure from Seung and Smith, 2019.

1.8 Carbon availability during infection: implications for host and pathogen

1.8.1 Host carbohydrates as a resource for pathogens

When colonising the host, a phytopathogen will rely heavily on host metabolism as a source of nutrients, and thus metabolism of the host and pathogen become interlinked (Duan et al., 2013). A striking example of the capacity of pathogens to hijack host processes for carbohydrate acquisition is the exploitation of host *SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTER* (*SWEET*) genes. SWEETs are plasma membrane localised sugar transporters which function in the translocation of sugar molecules unidirectionally out of cells. A number of pathogens secrete TALEs to upregulate certain *SWEET* genes, leading to an increased abundance of SWEETs and amplified supply of sugar to the pathogen, promoting successful colonisation (Chen et al., 2010). Thus, some

SWEETs can be considered susceptibility genes and have been variously engineered to prevent targeting by TALEs, for example through mutation of the effector-binding element of the *SWEET* promoter to prevent TALE binding (Gupta, 2020). The most well studied example of *SWEET* targeting is that of *OsSWEETs* by *Xanthomonas oryzae* pv. *oryzae* effectors (Streubel et al., 2013), but similar instances are also seen in other pathosystems, including the infection of Arabidopsis by *Pseudomonas syringae* pv. *tomato* DC3000, and fungal pathogen *Botrytis cincerea*, exemplifying that this mechanism may not be restricted to biotrophy (Chen et al., 2010, Chong et al., 2014).

The importance of host carbohydrate levels and distribution in the *C. higginsianum*-Arabidopsis pathosystem is highlighted in work by Gebauer and colleagues looking at infection of *sweet11/sweet12* double mutants (Gebauer et al., 2017). These Arabidopsis mutants show an increased soluble sugar and starch content, and increased resistance to *C. higginsianum* during both biotrophic and necrotrophic infection stages. This could suggest that *C. higginsianum* benefits from host sugar exporters for nutrient uptake during infection. These mutants display an increased SA response shortly after pathogen challenge, and SA-deficient *sweet11/sweet12/sid2* triple mutants show an enhanced susceptibility to *C. higginsianum*, comparable to a *sid2* single mutant. The authors conclude that the resistance seen in *sweet11/sweet12* lines is SA-dependent, implicating SWEET11/SWEET12 in SA-mediated defence responses (Gebauer et al., 2017). However, it seems possible that the hyper-susceptibility conferred by *sid2* is simply too great for any theoretically unrelated resistance benefit of *sweet11/sweet12* to show in *sweet11/sweet12/sid2* lines. During the necrotrophic stage of Arabidopsis infection with *C. higginsianum* (four days post infection), *SWEET12* is strongly upregulated proximal to the infection site (Gebauer et al., 2017). Thus, some of the loss of susceptibility in *sweet11/sweet12* lines may be due to the pathogen access to sugars being limited, an increased SA defence response, or aspects of both of these potential mechanisms.

Engelsdorf and colleagues investigated the role of carbohydrate availability on *C. higginsianum* infection of Arabidopsis (Engelsdorf et al., 2013). They showed that low host leaf soluble sugar availability may correlate with an enhanced rate of infection in this pathosystem (Engelsdorf et al., 2013). Availability of host carbon may be either positively or negatively correlated with infection depending on the lifestyle of the pathogen. While starch-free Arabidopsis mutants have been seen to be more susceptible than the wild-type to infection by hemibiotroph *C. higginsianum*, the same mutants have shown an increased resistance to the biotroph *E. cruciferarum* (Engelsdorf et al., 2013). As mentioned in Section 1.3.2, it is possible that while biotrophic pathogens rely on the host for carbon uptake during the biotrophic phase, this may be less critical for some hemibiotrophs. Further work by Engelsdorf *et al*. implicates alterations in the host cell wall composition of Arabidopsis starch mutant lines in their hypersusceptibility to *C. higginsianum*, pointing to a link between carbohydrate availability and penetration resistance (Engelsdorf et al., 2017).

Further, the seemingly counterintuitive increase of hemibiotroph virulence in the absence of host sugars may be reconciled in the context of sugar signalling and priming. Sugars play various roles as signals within plants and have been implicated as key signals in the defence response, regulating phytohormones, including in an infection context (Formela-Luboinska et al., 2020). To exemplify the impact of sugars on defence signalling, Gómez-Ariza and colleagues demonstrated priming of the defence response by application of exogenous sucrose to rice plants pre-infection, resulting in increased resistance to *M. oryzae* (Gomez-Ariza et al., 2007). This highlights the complexity in disentangling host sugar signalling in defence responses benefiting plant resistance, perturbations in host carbohydrates altering cell wall composition and therefore physical defence against pathogen penetration, and increased carbohydrate availability to the pathogen facilitating virulence. Thus, the potential effects of altered host carbohydrate availability or distribution on infection outcomes may be complex and antagonistic.

1.8.2 Starch hyperaccumulates during infection in some pathosystems

In some pathosystems, infection correlates with a hyperaccumulation of starch. For example, the infection of citrus plants by the phloem-limited bacterium *Candidatus* Liberibacter asiaticus causes the disease huanglongbing (HLB), one of the major symptoms of which is the overaccumulation of starch in leaves. HLB infection also correlates with inhibited growth of sink tissues and disruption of phloem function (Keeley et al., 2022). While the hyperaccumulation of starch in HLB has been linked to a specific effector, Las5315, the mechanisms by which this effector function are unclear (Pitino et al., 2018). It is possible that, as opposed to providing an increase in accessible carbon for the pathogen, the starch phenotype seen during this infection process is a side effect of the isolation of source tissues. Overaccumulation of soluble sugars in leaves would be damaging, and so their conversion to insoluble starch, which subsequently hyperaccumulates, may represent a host osmotic stress response. While the starch granules observed in citrus with HLB have morphological similarities to granules formed by phloem blockage by mechanical injury outside of an infection context, they are seen to have subtle differences, such as increased amylopectin chain lengths (Gonzalez et al., 2012).

In gall formation, pathogens have been suggested to manipulate host starch production. *A. tumefaciens* colonisation has been shown to reprogramme host cellular metabolism, causing the accumulation of *S. lycopersicum* metabolites which the bacterium may utilise (Gonzalez-Mula et al., 2018). Galls are commonly sites of starch accumulation (Murakami et al., 2021, Jankiewicz et al., 2021). As an example, *Spongospora subterranean* is a soilborne protist that infects potato. As part of the infection process, *S. subterranean* induces the formation of galls on potato roots. Recent research by Kamal and colleagues highlights the importance of starch in the processes of

gall formation and development (Kamal et al., 2023), with the pathogen potentially using starch granules as a carbon source for formation of sporosori. While amylopectin-enriched granules accumulated early in gall development, starch then decreased later in development despite an upregulation of a number of starch biosynthesis genes during infection (Kamal et al., 2023).

Further, during infection of plants by viruses, differential starch accumulation is seen, and carbohydrate metabolism can provide an avenue from which to reduce host susceptibility to viral infection (Zhao et al., 2022, Tecsi et al., 1992, Handford and Carr, 2007). It is possible that analysing starch granule phenotypes seen due to infection could provide an additional avenue for exploration beyond the natural variation of starch granules seen in uninfected plants. Thus, starch, and potentially the initiation of starch granules, has wide-ranging implications in plant-pathogen interactions more broadly.

1.9 Putative effectors ChEC153 and HaRxL94b localise to chloroplastic puncta

This project focusses specifically on two phytopathogen putative effector proteins, namely ChEC153 (*C. higginsianum* effector candidate 153) from *C. higginsianum*, and HaRxL94b from *H. arabidopsidis*. These putative effector proteins became of particular interest based on their similar, distinctive subcellular localisations revealed in screens carried out in the Faulkner lab (Ohtsu et al., 2023, Liu et al., 2022).

1.9.1 *C. higginsianum* putative effector ChEC153

Drs Mina Ohtsu and Joanna Jennings carried out a screen of putative *C. higginsianum* effectors in the Faulkner lab to identify those which are cell-to-cell mobile (Ohtsu et al., 2023). Alongside this, the screen revealed a number of putative effectors with otherwise interesting subcellular localisations, one of which became the focus of this work. To identify putative effectors, they mined published transcriptome data (O'Connell et al., 2012, Dallery et al., 2017), first excluding any genes that encoded proteins without signal peptides for secretion from the fungus as predicted using SignalP 4.1 (Petersen et al., 2011). Further selection was applied by excluding genes encoding proteins predicted with PredGPI to be membrane-associated (Pierleoni et al., 2008) or to have transmembrane domains. Candidates were classified as potential effector genes based on increased expression in *in planta* appressoria (PA) and in the biotrophic phase of infection (BP) relative to both in *in vitro* appressoria (VA) and in the necrotrophic infection phase (NP). They thereby obtained effector candidates by selecting for $\log_2$ expression ratios for PA/VA, BP/VA, PA/NP, and BP/NP of greater than two using publicly available RNA-Seq data (O'Connell et al., 2012, Dallery et al., 2017). Selected effector candidate coding sequences were synthesised and cloned

with C-terminal GFP tags for localisation via *Agrobacterium*-mediated transient expression in *N. benthamiana* (Jennings, 2021, Ohtsu et al., 2023).

One of these proteins, ChEC153 (OBR05549; locus ID: *CH63R_12252*, previous gene ID: *CH063_01906*), was observed to have a striking subcellular localisation and thus became one of the two putative effector proteins on which this project focuses (Jennings, 2021). *ChEC153* is expressed in early infection, according to published transcriptomic data (O'Connell et al., 2012, Dallery et al., 2017), with the relevant expression values for *ChEC153* reported in Table 1.1. This gene is not expressed in *in vitro* appressoria, nor during the necrotrophic phase, but is highly upregulated in *in planta* appressoria and during the biotrophic phase.

Table 1.1: **Expression of *ChEC153* during different stages of infection.** VA: *in vitro* appressoria; PA: *in planta* appressoria; BP: biotrophic phase; NP: necrotrophic phase. Normalised read counts and $\log_2$ expression ratios reported from O'Connell *et al.* 2012.

Normalised read counts				Expression ratio ($\log_2$)			
VA	PA	BP	NP	PA/VA	BP/VA	PA/NP	BP/NP
1	1357	804	0	11.31	10.56	12.12	11.37

ChEC153-eGFP was seen to localise to multiple spatially separated regions within the chloroplast during *Agrobacterium*-mediated transient expression in *N. benthamiana* (Jennings, 2021, Ohtsu et al., 2023). I repeated this localisation experiment and present representative confocal microscopy images here for illustrative purposes (Figure 1.5). ChEC153-eGFP localises to a number of discrete, punctate foci within the chloroplasts of each transformed cell, regions which are hereafter generally referred to simply as puncta.

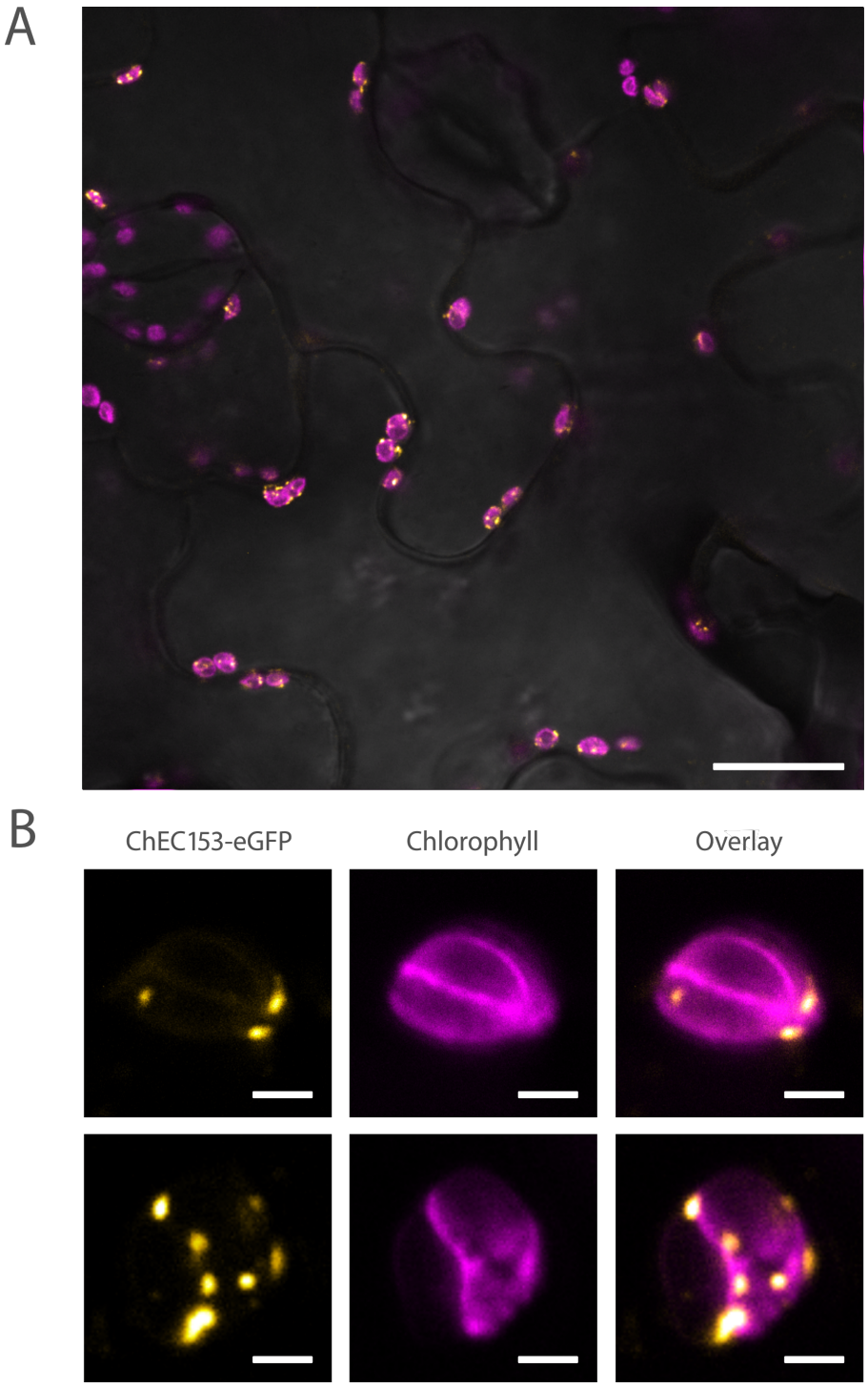


Figure 1.5: **ChEC153 localises to a number of discrete puncta within chloroplasts.** Confocal images showing *35S::ChEC153-eGFP* expression in epidermal cells during *Agrobacterium*-mediated transient expression in *N. benthamiana*. eGFP signal is shown in yellow, with chlorophyll autofluorescence in magenta. Dark, shadow-like regions in the chlorophyll autofluorescence are starch granules. **(A)** Image showing several *N. benthamiana* cells. Scale bars, 20 µm. **(B)** Images of individual chloroplasts at a higher magnification. Scale bars, 2 µm.

1.9.2 *H. arabidopsidis* putative effector HaRxL94b

HaRxL94b was identified as a putative effector protein of *Hpa* by Asai and colleagues in 2014 (Asai et al., 2014). Effector candidates were categorised as summarised: HaRxLs (high-confidence effector candidates), HaRxLLs (RxLR-like candidates with non-canonical features), HaRxLCRNs (Crinkler-homologous genes with RxLR motifs), and those with 5' amino acid sequence similarity including the RxLR motif and signal peptide, e.g., HaRxL1b. In this work, the authors produced transcriptomic data showing the expression of these putative effectors during infection of Arabidopsis Col-0 by avirulent and virulent *Hpa* isolates Emoy2 and Waco9, respectively (Asai et al., 2014). Dr Xiaokun Liu carried out an effector screen in the Faulkner lab to deduce the cell-to-cell mobility of the 87 induced *Hpa* effectors identified by Asai and colleagues (Liu et al., 2022).

Published expression data for one of these putative *Hpa* effector genes, *HaRxL94b,* is recounted here in **Table 1.2**. *HaRxL94b* was categorised as an induced effector gene based on its above twofold induction at three days post inoculation of Col-0 relative to expression in conidiospores for the virulent isolate Waco9 (Asai et al., 2014). Expression of *HaRxL94b* in the avirulent isolate Emoy2 appears restricted to conidiospores, with no expression detected one day after inoculation. In contrast, *HaRxL94b* expression was detected at every timepoint sampled in the infection with Waco9, peaking at three days post inoculation.

Table 1.2: **Expression of *HaRxL94b* during infection.** Expression levels are listed as TPM (tags per million) for Arabidopsis Col-0 infected with *Hpa* isolates Emoy2 and Waco9 at 1-, 3-, or 5-days post inoculation (dpi), or in conidiospores (cs). Data from Asai *et al.* 2014.

Gene	Expression level (TPM)					
	Emoy2		Waco9			
	cs	1 dpi	cs	1 dpi	3 dpi	5 dpi
HaRxL94b	17.418	0	15.477	23.245	32.038	5.132

Upon transient expression in *N. benthamiana*, HaRxL94b-eGFP was seen by Dr Xiaokun Liu to localise to distinct foci within chloroplasts. I repeated this localisation experiment and present representative confocal microscopy images here for illustrative purposes (**Figure 1.6**). Alongside the punctate chloroplastic localisation, some localisation of HaRxL94b-eGFP to the nucleus is also seen indicating dual-targeting by this putative effector (**Figure 1.6 A**, white arrowhead). Despite not being an obvious candidate for cell-to-cell mobility, HaRxL94b remained of interest due to the sub-chloroplastic localisation observed, and its marked similarity to that of ChEC153. These two putative effector proteins and their punctate localisation thereby form the basis of this project.

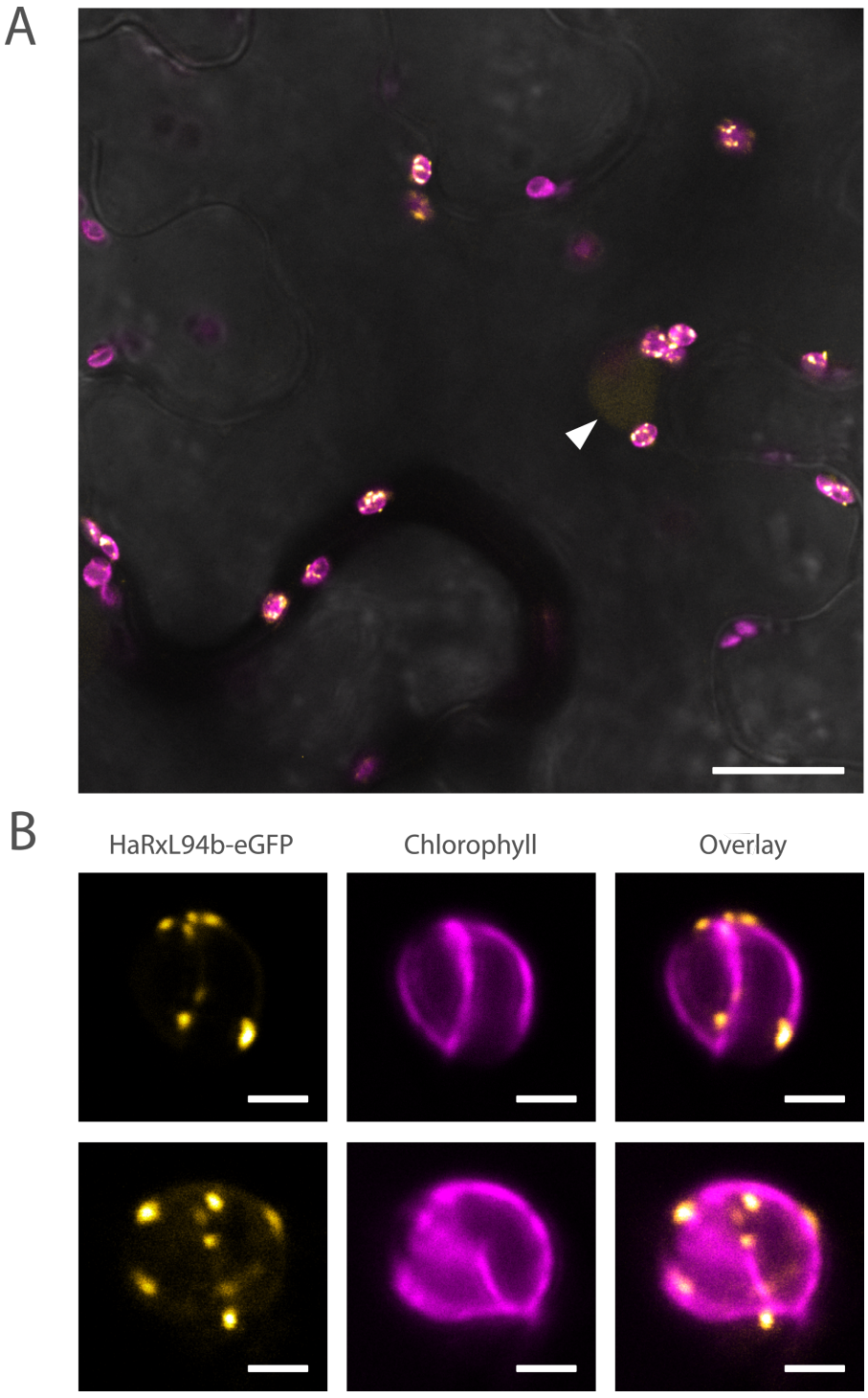


Figure 1.6: **HaRxL94b localises to a number of discrete puncta within chloroplasts, as well as to the nucleus.** Confocal images showing *35S::HaRxL94b-eGFP* expression in epidermal cells during *Agrobacterium*-mediated transient expression in *N. benthamiana*. eGFP signal is shown in yellow, with chlorophyll autofluorescence in magenta. Dark, shadow-like regions in the chlorophyll autofluorescence are starch granules. **(A)** Image showing several *N. benthamiana* cells. Scale bar, 20 µm. White arrowhead indicates HaRxL94b-eGFP localisation to the nucleus. **(B)** Images of individual chloroplasts at a higher magnification. Scale bars, 2 µm.

1.10 Proteins that form distinct, chloroplastic puncta

Alongside some of the starch granule initiation proteins introduced in Section 1.7.2, a number of proteins that localise to puncta within chloroplasts have been identified. Which of these sets of puncta are disparate, or whether multiple of these proteins/processes are spatially linked to one another remains largely undetermined.

1.10.1 Nucleoid-associated proteins and plastidial transcription machinery

Nucleoids are the DNA-containing regions of the chloroplast and resemble puncta when stained with 4′,6-diamidino-2-phenylindole (DAPI; Coleman, 1979) or marked with PLASTID ENVELOPE DNA-BINDING PROTEIN (PEND)-fluorophore fusion proteins (Terasawa and Sato, 2005). The number and morphology of nucleoids appears to change during chloroplast development and division (Kuroiwa et al., 1981), which has more recently been imaged in *Chlamydomonas reinhardtii* with a HEAT UNSTABLE -YFP reporter employed to mark nucleoids (Kamimura et al., 2018). Proteins which localise to nucleoids, such as MutS Homolog1 therefore display a punctate localisation that correlates with DAPI-staining and co-localises with PEND (Xu et al., 2011, Virdi et al., 2016).

Plastid-encoded genes are transcribed by both nuclear-encoded RNA polymerases and a plastid-encoded RNA polymerase (PEP). At least 12 nuclear-encoded PEP-associated proteins (PAPs) have been identified as part of the PEP complex. Some of these PAPs have been identified as potentially nucleoid-associated, such as PAP4, also known as IRON SUPEROXIDE DISMUTASE 3 (FSD3), which may form a complex with PAP9/FSD2, and localises to chloroplastic puncta that correlate with PEND-marked nucleoids (Myouga et al., 2008). Additionally, PEP-complex proteins FRUCTOKINASE-LIKE PROTEIN 1 (FLN1, also known as PAP6) and FLN2 are seen to localise to punctate foci in the chloroplast, co-localising with PEND and thereby being suggested to be associated with nucleoids (Arsova et al., 2010). Further, PAP1 (Yagi et al., 2012), PAP2 (Koussevitzky et al., 2007), and PAP3 (Wang et al., 2023c) have also been associated with punctate chloroplastic localisations. Furthermore, other PEP complex proteins such as PLASTID TRANSCRIPTIONALLY ACTIVE CHROMOSOME 5 (pTAC5; Zhong et al., 2013), pTAC13 (Xiong et al., 2022), GENOMES UNCOUPLED 1 (GUN1, Koussevitzky et al., 2007), MITOCHONDRIAL TRANSCRIPTION TERMINATION FACTOR 3 (mTERF3; Jiang et al., 2020), and mTERF9 (Méteignier et al., 2021) also show similar punctate localisations.

Another protein which may form part of the PEP complex is an RNA-binding protein CHLOROPLAST STEM-LOOP BINDING PROTEIN OF 41 KDA (CSP41). The wheat ortholog (TaCSP41a) was implicated as a *Pst* susceptibility factor in wheat, and localises to striking chloroplastic foci when transiently

expressed in *N. benthamiana* (Corredor-Moreno et al., 2022). CSP41a stabilises RNAs and promotes chloroplastic transcription and translation (Bollenbach et al., 2009). CSP41a has also been seen to be enriched in nucleoids (Majeran et al., 2012), but its putative association with the PEP complex is not yet clear (Leister, 2014). Furthermore, another RNA-binding protein, DEVH-type RNA-helicase INCREASED SIZE EXCLUSION LIMIT 2 (ISE2) which is implicated in plasmodesmal regulation, appears at chloroplast-localised foci (Burch-Smith et al., 2011). These puncta are described elsewhere as being cytoplasmic granule-like structures which appear proximal to chloroplasts (Kobayashi et al., 2007), but their spatial separation from the chloroplast is unclear. While a role of ISE2 in the PEP complex is not proven, cells with lowered ISE2 levels have been reported to show transcriptional changes similar to PEP mutants, and ISE2 appears to play a key role in RNA metabolism whether acting directly in transcription or post-transcriptionally (Bobik et al., 2017).

1.10.2 Punctate proteins implicated in infection or immunity

In addition to TaCSP41a, several other proteins connected to infection, whether as potential pathogen effectors or host susceptibility factors, have been seen in the literature to localise to puncta within chloroplasts. An effector from *Puccina striiformis* f. sp. *tritici* (*Pst),* CHLOROPLAST TARGETING EFFECTOR 1 (PstCTE1, also known as PstHa12j12), localises to chloroplastic puncta despite having no predicted chloroplast transit peptide (Andac et al., 2020). Similarly, Petre *et al*. showed that CHLOROPLAST-TARGETED PROTEIN 3 (CTP3, Melli_sc2834) a *Melampsora lini* effector protein, also localises to puncta within the chloroplast, and additionally showed co-localisation of CTP3 with *Populus trichocarpa* COPROPORPHYRINOGEN III OXIDASE (PtCPO) which they defined as a "marker for undetermined discrete chloroplast bodies" (Petre et al., 2016a, Lorrain et al., 2014). CPO is an enzyme involved in tetrapyrrole biosynthesis and has been associated with a potential role in susceptibility in an SA-dependent manner, with plants in which CPO is mutated having enhanced resistance to *H. arabidopsidis* (Guo et al., 2013).

1.10.3 Plastid division machinery

A key component of the chloroplast division machinery, MinD, has also been seen to localise to puncta, with a point mutation being sufficient to alter the number of puncta forming (Zhang et al., 2021b). Similarly, the Arabidopsis PARALOG OF ARC6 (PARC6) appears as single or several foci throughout chloroplast division (Ishikawa et al., 2020). The punctate localisation of these chloroplast division proteins may reflect the need for polarity determination in the division process. However, TaPARC6 has also been seen to form numerous puncta irregularly spaced within

the chloroplast toward the periphery (Esch et al., 2023), with greater resemblance to the localisation of starch granule initiation proteins.

1.10.4 Other chloroplastic punctate plant proteins

Plastoglobules are sub-chloroplastic lipoprotein bodies which are associated with the thylakoid membrane (Austin et al., 2006) and contain structural fibrillin proteins (Ytterberg et al., 2006). A number of puncta-forming proteins have been theorised as being plastoglobule-associated, such as VARIEGATED 3 (Naested et al., 2004, Ytterberg et al., 2006). Plastid PYRUVATE DEHYDROGENASE COMPLEX E1 COMPONENT SUBUNIT α1, which is involved in galactolipid biosynthesis, also localises to puncta within the chloroplast (Lei et al., 2022). While the authors do not attribute this localisation to plastoglobules or any other punctate structure, given its association with galactolipid biosynthesis, PYRUVATE DEHYDROGENASE COMPLEX E1 COMPONENT SUBUNIT α1 may be plastoglobule-associated. Also, AtPII-GFP localises to distinct puncta referred to as PII foci in Arabidopsis and *N. benthamiana* chloroplasts, but the physiological role of PII in plants is not yet understood (Krieger et al., 2021).

1.10.5 Punctate chloroplastic proteins in *Chlamydomonas*

Besides plant proteins, Wang *et al.* identified 581 *Chlamydomonas reinhardtii* proteins in total localising to the chloroplast, of which 18 localised to chloroplastic puncta. They categorised these punctate chloroplastic proteins as corresponding to 13 separate chloroplastic punctate structures based on the number, size, and positioning of the foci, and consider these punctate structures to be functionally specialised (Wang et al., 2023c). It may be assumed that a number of distinct sets of punctate structures also exist within the plant chloroplast.

1.10.6 Summary

In summary, while the nature of many of these chloroplastic puncta remains elusive, proteins localising to puncta have been implicated in a number of processes in addition to starch granule initiation, including infection and immunity. It is not clear whether some of these localisations represent the same targeting as one another, or whether each of these sets of puncta are both spatially and functionally discrete.

1.11 Project aims

ChEC153 and HaRxL94b were identified in the Faulkner lab as putative effector proteins, meaning that they are expected to contribute to the virulence of their respective pathogens in some way. Despite being predicted to be produced by distantly related pathogens, ChEC153 and HaRxL94b display a remarkably similar localisation in *N. benthamiana*, resembling that of key starch granule initiation proteins. The main aim of this project was to characterise these putative effector proteins, and to investigate the nature of their punctate localisation. Further, this work aimed to evaluate whether targeting of starch granule initiation by filamentous pathogens may present a route to avail them of greater carbon accessibility. A particular emphasis was put on the *C. higginsianum* putative effector ChEC153 due to the genetic tractability and ease of axenic culturing of the pathogen. Toward the end of the project, new questions were raised regarding the *ChEC153* gene model, its classification as an effector, and its role in the host or pathogen, which I seek to address in my final results chapter.

CHAPTER 2

Materials & Methods

This chapter contains information pertaining to general materials and methods used throughout this work. Information regarding more specialised materials and methods can be found within the relevant results chapters.

2.1 Arabidopsis mutant and transgenic lines

Table 2.1: Arabidopsis lines used.

Line	Gene ID	Acquired from	Reference
Col-0	N/A	J. Jennings (JIC)	N/A
35S::ChEC153-eGFP (T3)	*CH63R_12252*	J. Jennings (JIC)	Jennings, 2021
mrc-3 (SAIL_1151_E06)	*AT4G32190*	D. Seung (JIC)	Seung et al., 2018
ss4-1 (GABI_290D11)	*AT4G18240*	D. Seung (JIC)	Roldan et al., 2007
pgm	*AT5G51820*	D. Seung (JIC)	Caspar et al., 1985
SALK_099429C (*crbic*)	*AT1G53120*	NASC (N661044)	Alonso et al., 2003
SALK_041981C	*AT1G53120*	NASC (N669607)	Alonso et al., 2003

2.2 Plant growth

Arabidopsis plants were grown on Levington F2 Starter (SCOTTS) soil. Seeds were stratified at 4 °C for three days before being transferred to short-day (10 h: 14 h, light: dark) conditions at 22 °C. For growth on plates, seeds were sterilised in 10% (v/v) sodium hypochlorite solution for ten minutes after washing in 70% ethanol for one minute. Seeds were then rinsed in sterile water three times prior to plating on 0.8% agar, 1% sucrose Murashige and Skoog (MS) plates, with 10 µg/mL phosphinothricin where appropriate (*35S::ChEC153-eGFP* lines). Following stratification, plates were transferred to short-day growth cabinets (10 h: 14 h, light: dark; 22°C). Two-week-old seedlings were then transferred to soil for continued growth.

N. benthamiana plants were grown under long-day (16 h: 8 h, light: dark) conditions in a controlled environment room at 22 °C, with 80% relative humidity.

2.3 Bacterial strains and growth conditions

E. coli strain DH5α was grown at 37 °C, with shaking at 200 rpm for liquid cultures. *Agrobacterium tumefaciens* strain GV3101 was grown at 28 °C, with shaking at 220 rpm for liquid cultures. In the case of fungal transformation, *Agrobacterium* strain AGL1 was used in lieu of GV3101, as described in Section 4.2.3.2.

Glycerol stocks of bacterial strains (20% glycerol, made from dense overnight cultures of single colony inoculations) were stored at -70 °C. Recovery of glycerol stocks was carried out by streaking stocks onto LB (lysogeny broth) plates containing appropriate antibiotics and incubating until colonies were seen. Sterile toothpicks were then used to inoculate 10 mL liquid cultures (LB + appropriate antibiotics) with single, isolated bacterial colonies. *E. coli* cultures were then incubated at 37 °C overnight, and *Agrobacteria* cultures at 28 °C for approximately 36 hours.

Table 2.2: Concentrations of antibiotics used for bacterial selection.

Antibiotic		Final working concentration
Spectinomycin	In dH_2O	50 µg/mL
Carbenicillin	In dH_2O	100 µg/mL
Kanamycin	In dH_2O	50 µg/mL
Rifampicin	In DMSO	50 µg/mL
Gentamycin	In dH_2O	10 µg/mL

2.4 Bacterial transformation

Electrocompetent DH5α and GV3101 cells were transformed with approximately 100 ng of purified plasmid DNA by electroporation using a Gene Pulser™ electroporation system (Bio-Rad) at 200 Ω, 25 µF, at a voltage of 2.5 or 1.8 kV, respectively. Electroporated *E. coli* were recovered by outgrowth in 500 µL SOC medium for 1 hour at 37 °C, and *Agrobacterium* for 2 hours at 28 °C. Bacteria were plated on pre-warmed LB agar plates supplemented with appropriate antibiotics (**Table 2.2**) and/or IPTG and X-Gal for blue-white selection (238 µg/mL IPTG; 20 µg/mL X-Gal). Generally, 120 µL of transformation cultures were plated on selective LB plates using a sterile spreader in order to obtain isolated single colonies. Where necessary, the remaining culture was later plated to obtain more colonies, or overgrown bacterial lawns were streaked out to better separate colonies.

2.5 Cloning

The cloning strategy primarily employed in this work was Golden Gate cloning (Engler et al., 2008). This was greatly facilitated by the TSL SynBio platform's repository of Golden Gate backbones and parts, as well as those available to me within the Faulkner lab stocks, which use the standard plant Golden Gate cloning system described by Engler and colleagues (Engler et al., 2014) and Patron and colleagues (Patron et al., 2015). This Golden Gate system employs antibiotic screening of Level 0 assembly transformant colonies using spectinomycin, and of Level 1 assembly transformant colonies with carbenicillin. Additionally, replacement of the *lacZ* cassette in receiver vectors by the desired assembly parts enables blue-white selection.

2.5.1 Amplification reactions

To generate Level 0 parts for alternate assembly positions, or from cDNA, gDNA, or parts of other plasmids, amplification using primers to introduce Golden Gate cloning extensions was carried out. Phusion® High-Fidelity DNA polymerase (NEB, M0530) was used with the Phusion® HF buffer in final reaction volumes of 20 µL, according to the manufacturer's protocols. Standard thermocycling conditions recommended by the manufacturer were used, with 35 cycles and allowing one minute of extension time per kb of expected amplicon. The NEB Tm calculator (https://tmcalculator.neb.com) was used to determine the annealing temperature to be used with each primer pair. Amplicons were typically separated by agarose gel electrophoresis on 1% gels stained with ethidium bromide (EtBr). Desired bands were excised using a UV transilluminator, and DNA extracted using a QIAquick Gel Extraction Kit (QIAGEN, 28704) to provide the input for Golden Gate assembly reactions.

2.5.2 Golden Gate reactions

For Golden Gate cloning, compatible parts were combined at an equimolar concentration (40 fmol), with 2:1 insert:acceptor molar ratios (acceptor vectors: 20 fmol). These parts were added to reaction mixtures containing 1.5 µL 10× BSA, 1.5 µL T4 DNA ligase buffer (NEB, M0202), 0.5 µL T4 DNA ligase (NEB, M0202) and 0.5 µL of the appropriate type IIS restriction enzyme (Level 0 reactions: BpiI-HF (ThermoFisher, ER1011) Level 1 reactions: Bsa1-HF-v2 (NEB, R3733)). Reactions were made to a final volume of 15 µL with dH_2O and the following thermocycler protocol was run: 20 seconds at 37 °C, 30 cycles of 3 minutes at 37 °C then 4 minutes at 16 °C, 5 minutes at 50 °C, 5 minutes at 80 °C, and the reactions then held at 16 °C. Transformation of assembled constructs into DH5α *E. coli* cells was carried out as described in **Section 2.4** using 1 µL of the Golden Gate reaction.

2.5.3 Construct verification

Candidate transformant colonies were first screened by PCR directly from colonies using GoTaq® G2 Green Master Mix (Promega, M7822) with standard primers 0015 and 0016 (Level 0 constructs) or 0299 and 0230 (Level 1 constructs; see **Appendix 1 Tab. E**). Amplicon sizes were verified by electrophoresis of PCR products on 1% agarose TBE + EtBr gels and visualisation with a UV transilluminator. Plasmids were isolated from *E. coli* using a QIAprep Spin Miniprep Kit (QIAGEN, 27106) according to the manufacturer's instructions. Newly assembled Level 0 constructs were sequence-verified using Sanger sequencing across the cloning junctions using primers 0015 and/or 0016. Sanger sequencing was largely carried out using the Mix2Seq kit service provided by Eurofins Genomics. Sequencing results were evaluated by aligning chromatogram sequence traces with templates of expected DNA sequences using the online software Benchling (benchling.com).

Details of plasmids used in this work (generated through this work or acquired from others) are presented in **Appendix 1**, including Level 0 Golden Gate constructs (**Appendix 1 Tab. A**), Golden Gate acceptor plasmids (**Appendix 1 Tab. B**), and Level 1 Golden Gate constructs (**Appendix 1 Tab. C**). Gateway constructs acquired from others are listed in **Appendix 1 Tab. D**. Primers used for cloning and construct-verification can be found in **Appendix 1 Tab. E**. Finally, gBlocks synthesised during this project are detailed in **Appendix 1 Tab. F**.

2.6 *Agrobacterium*-mediated transient transformation of *N. benthamiana*

Single colonies of GV3101 *Agrobacterium tumefaciens* strains carrying vectors for gene expression were used to inoculate 10 mL of liquid media (LB + appropriate antibiotics), with cultures being grown for approximately 48 hours. Cells were then harvested by centrifugation and washed twice

in 10 mM $MgCl_2$ before being resuspended in 10 mM $MgCl_2$ + 100 µM acetosyringone, or infiltration media (10 mM MES, 10 mM $MgCl_2$ + 100 µM acetosyringone). Strains were combined to give a final concentration of each strain at OD_{600nm} = 0.5, in addition to an *Agrobacterium* strain carrying the *p19* silencing suppressor (Win and Kamoun, 2004), also at OD_{600nm} = 0.5. Generally, prepared *Agrobacterium* solutions were then allowed to recover at room temperature for at least one hour prior to infiltration. Approximately 4-week-old *N. benthamiana* plants were infiltrated with a needleless 1 mL syringe and the abaxial side of leaves imaged with confocal microscopy three days after infiltration. This process is illustrated in **Figure 2.1**.

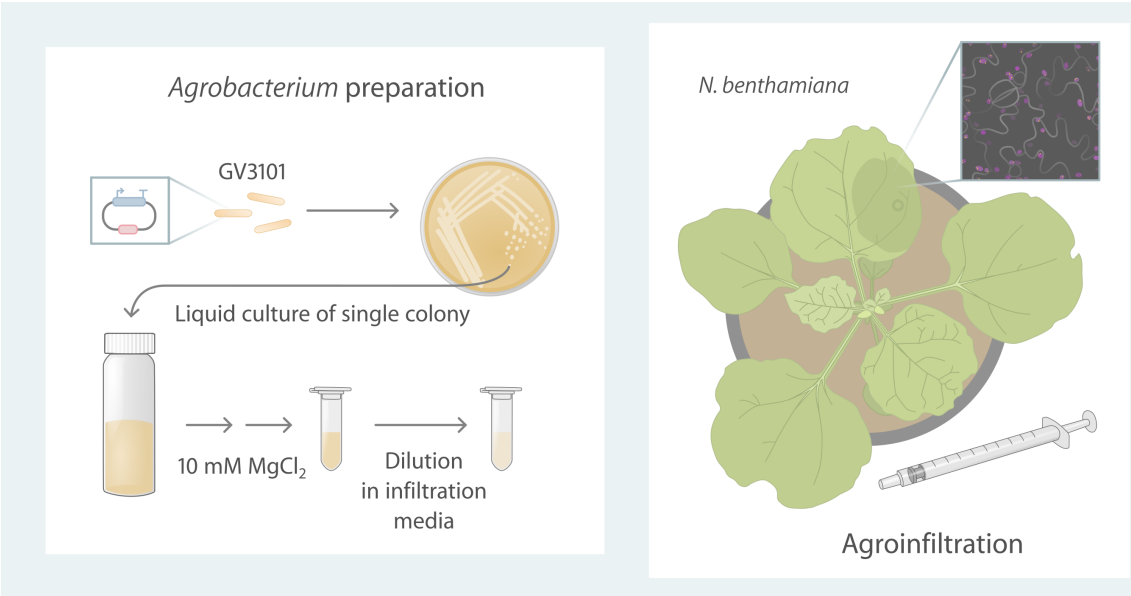


Figure 2.1: Schematic of *Agrobacterium*-meditated transient expression in *N. benthamiana*. Single colonies of *Agrobacteria tumefaciens* transformed with expression vectors are isolated and grown in 10 mL liquid cultures. Cells are harvested by centrifugation and washed twice in 10 mM $MgCl_2$ prior to resuspension and dilution to a given optical density in infiltration media. Leaves of *N. benthamiana* leaves are infiltrated with bacterial suspensions using a needleless syringe. Samples are harvested three days post infiltration, for example for examination with confocal microscopy, among other applications.

2.7 Confocal microscopy

Confocal microscopy was carried out with a Zeiss LSM 800 confocal laser scanning microscope with a 40× water-dipping objective (W PlanApochromat 40×/1.0 DIC VIS-IR M27-water), or 63× water-dipping objective (W Plan-Apochromat 63×/1.0 M27-water). eGFP excitation was with a 488 nm solid-state laser, with collection at 410-530 nm, chlorophyll autofluorescence was detected by excitation at 488 nm and collected at 656-700 nm, and RFP or mCherry excitation was at 587 nm with collection at 585-617 nm.

2.8 Protein purification and Western blotting

Arabidopsis leaf tissue samples were manually homogenised on ice into extraction buffer (10 µL/mg fresh weight; 50 mM Tris-HCl pH 8, 150 mM NaCl, 1% Triton X-100, 1× proteinase inhibitor cocktail, 1 mM dithiothreitol (DTT)). Insoluble material was pelleted by centrifugation at 4 °C at maximum speed in a benchtop microcentrifuge (~21,000 ×g) for 10 minutes and resuspended in 1× sample loading buffer (see Section 2.8.1; to a volume approximately equal to that of the extraction buffer used), while the soluble fraction was transferred to a new tube. Resuspended pellets were boiled at 95 °C for 10 minutes. Soluble fractions were centrifuged for a second time, and again the supernatant was transferred to a new tube. 32 µL of the soluble fraction was retained as the "input" sample and boiled at 95 °C for 10 minutes with 8 µL of 5× loading buffer.

For immunoprecipitation (IP), the remaining soluble fraction was incubated tumbling in a cold room (5 °C) for 2 hours with 25 µL of freshly washed GFP-Trap® Magnetic Agarose beads (ChromoTek, gtma). Beads were then incubated in a magnetic rack, and 32 µL of the supernatant was taken as the "flow through" sample and boiled at 95 °C for 10 minutes with 8 µL of 5× loading buffer. The remaining supernatant was discarded and the beads washed with fresh, cold extraction buffer a total of four times. Following the final wash, all remaining liquid was discarded and 30 µL of extraction buffer and 10 µL of 5× loading buffer were added to the beads, which were then boiled for 15 minutes at 95 °C (IP sample). 10 µL of each sample and 6 µL of PageRuler Plus pre-stained protein ladder (NEB, 26619) were loaded onto a 10% polyacrylamide gels (see Section 2.8.1) for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Separated proteins were then transferred to PVDF membrane (Bio-Rad: Immun-Blot® PVDF membrane) in a wet transfer at 100 V for 1 hour 15 minutes.

To detect GFP-tagged proteins, the primary anti-GFP mouse antibody (Roche, 11814460001) was diluted 1:1000 in 5% milk + TBST and blot was incubated in this solution overnight at 4 °C after blocking for 1 hour at room temperature in 5% milk + TBST. After washing in 5% milk + TBST (3× 5 minutes), the secondary antibody: anti-mouse IgG-Peroxidase (Sigma, A0168) was used at a 1:10,000 dilution with 1.5 hours' incubation at room temperature. To detect GBSS, a primary anti-GBSS rabbit antibody (Seung et al., 2015) was used at a 1:200 dilution in 3% milk + TBST, with incubation overnight at 4 °C. After washing the membrane in 5% milk + TBST (3× 5 minutes), an anti-rabbit IgG-Peroxidase secondary antibody (Sigma, A0545) was used at a 1:20,000 dilution for 2.5 hours' incubation at room temperature.

Following incubation with antibodies, membranes were washed in TBST (3× 5 minutes), and then 200 µL of each HRP substrate (Thermo: SuperSignal™ West Femto kit, 34094) were applied to the membrane and the blot was imaged using an ImageQuant™ LAS 500 imager (GE Healthcare) incrementally until over-exposed.

2.8.1 Western blotting solutions and SDS polyacrylamide gels

For SDS PAGE, 10% polyacrylamide gels were prepared with components listed in **Table 2.3**. Compositions of general buffers used in Western blotting are listed in **Table 2.4**.

Table 2.3: Composition of 10% polyacrylamide gels used for SDS-PAGE. SDS: Sodium dodecyl sulphate; APS: Ammonium persulfate; TEMED: Tetramethylethylenediamine.

Stacking gel		Resolving gel	
Dissolved in dH_2O:		Dissolved in dH_2O:	
Tris-HCl (pH 6.8)	0.13 M	Tris-HCl (pH 8.8)	0.39 M
Acrylamide:bisacrylamide (30%; 37.5:1)	5%	Acrylamide:bisacrylamide (30%; 37.5:1)	10%
SDS	0.1%	SDS	0.1%
APS	0.1%	APS	0.1%
TEMED	0.1%	TEMED	0.1%

Table 2.4: Composition of buffers used for Western blotting

Buffer	Final composition
Sample loading buffer (6×) In dH_2O Used as 5× or 1× diluted from 6× by the addition of water and β-mercaptoethanol	120 mM Tris-HCl, pH 6.8 50% Glycerol 6% SDS 0.01% Bromophenol blue
Running buffer In dH_2O	25 mM Tris 200 mM Glycine 0.1% SDS
Transfer buffer In dH_2O	10 mM Tris 200 mM Glycine 20% Methanol
TBST In dH_2O	20 mM Tris-HCl, pH 7.5 150 mM NaCl 0.1% Tween20

2.9 Statistics, data visualisation, and figure preparation

All statistical analyses were carried out in R Statistical Software (v4.1.1) in RStudio (R Core Team, 2021). Generally, data were analysed using linear and linear mixed effects models with lm() and lmer() functions. For linear mixed effects models comparing two groups, data were analysed using t-tests using Satterthwaite's method, and for linear mixed effects models comparing more than two groups, data were analysed using an ANOVA followed by pairwise comparisons on the estimated marginal means using post hoc Tukey HSD. The emmeans package (Lenth, 2023) was used for pairwise comparisons (emmeans() function), while the lmerTest package was used for the linear mixed effects models (Kuznetsova et al., 2017). Random effects are specified in each experiment. Data were visualised using R, predominantly with the ggplot2 package (Wickham, 2009). Where boxplots are presented, the box shows the first quartile, median (horizontal line), and third quartile. Whiskers of the boxplots indicate the data minimum and maximum (Q1 - 1.5 × interquartile range and Q3 + 1.5 × interquartile range, respectively), with potential outliers shown outside of the whiskers. Where the mean was plotted, this is shown by a grey diamond.

Microscopy images were adjusted for figure preparation in ImageJ (Schindelin et al., 2012). All figures were compiled, and schematics drawn, using Inkscape (Inkscape.org).

Chapter 3

Probing the cell biology of putative effectors ChEC153 and HaRxL94b

3.1 INTRODUCTION

Plant pathogens secrete effector proteins into their hosts and may use these to increase their access to carbohydrates for uptake. As introduced in **Chapter 1**, ChEC153 and HaRxL94b were identified as putative effector proteins from *C. higginsianum* and *H. arabidopsidis*, respectively. Stemming from preliminary work carried out by Dr Joanna Jennings, Dr Mina Ohtsu, Dr Xiaokun Liu, and Dr Andrew Breakspear within the Faulkner lab, the striking localisations of these proteins to punctate foci within the chloroplast (**Figure 1.5** and **Figure 1.6**), reminiscent of the localisation of several starch granule initiation proteins, prompted their further investigation within this project. I sought to test the hypothesis that these effector candidates may be targeting starch granule initiation in the host.

3.1.1 Protein translocation into the chloroplast: a route for pathogen effectors?

In order for pathogen effectors to manipulate the processes of the chloroplast, we would expect many of them to first be imported into the chloroplast, translocating through the chloroplast envelope. The majority of native chloroplastic proteins are produced in the cytosol before being translocated into the chloroplast. Effector proteins present in the host cell cytosol may hijack endogenous mechanisms of chloroplast targeting and import. Proteins destined for the stroma of the chloroplast must first cross the chloroplast outer membrane and inner membrane. Further translocation is required for directing proteins to the thylakoid lumen.

3.1.1.1 Overview of the general import pathway

Following cytosolic translation, chloroplast targeting is generally dictated by an N-terminal sequence of amino acids within the precursor protein (pre-protein) known as a chloroplast transit peptide (cTP). Cytosolic chaperones and chaperone-assisted complexes are believed to aid the trafficking of preproteins with cTPs to the chloroplast surface from the ribosome, preventing their aggregation or degradation and allowing them to maintain an at least partially unfolded, import-competent, conformation prior to translocation (Flores-Pérez and Jarvis, 2013). Translocation of folded proteins may also be possible (Ganesan et al., 2018), and may be of particular relevance in the consideration of translocation of effector proteins which presumably may exist in the cytosol in a folded state prior to targeting to the chloroplast.

Translocation takes place via unidirectional translocation protein complexes at the outer membrane (*Translocon at the Outer envelope membrane of Chloroplasts*; TOC) and inner membrane (*Translocon at the Inner envelope membrane of Chloroplasts*; TIC). The TOC complex

transfers proteins into the intermembrane space, and transfer via the TIC complex into the chloroplast stroma is suggested to take place simultaneously. This is expected to prevent mislocalisation of preproteins to the intermembrane space. Protein import is an energetically costly process, with an estimated hydrolysis of approximately 650 ATP molecules being required for the import of a single pre-protein (Shi and Theg, 2013). While GTPase activity of TOC components is evident and GTP hydrolysis has a role in the formation of early-import intermediates, ATP hydrolysis mediates translocation and is sufficient to drive protein import (Young et al., 1999).

Following translocation into the chloroplast stroma, the cTP is cleaved from pre-proteins by a stromal processing peptidase (SPP) to produce a mature protein (Vandervere et al., 1995, Richter and Lamppa, 1998). Proteins targeting the thylakoid lumen may possess a thylakoid signal peptide which is exposed following cTP cleavage, leading to their targeting to the lumen by the Sec or Tat pathways (reviewed in Albiniak et al., 2012). Proteins imported into the chloroplast may be targeted for insertion into the thylakoid membranes (reviewed by Zhu et al., 2022).

3.1.1.2 Regulation of chloroplast protein import

Import of proteins encoded in the nucleus to the chloroplasts can present an opportunity for regulation of chloroplastic processes, further to NECG transcriptional regulation. Relatively recently, the process of proteolytic degradation of outer envelope membrane proteins such as components of the TOC complex has begun to be characterised, and termed chloroplast-associated protein degradation (CHLORAD, Ling et al., 2019). In this process, the chloroplast-localised E3 ligase SUPPRESSOR OF PPI1 LOCUS 1 guides ubiquitination of outer membrane proteins and then acts in a complex with SUPPRESSOR OF PPI1 LOCUS 2 and CELL DIVISION CYCLE PROTEIN 48 to extricate these proteins from the outer membrane (retrotranslocation) for degradation by the 26S proteasome in the cytosol (Ling et al., 2019). SUPPRESSOR OF PPI1 LOCUS 1 is believed to inhibit import in response to certain abiotic stresses by degradation of TOC machinery (Ling and Jarvis, 2015), and TIC. It is possible that similar phenomena may occur in responses to biotic stresses.

3.1.1.3 Effector translocation into the chloroplast

While some putative effector proteins possess N-terminal chloroplast transit peptides which presumably allow their import into the chloroplast by the general import pathway via TOC/TIC (e.g. CTP1, CTP2, and CTP3 (Petre et al., 2016a)), other effectors seen experimentally to localise to the chloroplast lack clear transit peptides (Sperschneider et al., 2017, Petre et al., 2015, Petre et al., 2016b). This may highlight limitations in our understanding and prediction of chloroplast transit

peptides, or these proteins may use alternative routes to localise to the chloroplast. Several Arabidopsis chloroplast proteins have been described whose targeting of the chloroplast is experimentally confirmed despite their lack of a predicted cTP, however the prevalence of such noncanonical chloroplast proteins remains unclear (Armbruster et al., 2009).

3.1.2 Starch granules in Arabidopsis: wild-type and granule initiation mutants

All of the known starch granule initiation proteins are encoded in the nucleus and translocated into the chloroplast, and have canonical transit peptides (Seung et al., 2018, Seung et al., 2017, Gamez-Arjona et al., 2014, Abt et al., 2020).

In Arabidopsis, transitory starch granules in leaves have an irregular, lenticular morphology and are relatively uniform in size, with wild-type Arabidopsis leaf starch granules being approximately 2 µm in diameter at the end of the day (Seung and Smith, 2019, Zeeman et al., 2002). While leaf starch granules are unimodal (Vandromme et al., 2019), the starch granules seen in starch storage organs of some species are often bimodal or compound (originating from more than one initiation point). As recently summarised by Chen and colleagues, unimodal, bimodal, and compound starch granules are seen in the endosperm of maize, wheat, and rice, respectively, and granules may be round, ellipsoid, or polyhedral (Chen et al., 2021a).

There are numerous Arabidopsis mutants in various starch-related genes. An Arabidopsis mutant in plastid phosphoglucomutase, *pgm*, is essentially starch-free (Caspar et al., 1985). Similarly, the Arabidopsis *adg1* mutant of plastidial ADP-glucose pyrophosphorylase is largely unable to synthesise starch (Lin et al., 1988). Conversely, mutants with impeded starch degradation (Caspar et al., 1991) including mutation in *STARCH EXCESS 4* (*sex4*; Zeeman et al., 1998) have been characterised, with constitutively elevated starch levels. A number of Arabidopsis mutants for starch granule initiation proteins have been characterised. These include mutants in *MRC* and *SS4* genes, both of which share a starch granule phenotype with fewer, larger starch granules per chloroplast than Col-0, though this is more extreme in *ss4* relative to *mrc* lines, with many *ss4* chloroplasts lacking starch granules (Seung et al., 2016). Similar but less extreme phenotypes are seen for leaf starch of *TaMRC* mutants (Chen et al., 2022). The roles of these starch granule initiation proteins in the formation of starch in storage tissues such as grains, however, sometimes contrasts with their roles in leaves, pointing to a diversification of roles for granule initiation proteins (Chen et al., 2022). Additional similar starch granule initiation mutant phenotypes include those of *ptst2* and *ptst3.* Particularly strikingly, Arabidopsis *ptst2* lines generally have a single large starch granule or zero starch granules within a chloroplast, while *ptst3* mutant phenotypes are more subtle, similar to *mrc* phenotypes, with a slightly reduced number of granules per chloroplast

(Seung et al., 2017). Conversely, PTST2 overexpression induces the formation of a large number of small starch granules (Seung et al., 2017). While these phenotypes inform on the fundamentals of starch granule initiation or metabolism, they may also be used as tools with which to explore the impact of varying starch granule size and number per chloroplast, or starch content, on a number of other processes. Within this work the potential for starch granule morphology to affect carbohydrate availability is of particular interest.

3.1.3 Starch-related proteins do not necessarily contain starch-binding or catalytic domains

Proteins associating directly with starch granules are expected to possess carbohydrate binding modules. However, a number of proteins involved in starch granule initiation or starch metabolism are not known to interact directly with starch itself. For example, GBSS is targeted to the starch granule through interaction with PTST1 (Seung et al., 2015). PTST2 is believed to bind glucans as well as MFP1 and MRC, facilitating their roles in granule initiation (Seung et al., 2017). Several proteins involved in starch initiation or metabolism possess no clear catalytic domains, but rather are suspected to play structural roles in the process – for example, MRC is not believed to have either catalytic or starch-binding capacity, but instead may function in correctly positioning the enzyme SS4 within the starch granule initial (Seung et al., 2018, Vandromme et al., 2019). MRC instead possesses long coiled-coil motifs which may facilitate protein-protein interactions at the starch granule initial (Seung et al., 2018). Further, CBMs are sometimes misannotated as RNA-binding domains due to the similarity in the way aromatic rings of proteins stack and interact with the glucose monomers of glucans or ribose molecules of RNA. Thus, prediction of proteins involved in these processes based on analysis of protein sequences or folds is not facile.

3.1.4 Chapter Aims

In this Chapter, I aimed to investigate ChEC153 and HaRxL94b from a cell biology perspective and in the context of starch granule initiation. I analyse the sequences of ChEC153 and HaRxL94b using *in silico* tools, and probe the localisation of these putative effector proteins, particularly in the context of sites of starch granule initiation. Further, I evaluate the impact of the expression of the *C. higginsianum* putative effector ChEC153 on host starch content and granule size.

3.2 MATERIALS & METHODS

3.2.1 Particle bombardment

Plasmids for expression of *35S::ChEC153-eGFP* and *35S::HaRxL94b-eGFP* were isolated from DH5α cells using a QIAGEN plasmid maxi kit (12162) according to the manufacturer's instructions. 8747 ng and 9502 ng respectively of either putative effector plasmid for ChEC153-eGFP and HaRxL94b-eGFP, and 2577 ng of an ER-RFP transformation marker (pB7WG2.0 mRFP::KDEL) were co-precipitated onto 1 µm gold microcarriers (Bio-Rad, 1652263) as described by Tee *et al.* (Tee et al., 2022). Briefly, 2.5 µL of either plasmid was mixed with a 25 µL aliquot of resuspended gold particles and sonicated briefly. 25 µL of 2.5 M $CaCl_2$ was added and samples mixed gently for 10-20 seconds to precipitate the DNA onto the gold particles. 10 µL of 0.1 M spermidine was added, and samples mixed gently for two minutes before being incubated on ice for 10-20 seconds. The gold particles were then washed twice in 100% ethanol by centrifugation and resuspension. Finally, particles were resuspended in 100 µL 100% ethanol and sonicated again before being aliquoted onto macrocarriers (Bio-Rad, 1652335). The abaxial side of Arabidopsis leaves from Col-0, *mrc*, *ss4*, and *pgm* lines (provided by the Seung lab; **Table 2.1**) were bombarded from a distance of 9 cm as described by Tee *et al.* using a Biolistic PDS-1000/He particle delivery system (Bio-Rad, 1652257) or "gene gun" (Tee et al., 2022). Bombardments were carried out at 1100 psi (rupture disks: Bio-Rad, 1652329). Bombarded leaves were imaged 16-20 hours post bombardment using confocal microscopy. A schematic to illustrate the particle bombardment experiment is presented in **Figure 3.1**.

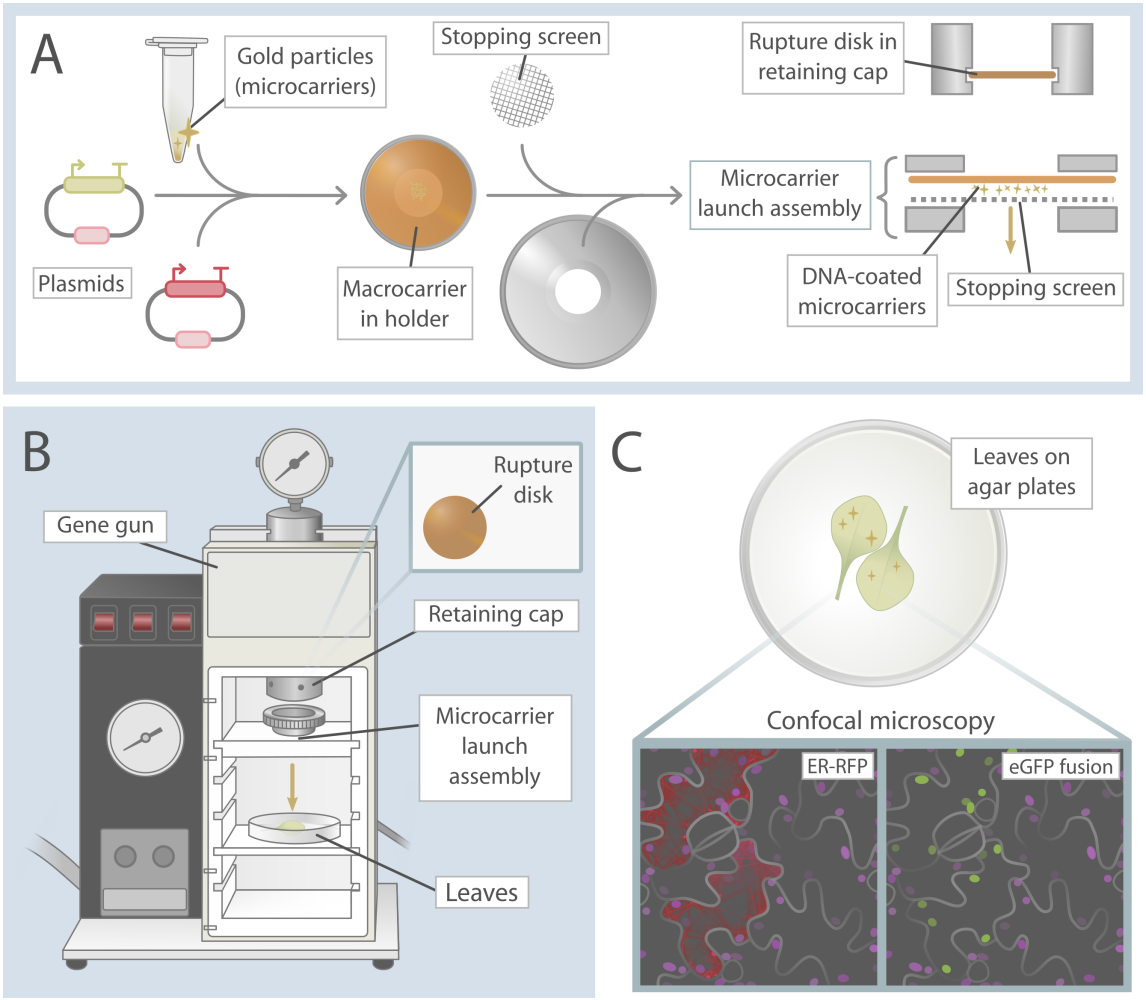


Figure 3.1: Schematic depicting particle bombardment of Arabidopsis leaves. (A) Vectors for high expression of eGFP fusion proteins are isolated at high concentrations and, with an ER-RFP marker, co-precipitated onto 1 µm gold particle microcarriers. DNA-coated microcarriers are aliquoted onto the macrocarrier within the macrocarrier holder. The microcarrier launch assembly is set up such that the microcarriers are facing toward the sample, separated from it only by a stopping screen. **(B)** Leaf samples are arranged on 0.6% MS agar plates with the abaxial side facing upwards. Rupture disks are positioned within the retaining cap, which is tightly screwed onto the gas acceleration tube. The vacuum pump is used to generate a vacuum within the chamber. Helium pressure is then built up until the rupture disk bursts, then allowed to drop and the vacuum released. DNA-coated gold particles are thereby fired at high velocity into the surface of the leaf. Cells in which the nucleus is hit by the gold particle are transformed and begin to express the genes encoded on the plasmids with which the gold particles were coated. **(C)** The samples are retained on 0.6% MS agar plates to maintain hydration prior to imaging. After 16-20 hours, the ER-RFP marker and eGFP fusions are visible in transformed cells via live-cell imaging. Schematic and methodology adapted from the Bio-Rad instruction manual and Tee *et al.* (Tee et al., 2022).

3.2.2 Starch and sugar quantification from *N. benthamiana* and Arabidopsis

3.2.2.1 N. benthamiana

Nicotiana benthamiana plants were grown under 12 h: 12 h light: dark conditions prior to infiltration. *Agrobacterium* suspensions for *p19* and *ChEC153-eGFP* were combined to a final OD_{600nm} of 0.5 (*p19* OD_{600nm} = 0.2 and *35S::ChEC153-eGFP* OD_{600nm} = 0.3), or *p19* was used alone at OD_{600nm} = 0.5. Two leaves of each of nine plants were infiltrated with the same *Agrobacterium* solution, and control- and *ChEC153*-expressing plants were staggered around the cabinet before and after infiltration to account for potential variation in light intensity. For each infiltrated plant, two leaf discs of 16 mm ⌀ were harvested three days post infiltration at the end of the photoperiod, weighed, and flash-frozen in liquid nitrogen.

3.2.2.2 Arabidopsis stable lines

Arabidopsis plants were germinated on MS + phosphinothricin (PPT, 10 µg/mL) plates (0.8% agar, 1% sucrose), and Col-0 were germinated on MS-only plates (0.8% agar, 1% sucrose). Following cold stratification (2 days at 5 °C) and two weeks of growth under short day conditions, seedlings were transplanted to 40-cell soil trays and grown in either a Sanyo growth cabinet (MLR-351H) at 12 h: 12 h (light: dark) conditions with light intensity at the maximum setting (5), or controlled environment room with long day conditions (16 h: 8h, light: dark) with 400 µmol photons m^{-2} s^{-1} light. Whole rosettes were harvested, weighed, and flash-frozen in liquid nitrogen at approximately 4.5 weeks post germination, at the end of the day and end of the night.

3.2.2.3 Starch and sugar extraction and quantification

Starch and sugar were separately extracted from approximately 100 - 250 mg total leaf tissue as described by Smith and Zeeman (Smith and Zeeman, 2006). Tissue was homogenised in 0.7 M cold perchloric acid. 200 µL of each starch sample was gelatinised by incubation at 95 °C for 12 minutes, briefly cooled at room temperature and then digested by incubation with α-amylase (10 U) and amyloglucosidase (1.26 U) in 0.1045 M sodium acetate buffer (pH 4.8) for two hours at 37 °C. Non-digested control samples were likewise gelatinised and incubated at 37 °C for two hours in 0.1045 M sodium acetate.

Starch content was quantified through an enzymatic assay of digested sample glucose content, by monitoring NADH production via absorbance at 340 nm in 96-well format using an Omega FLUOstar microplate reader in a hexokinase/glucose-6-phosphate dehydrogenase-based assay.

Technical triplicates were included for each digested sample, plus a single well per non-digested sample. Starch content in mg per gram of fresh weight of sample was calculated from the glucose assay values. Soluble sugars were similarly quantified by enzymatic assay of extracted sugars to measure glucose, fructose and sucrose sequentially following addition of glucose-6-phosphate dehydrogenase, phosphoglucose isomerase and invertase, respectively. Technical triplicates were measured for each sample. The means of the technical triplicates were plotted in R with ggplot (Wickham, 2009) and statistical differences between control and *ChEC153*-expressing samples evaluated relative to *p19*-only or Col-0 controls using unpaired Welch two sample t-tests. In plots, asterisks denote p-values: ns, $p > 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Outliers were retained within the data (shown by grey points in plots).

3.2.3 Starch purification from Arabidopsis rosettes

Whole rosettes of 60 approximately three-week-old Arabidopsis plants were harvested into liquid nitrogen at the end of the light period. Leaf material was ground in liquid nitrogen in a pestle and mortar and homogenised in 100 mL starch buffer II (50 mM Tris-HCl pH 8.0, 0.2 mM EDTA, 0.5% v/v Triton X-100). This suspension was filtered through one layer of Miracloth (Millipore, 475855) and then a 60 µm nylon net. The collected filtrate was centrifuged at 3000 ×g for 15 minutes, and the resultant pellets resuspended in 40 mL starch buffer II. Samples were then filtered through a 15 µm nylon net and filtrates were immediately centrifuged at 2500 ×g for 15 minutes over a Percoll cushion (95% v/v Percoll, 5% v/v 0.5 M Tris-HCl, pH 8.0). Pellets were resuspended in water and transferred to 1.5 mL microfuge tubes. Samples were centrifuged at maximum speed in a benchtop microcentrifuge (~20,000 ×g) for 1 minute. Supernatants were discarded and the surface of the starch cleared of visible cell remnants using a pipette, repeating this process until the visible dirt was removed. Starch was then washed twice in 0.5% w/v SDS in water and twice in dH_2O to remove SDS. Finally, for particle size analysis, purified starch resuspended in dH_2O was stored at -20 °C.

3.2.4 Determination of starch granule particle size distributions

Starch granule size distributions were determined from purified starch using a Multisizer 4e Coulter counter (Beckman Coulter, B43905) particle size analyser calibrated with 2 µm latex beads (size standard L2: Beckman Coulter, C72513). Starch was suspended in Isoton II diluent (Beckman Coulter, 8546719) and the Coulter counter was set to measure at least 100,000 starch granules using a 30 µm aperture. Each sample was measured in technical triplicates. Data were visualised as equivalent spherical diameters of starch granules from volume measurements in R using ggplot2 (Wickham, 2009). A Python script written by Rose McNelly for the analysis of unimodal size

distribution traces was used to extract mean starch granule diameters from the raw data using lognormal curve fitting. The script used is available at: https://github.com/DavidSeungLab/Coulter-Counter-Data-Analysis (Esch et al., 2023). Statistical analysis of the average granule diameter from the three technical replicates was then used to compare the starch granule sizes from different lines: particle size data were analysed using an ANOVA on a linear model followed by post-hoc Tukey tests. A schematic of the process of starch granule size determination is presented in **Figure 3.2** for illustrative purposes.

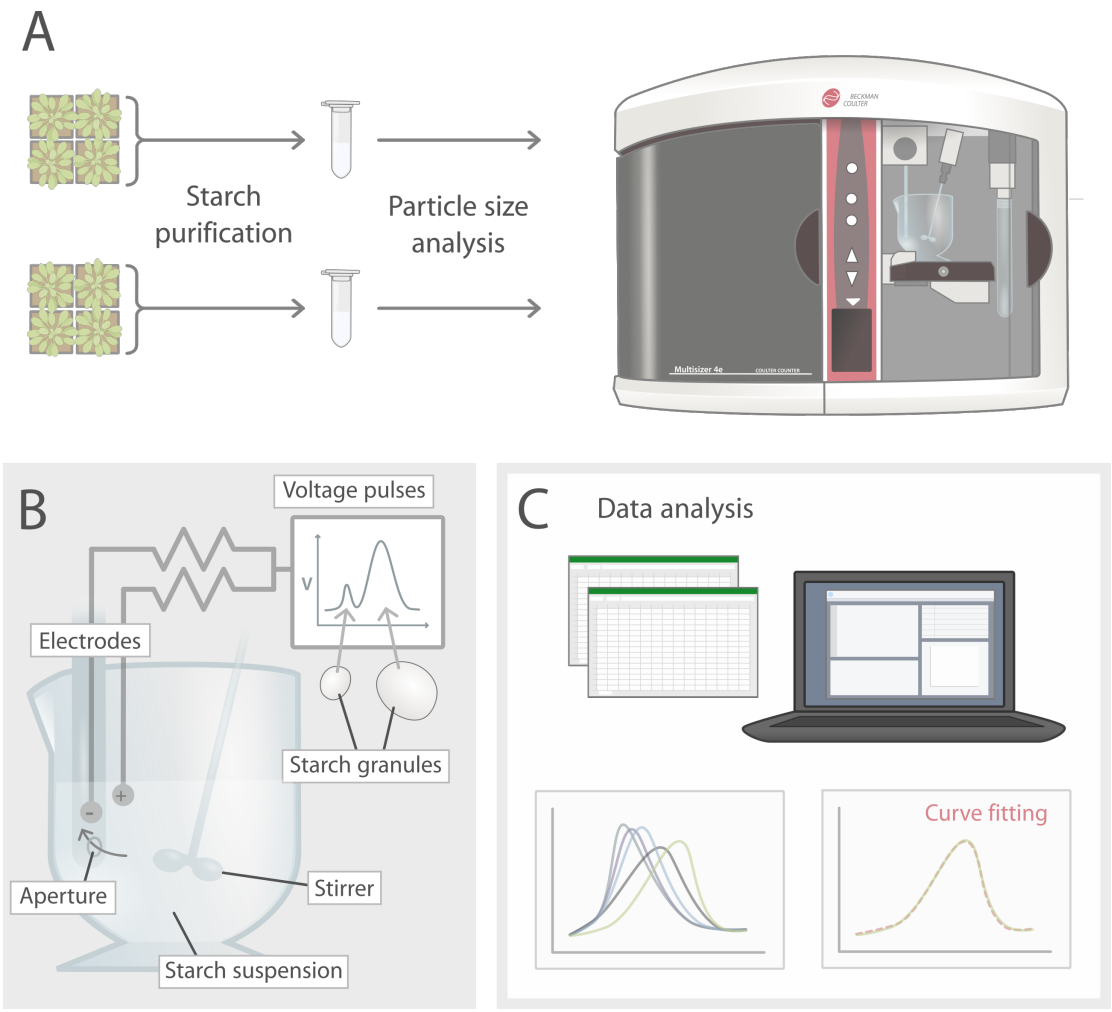


Figure 3.2: Schematic to illustrate starch granule particle size analysis. A Multisizer 4e Coulter counter particle size analyser was used for the determination of starch granule size distributions. **(A)** 60 Arabidopsis rosettes of the same genotype were pooled and used for starch purification. Purified starch suspended in Isoton II diluent was used as the input for the Coulter counter. Samples were measured in technical triplicate. **(B)** The particle size analyser works on the basis of the Coulter principle whereby changes in the electrical conductance seen as particles pass through a small aperture can be measured and converted into a metric of particle volume. **(C)** From resulting data, curve fitting analysis was performed to extract values of average granule diameter for statistical analysis, and plots were generated in R.

3.2.5 Scanning electron microscopy (SEM)

Scanning electron microscopy of purified starch was carried out by Dr David Seung. Starch granules were purified as described in **Section 3.2.3** and stored in water at -20 °C. A droplet of this starch suspension was dried onto a glass coverslip on an SEM stub. A Nova NanoSEM 450 (FEI) scanning electron microscope was used to image these starch granules.

3.3 RESULTS

3.3.1 ChEC153 and HaRxL94b localise to the same discrete, chloroplastic puncta

Despite *C. higginsianum* and *H. arabidopsidis* being of different kingdoms, ChEC153 and HaRxL94b have a remarkably similar localisations when transiently expressed in *N. benthamiana* (**Figure 1.5** and **Figure 1.6**, respectively). That both ChEC153-eGFP and HaRxL94b-eGFP localise to foci within the chloroplast raises the question as to whether the puncta of the putative effectors represent the same localisations within the chloroplast, or whether they exist in two distinct populations of puncta. In order to address this question, I used *Agrobacterium*-mediated transient expression to co-express both putative effectors in *N. benthamiana* as either ChEC153-eGFP + HaRxL94b-mCherry or ChEC153-mCherry + HaRxL94b-eGFP. Confocal microscopy revealed comprehensive overlap between the eGFP and mCherry signals in either combination, with all observations of the proteins showing co-localisation. The putative effectors co-localise with one another when co-expressed (**Figure 3.3**). Therefore, these effector candidates localise to the same structures within the chloroplast and may be targeting the same host proteins or processes as part of their function.

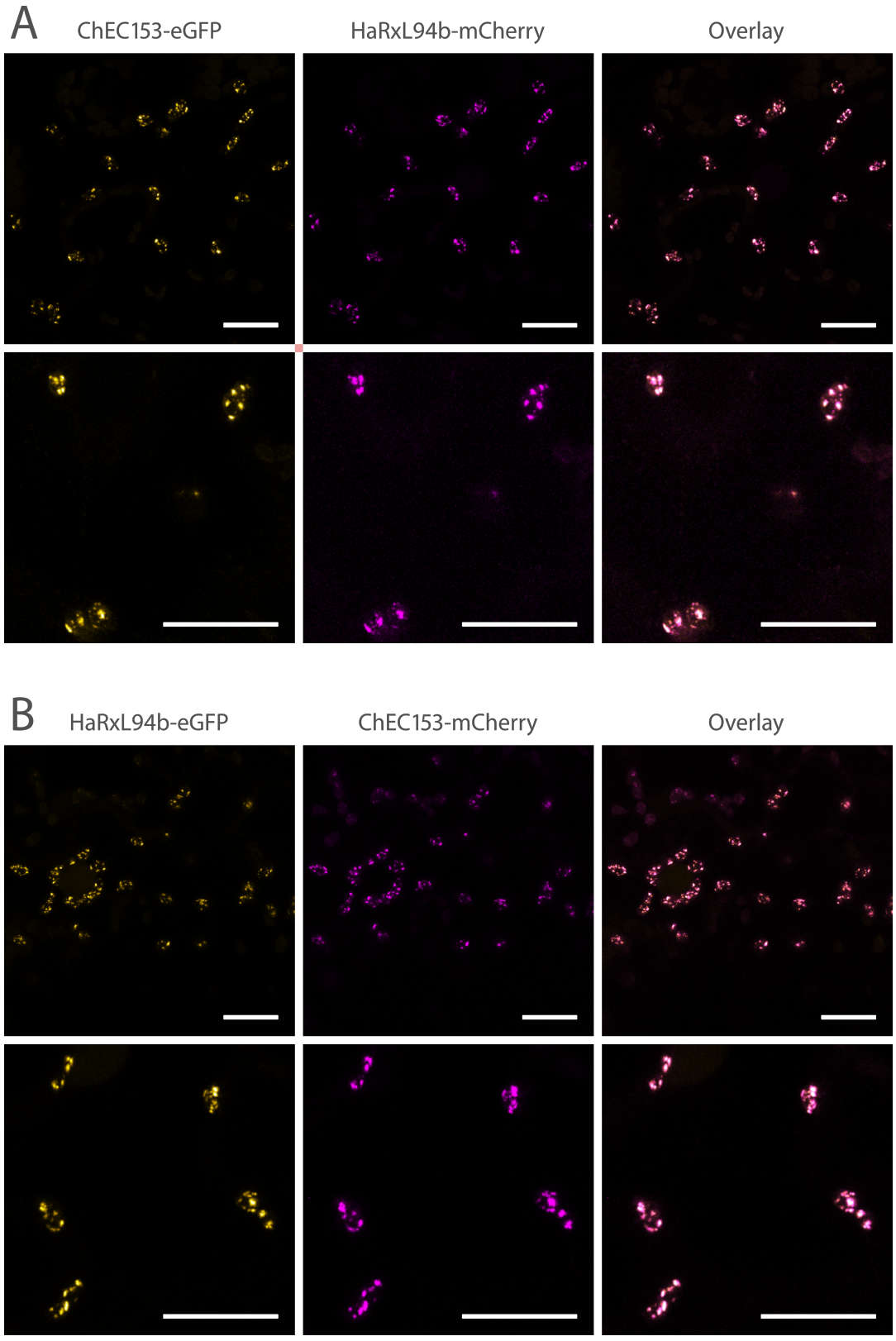


Figure 3.3: ChEC153 and HaRxL94b co-localise with one another. Confocal images of putative effector constructs co-expressed during *Agrobacterium*-mediated transient expression in *N. benthamiana*, as either combination of **(A)** *35S::ChEC153-eGFP* + *35S::HaRxL94b-mCherry* and **(B)** *35S::HaRxL94b-eGFP* + *35S::ChEC153-mCherry*. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

3.3.2 *In silico* characterisation of the putative effectors ChEC153 and HaRxL94b

Based on the co-localisation of ChEC153 and HaRxL94b and their potential to target the same host processes (**Figure 3.3**), I sought to analyse the protein sequences *in silico* to look for any similarities between them or hints as to their functions. At the nucleotide and amino acid levels, the two putative effectors have no sequence similarity.

3.3.2.1 ChEC153 is predicted to be secreted from *C. higginsianum*

The classification of ChEC153 as a putative effector by Dr Mina Ohtsu and Dr Joanna Jennings was partly based on it having a predicted signal peptide for secretion from the fungus. SignalP 4.1 was used to confirm the likelihood of ChEC153 it to possess an N-terminal signal peptide, which is predicted to be cleaved between residues 20 and 21 (Ohtsu et al., 2023, Petersen et al., 2011). SignalP 3.0 agrees with this cleavage point prediction when using the neural network predictors mode (Bendtsen et al., 2004). Interestingly, signal peptide prediction using newer tools gives a slightly different prediction of the signal peptide cleavage point (Armenteros et al., 2019b, Armenteros et al., 2019a, Teufel et al., 2022, Bendtsen et al., 2004). An analysis by Sperschneider *et al.* indicated that the neural network based prediction functions of SignalP 3.0 are generally better suited for use in predicting fungal signal peptides than hidden Markov model predictor options (Sperschneider et al., 2015b). For the purposes of this study, I therefore assume the signal peptide cleavage point to be between residues 20 and 21 as originally predicted by Dr Mina Ohtsu and Dr Joanna Jennings using SignalP 4.1. Thus, the majority of this work uses the mature protein sequence of ChEC153 corresponding to residues 21-521.

Further, within the effector candidate screening criteria, ChEC153 was not predicted to possess any transmembrane regions or a GPI-anchor (Petersen et al., 2011 Pierleoni et al., 2008). I confirmed this using DeepTMHMM and NetGPI-1.1 (Hallgren et al., 2022, Gislason et al., 2021).

3.3.2.2 ChEC153 is a putative cytochrome P450 protein

In order to make inferences about the function of the putative effector ChEC153, the mature protein sequence without the signal peptide was used as an input for several domain prediction tools. InterPro predicts ChEC153 to be a member of the cytochrome P450 family, with E-class group I signatures. It also annotates much of the protein as CYP60B-like, and predicts a cytochrome P450 heme-iron ligand signature near the protein C terminus (Paysan-Lafosse et al., 2022, Jones et al., 2014). The IMI 349063 genome assembly ASM167251v1 (Zampounis et al., 2016) annotates

the protein at CH63R_12252 (ChEC153) as cytochrome P40 monooxygenase [sic], but I conclude that the protein is likely a cytochrome P450.

3.3.2.2.1. Structure predictions of ChEC153

In agreement with the domain predictions, the structure of ChEC153 is expected to have a characteristic cytochrome P450 fold according to homology modelling using Phyre2 (Kelley et al., 2015). The ChEC153 model .pdb file generated was loaded into ChimeraX for visualisation (Pettersen et al., 2021), with the structure presented in **Figure 3.4 A**.

The structure of ChEC153 was also predicted using AlphaFold2 via the ColabFold platform (Mirdita et al., 2022) in ChimeraX. The ChimeraX software was also used to visualise the resulting structure (Pettersen et al., 2021), which is presented in **Figure 3.4 B**. The vast majority of this structure is modelled with high or very high confidence (pLDDT values above ~80).

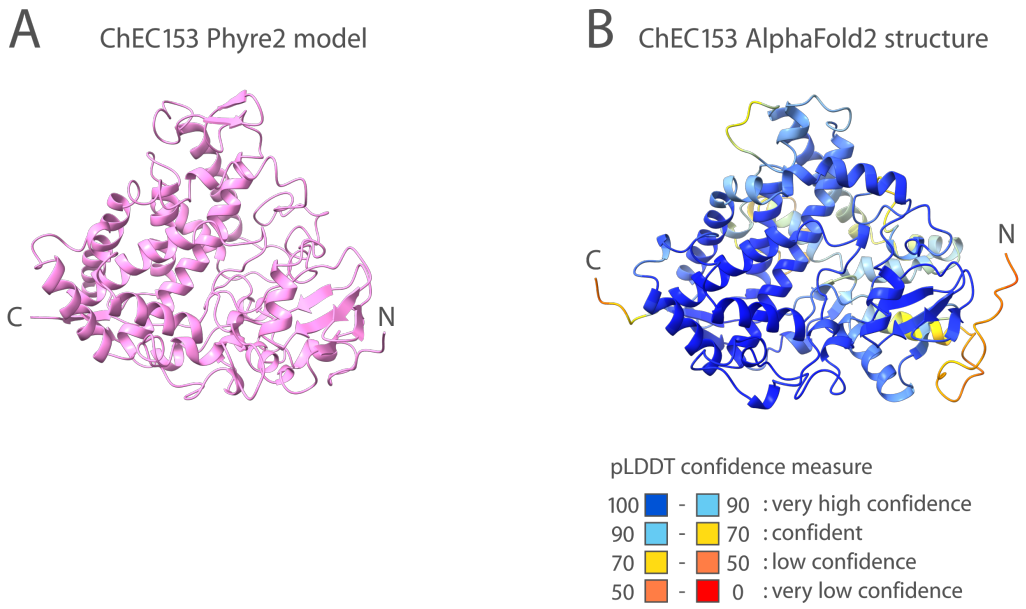


Figure 3.4: **Predicted structure of the mature ChEC153 protein. (A)** Phyre2 homology modelling. **(B)** AlphaFold2 predicted structure, model coloured according to predicted local distance difference test (pLDDT) scores. Models are oriented with respect to one another using the Matchmaker function in ChimeraX.

To my knowledge, no cytochrome P450 fungal effector proteins have yet been characterised (Carreón-Anguiano et al., 2020). Thus, should ChEC153 be a true effector protein, this would represent a novel discovery.

3.3.2.3 ChEC153 lacks a canonical chloroplast transit peptide

Despite the chloroplastic subcellular localisation seen for ChEC153 during transient expression in *N. benthamiana* (**Figure 1.5** and **Figure 3.3**), none of the localisation prediction tools I used predicted the mature ChEC153 protein to localise to the chloroplast. TargetP 2.0 did not recognise signatures of the protein targeting the chloroplast or mitochondria (Armenteros et al., 2019a), with DeepLoc versions 1.0 and 2.0 predicting localisation to the mitochondrial membrane and endoplasmic reticulum, respectively (Armenteros et al., 2017, Thumuluri et al., 2022). Further, LOCALIZER did not identify targeting of the chloroplast, mitochondria or the nucleus (Sperschneider et al., 2017). This suggests that the protein lacks a canonical N-terminal chloroplast transit peptide, rather targeting the chloroplast by a less well-characterised method.

3.3.2.4 HaRxL94b lacks clear predicted characterised domains and structural motifs

HaRxL94b possesses an RxLR motif near its N-terminus, critical in its classification as a putative effector. Shortly downstream of this RxLR motif is an EER motif, typical of RxLR effector proteins (Rehmany et al., 2005). RxLR motifs and effectors are introduced in general terms in **Section 1.4.2**, and this motif may be involved in protein secretion from oomycetes or translocation into the host plant. It is the C-terminal sequence downstream of the RxLR motif of oomycete RxLR effectors that is functionally diverse and could in theory offer inferences on the role of the effector. However, InterPro predicts no protein family membership for HaRxL94b, and only annotates a potential short, disordered region at the C terminus, with a coil just upstream of it (Paysan-Lafosse et al., 2022, Jones et al., 2014). HaRxL94b is therefore not predicted to have any specific characterised domains. This clearly sets it apart from the *C. higginsianum* putative effector ChEC153, which is consistently predicted to have a cytochrome P450 domain. HaRxL94b is not predicted to be GPI-anchored (Gislason et al., 2021, Pierleoni et al., 2008), and does not have any predicted transmembrane domains according to DeepTMHMM (Hallgren et al., 2022).

3.3.2.4.1 Structure predictions of HaRxL94b

Phyre2 homology modelling of HaRxL94b returned a high-confidence prediction (98.7% confidence) of the structure of only 43% of the sequence (Kelley et al., 2015). The model produced by Phyre2 is presented in **Figure 3.5 A**. This region of high confidence homology modelling corresponds to residues 183-399 of the HaRxL94b mature protein sequence. This low coverage may be due to limitations in the limited number of template structures in the Phyre2 library. The template of highest similarity to HaRxL94b was given as a crystal structure of *P. sojae* PHYTOPHTHORA SUPPRESSOR OF RNA SILENCING 2 (*Ps*PSR2, template c5gncA_). PSR2 is an RxLR

effector involved in RNA silencing suppression, which contributes to Arabidopsis susceptibility to *Phytophthora* but not *Pseudomonas* (Xiong et al., 2014). It appears that the region of PSR2 to which HaRxL94b is expected to have structural homology includes helix bundles 1 and 2 of the LWY2 and LWY3 motifs, extending into helix bundle 1 of LWY4. These LWY modules are suggested to be both structural and functional units common in *Phytophthora* effectors which may enable virulence and effector diversification (He et al., 2019). Despite predicting the structure of the two effectors HaRxL94b and PSR2 to be similar in this region, the percentage identity for amino acids is only 23%. *Hpa* and *P. sojae* are both species of the *Peronosporaceae* family of oomycetes, and thus may be considered fairly closely related. As chloroplasts do not contain RNAi machinery it appears unlikely that HaRxL94b is involved in RNA silencing suppression.

The mature amino acid sequence of HaRxL94b was also used for structure prediction with AlphaFold2 via the ColabFold platform (Mirdita et al., 2022), accessed via ChimeraX which was additionally used for structure visualisation (Pettersen et al., 2021). The predicted structure is presented in **Figure 3.5 B**. The majority of the sequence is modelled with a reasonable level of confidence, though some helices have particularly low confidence (pLDDT scores below 50) and so cannot be considered accurately modelled. To conclude, it appears that HaRxL94b shares some localised structural similarity with RxLR effectors of other species, despite a lack of characterised domains or fold.

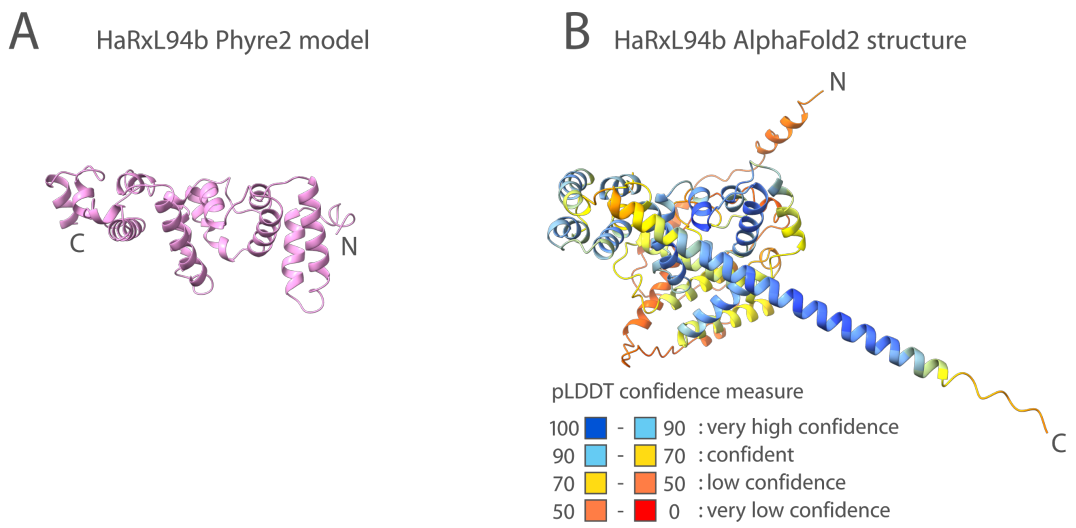


Figure 3.5: **Predicted structure of the mature HaRxL94b protein. (A) Phyre2 homology modelling.** This model includes only residues 183-399 of the HaRxL94b protein sequence. **(B)** AlphaFold2 predicted structure, model coloured according to predicted local distance difference test (pLDDT) scores. Models are oriented with respect to one another using the Matchmaker function in ChimeraX.

3.3.2.5 HaRxL94b has a canonical N-terminal chloroplast transit peptide

As shown in **Figure 1.6**, expressing the mature HaRxL94b protein coding sequences as a GFP-fusion in *N. benthamiana* reveals localisation to puncta within the chloroplasts, as well as to the nucleus. Consistent with this, where *in silico* localisation prediction tools returned a specific localisation, predictions for HaRxL94b tended to be chloroplastic (DeepLoc 1.0 (Armenteros et al., 2017)), or to detect a combination of chloroplastic, nuclear and mitochondrial localisation motifs (LOCALIZER (Sperschneider et al., 2017)). However, TargetP 2.0 does not predict a chloroplast localisation for HaRxL94b (Armenteros et al., 2019a). That some of these subcellular localisation prediction tools project HaRxL94b to localise to the chloroplast suggests that it may possess a typical N-terminal chloroplast transit peptide, corresponding to the chloroplastic localisation seen experimentally (**Figure 1.6** and **Figure 3.3**).

3.3.2.6 Further inferences about HaRxL94b from the literature

HaRxL94 was first identified as a candidate effector protein *H. arabidopsidis* using genomic data in 2010 (Baxter et al., 2010). Categorisation of HaRxL94 as a putative effector was later reiterated in the context of transcriptomics, alongside HaRxL94b (Asai et al., 2014). These two proteins have extensive sequence similarity, particularly at the amino acid level (97.58% identify with 100% query coverage). HaRxL94 has been shown using I-TASSER and DaliLite, along with four other putative *Hpa* effector proteins, to have some expected structural similarity with the AvrE1 effector of *Pseudomonas syringae*, despite a lack of sequence similarity and a difference in protein size (AvrE1: 1795 amino acids; HaRxL94: 496 amino acids, including the signal peptide; Deb et al., 2018b). Given the sequence similarity between HaRxL94 and HaRxL94b, it is highly likely that HaRxL94b also shares structural similarity with AvrE1.

HaRxL94b also has considerable sequence similarity to another putative *Hpa* effector, HaRxL92. These two putative effectors share 83.74% identity at the amino acid level with 90% query coverage. HaRxL92 was initially a candidate for localisation screening by Dr Xiaokun Liu, however the synthesis of the coding sequence failed and so the localisation in *N. benthamiana* was not determined. With the signal peptide cleavage point determined by SignalP 6.0, the mature HaRxL92 sequence localisation predictions include a primarily nuclear localisation (LOCALIZER) or a chloroplast membrane localisation (DeepLoc 1.0). It is possible therefore that HaRxL92 localises similarly to HaRxL94b. According to Boutemy and colleagues (Boutemy et al., 2011), HaRxL92 is predicted to have a WY-domain-like sequence, while HaRxL94 fell just below their threshold for characterisation as containing a WY-domain-like sequence. While HaRxL94b was not listed as an

HaRxLR protein tested for WY-domains-like sequences, we may assume it would score very similarly to HaRxL94 based on their considerable sequence similarity (Boutemy et al., 2011).

3.3.3 The chloroplastic localisation of HaRxL94b-eGFP is dependent on its N-terminal sequence

As shown in **Figure 1.6**, HaRxL94b-eGFP localises to chloroplastic puncta when transiently expressed in *N. benthamiana*, as well as a comparatively faint eGFP signal frequently being seen in the nucleus, suggesting dual targeting of the effector. HaRxL94b possess an N-terminal predicted signal peptide for secretion from the host, with the signal peptide (amino acids 1-23) being predicted by Dr Xiaokun Liu using SignalP 2.0 (Nielsen and Krogh, 1998, Liu et al., 2022).

Within the putative *H. arabidopsidis* effector protein localisation screen carried out by Liu and colleagues, HaRxL94, which differs from HaRxL94b by only 12 amino acids, was seen to localise predominantly to the nucleus (Liu et al., 2022). Aligning the sequences of the constructs cloned by Dr Xiaokun Liu, I found that the N-terminal region of HaRxL94 was truncated by 40 amino acids relative to HaRxL94b (**Figure 3.6**) due to a difference in signal peptide predictions using SignalP 2.0 (Nielsen and Krogh, 1998). The chloroplast transit peptide of HaRxL94b predicted by LOCALIZER is at residues 50-90 (Sperschneider et al., 2017), the beginning of the cTP being truncated in HaRxL94. This raised the question as to whether this N-terminal region of the protein is sufficient to dictate the differential localisations.

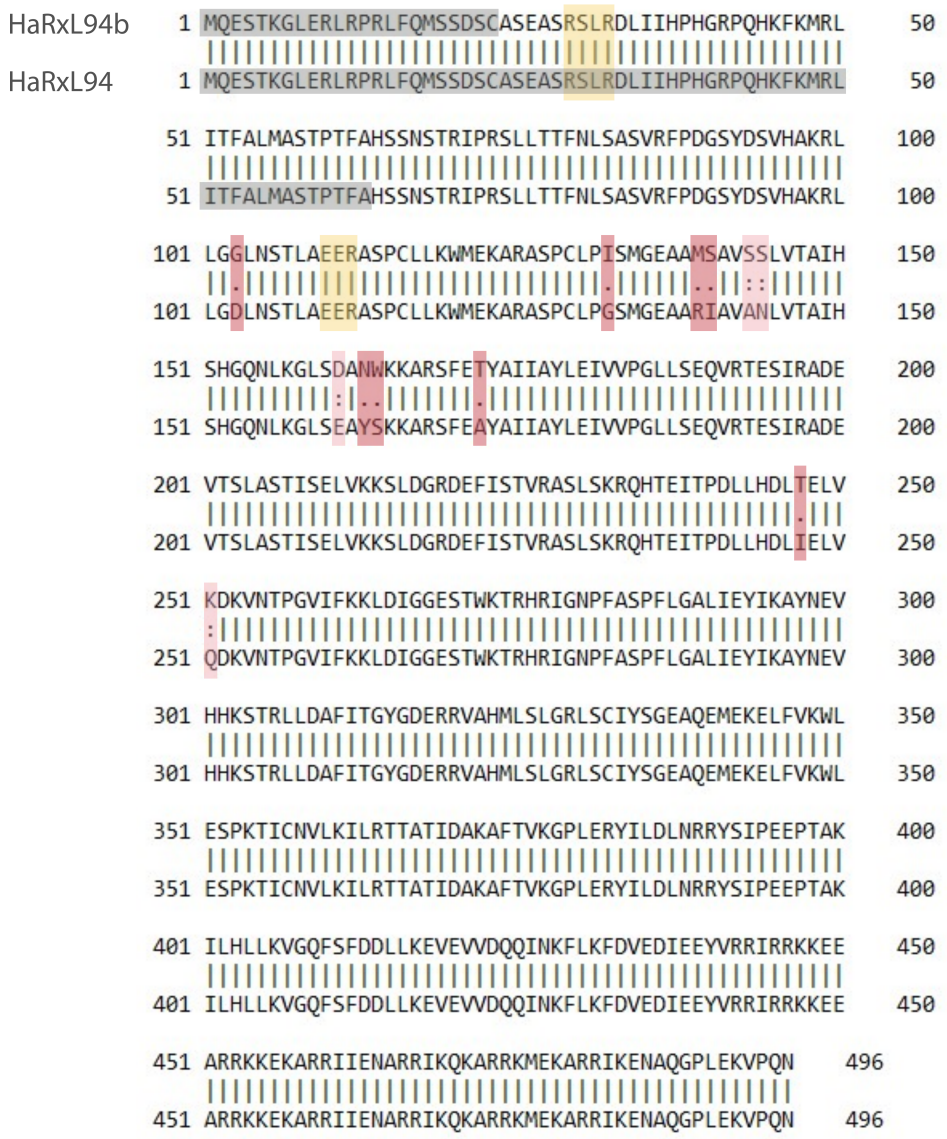


Figure 3.6: **HaRxL94b and HaRxL94 have considerable sequence identity at the amino acid level.** Pairwise alignment of amino acid sequences of HaRxL94b and HaRxL94 using EMBOSS Matcher. Sequences have 98.4% similarity with differences highlighted in: light pink, residues have largely similar properties; or dark pink, residues have less similar properties. The RxLR and EER motifs are highlighted in yellow. Residues absent from cloned constructs (putative signal peptides) are highlighted in grey.

In order to address the differences in localisations of these sequences, I first used the *in silico* effector localisation tool LOCALIZER (Sperschneider et al., 2017) to predict the subcellular localisations of these various forms of HaRxL94b and HaRxL94, with results presented in **Table 3.1**. The N-terminal region at residues 24-63 is required for either effector to be predicted to localise to the chloroplast. Each form of either putative effector is predicted to target the nucleus. In keeping with the results observed by Dr Xiaokun Liu, HaRxL94b$_{\Delta 1\text{-}23}$ is predicted to localise to the chloroplast as well as the nucleus, whereas HaRxL94$_{\Delta 1\text{-}63}$ is predicted to localise only to the nucleus.

I also cloned HaRxL94b with the same N-terminal truncation as seen in HaRxL94 (HaRxL94b$_{\Delta 1\text{-}63}$) and generated a construct for the HaRxL94 sequence with an N-terminal extension to match that of HaRxL94b (HaRxL94$_{\Delta 1\text{-}23}$). I expressed these constructs in *N. benthamiana* via *Agrobacterium*-mediated transient expression, with C-terminal eGFP fusions, as illustrated in **Figure 3.7 A**. The *in silico* localisation predictions (**Table 3.1**) proved to be very accurate, with the Δ1-23 forms of either effector localising to puncta within chloroplasts, in addition to some faint expression in the nucleus (**Figure 3.7 B** and **C**), and the truncated versions localising predominantly to nuclei (**Figure 3.7 D** and **E**). With these imaging settings, only when image contrast and brightness were adjusted did a faint signal associated with intracellular bodies become visible, possibly corresponding to mitochondria. Further imaging with different settings and a mitochondrial marker would be required to confirm this, but preliminary observations supported the predictions. Hence, and in agreement with the *in silico* predictions, the 12 amino acid differences between the HaRxL94b and HaRxL94 sequences do not impact the localisations of these proteins. Instead, the N-terminal sequence at residues 24-63 of these effectors is critical for chloroplast targeting, likely due to the removal of part of the canonical cTP in the Δ1-63 forms of the putative *Hpa* effectors.

Table 3.1: **Subcellular localisation predictions for HaRxL94b and HaRxL94 with various N-terminal truncations using LOCALIZER, and experimentally determined localisations.** For *in silico* predictions, input sequences are indicated with amino acid sequence lengths indicating the truncations: to the start of the original HaRxL94b sequence (Δ1-23) or truncated to the start of the original HaRxL94 sequence (Δ1-63). Outputs are summarised according to the presence (Y) or absence (N) of localisation prediction to the chloroplasts, mitochondria, or nucleus, with probability of predictions stated in brackets. Experimentally determined localisations were evaluated using transient expression in *N. benthamiana* with C-terminal eGFP fusions, with representative confocal images shown in **Figure 3.7**.

Sequence	*In silico* predictions (LOCALIZER)			Experimentally determined localisation
	Chloroplastic localisation	Mitochondrial localisation	Nuclear localisation	
HaRxL94b$_{\Delta 1\text{-}23}$	Y (1.0)	Y (0.996)	Y	Chloroplastic puncta, faint nuclear
HaRxL94$_{\Delta 1\text{-}23}$	Y (1.0)	Y (0.996)	Y	Chloroplastic puncta, faint nuclear
HaRxL94b$_{\Delta 1\text{-}63}$	N	Y (0.901)	Y	Nuclear
HaRxL94$_{\Delta 1\text{-}63}$	N	N	Y	Nuclear

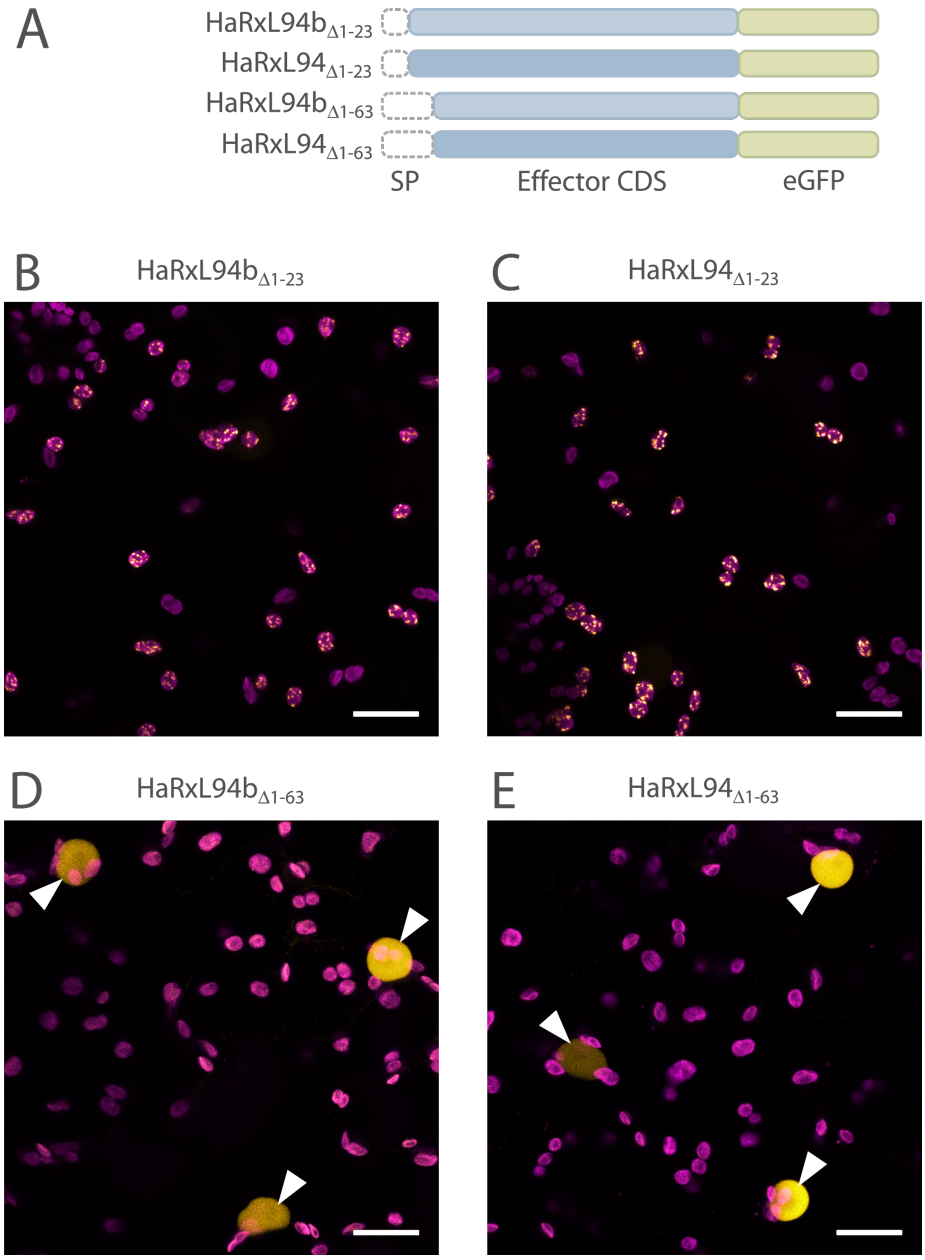


Figure 3.7: The N-terminus of HaRxL94b dictates its chloroplastic localisation. N-terminal truncations of HaRxL94b result in a nuclear localisation. **(A)** Schematic of regions cloned to test the importance of the N-terminal sequence for localisation. Sequences differ by the length of the N-terminal truncation of the putative signal peptide. **(B-E)** Confocal images collected during *Agrobacterium*-mediated transient expression in *N. benthamiana* of sequences shown in (A), driven by a 35S promoter. eGFP signal shown in yellow, with chlorophyll autofluorescence in magenta. Arrowheads indicate localisation to nuclei. Confocal images shown are maximum intensity projections of z-stacks. Scale bars, 20 µm.

The fact that different signal peptide sequences were predicted for HaRxL94b and HaRxL94 is surprising given their high sequence similarity. The first 101 amino acids are identical in these proteins (**Figure 3.6**), and these N-terminal residues (particularly those upstream of the RxLR motif)

are the ones considered important for the signal peptide prediction. SignalP 2.0, which was used by Dr Xiaokun Liu to predict the signal peptides, is no longer accessible so I was unable to confirm the differential signal peptide predictions. The prediction for the signal peptide of HaRxL94b appears more conventional, given that the signal peptide cleavage point was predicted to be just upstream of the RxLR motif. Generally, RxLR motifs are expected to be shortly downstream of the signal peptide cleavage point (Rehmany et al., 2005). It is possible that in reality both HaRxL94b and HaRxL94 share the 23 amino acid signal peptide seen in the HaRxL94b original prediction, and both target the host chloroplast. However, more recent versions of SignalP (SignalP versions 3.0, 4.1, 5.0 and 6.0 (Bendtsen et al., 2004, Petersen et al., 2011, Armenteros et al., 2019b, Teufel et al., 2022)) predict neither HaRxL94b nor HaRxL94 to have a signal peptide. While an analysis of the reliability of various signal peptide prediction tools cautions against using SignalP 4.1 to predict oomycete signal peptides based on a comparatively low sensitivity, they found both versions 2.0 and 3.0 to perform well with the application of hidden Markov model predictors (Sperschneider et al., 2015b). This knowledge and the presence of the RxLR and EER motifs in both of these sequences gives compelling evidence of their potential function as secreted effector proteins. In agreement with this, EffectorP 3.0 predicts both HaRxL94b and HaRxL94 to be cytoplasmic effector proteins (Sperschneider and Dodds, 2022). In all further experiments in this work, the HaRxL94b sequence used is that determined by Dr Xiaokun Liu, comprising residues 24-496: HaRxL94b$_{\Delta 1\text{-}23}$.

3.3.4 Do ChEC153 and HaRxL94b target sites of starch granule initiation?

The localisations observed for both ChEC153 and HaRxL94b are strikingly similar to those that have been seen for starch granule initiation proteins such as MRC. This raised the question as to whether these putative effector proteins are targeting the starch granule initial. Prior to my joining this project, Dr Andrew Breakspear therefore generated a pB7RWG2 *35S::MRC-RFP* construct and transiently expressed this alongside either *35S::ChEC153-eGFP* or *35S::HaRxL94b-eGFP* in *N. benthamiana*, seeing that both putative effector proteins co-localise with starch granule initiation protein MRC. I repeated this experiment and confirmed that both ChEC153-eGFP and HaRxL94b-eGFP co-localise with MRC-RFP during transient expression in *N. benthamiana*. Representative confocal images are shown in **Figure 3.8**. I conclude that the putative effectors are targeting MRC, or starch granule initiation in some way.

Within some but not all chloroplasts there appeared to be a greater number of HaRxL94b-eGFP puncta than MRC-RFP puncta, with some HaRxL94b-eGFP puncta being seen in the absence of an MRC-RFP punctum (**Figure 3.8 B**, white arrowheads). I did not find this to be the case for the *C. higginsianum* putative effector: all ChEC153-eGFP puncta overlapped with corresponding MRC-RFP puncta.

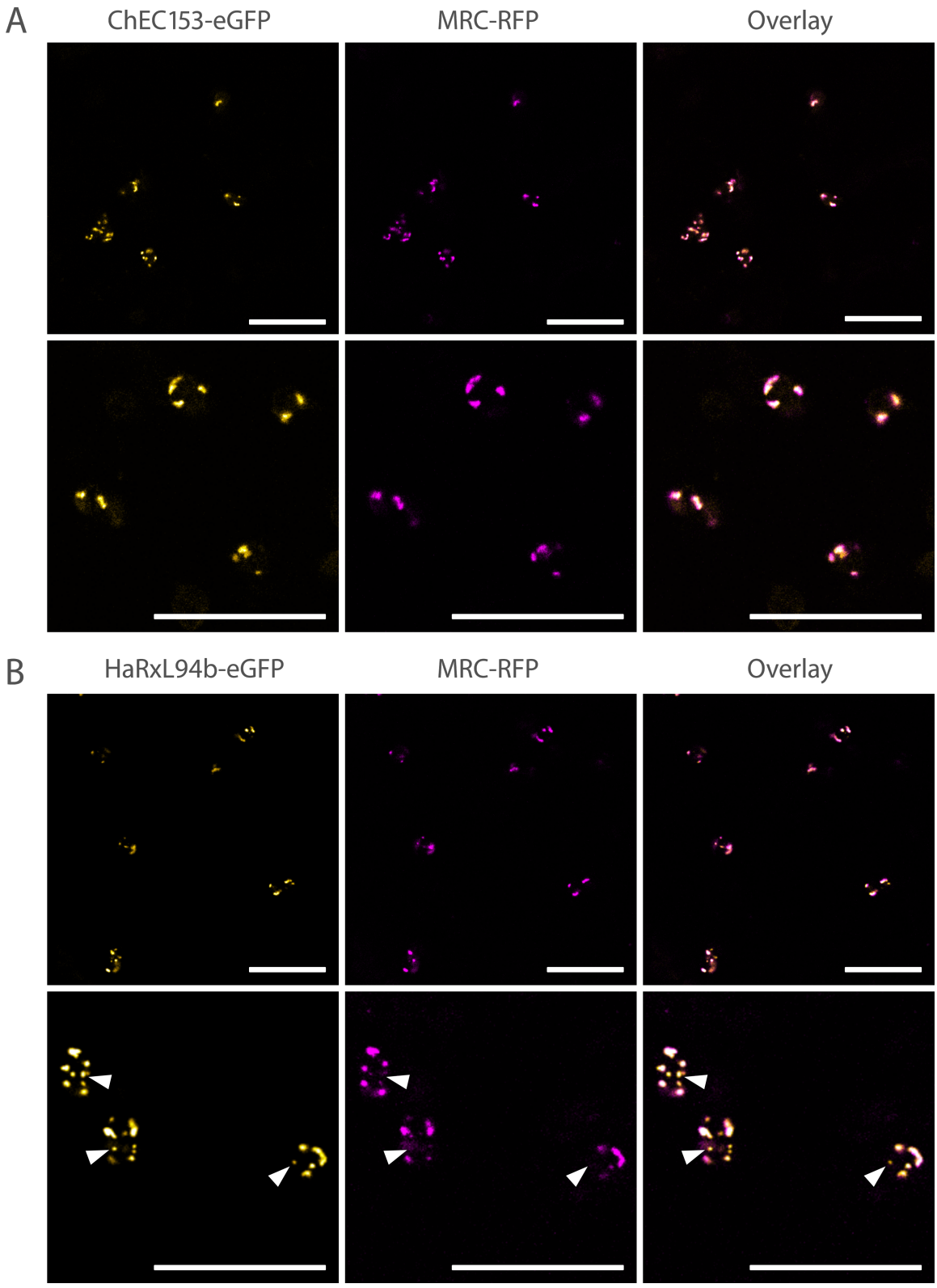


Figure 3.8: ChEC153 and HaRxL94b co-localise with MRC. Confocal images taken during *Agrobacterium*-mediated transient expression in *N. benthamiana* of *35S::MRC-RFP*, co-expressed with **(A)** *35S::ChEC153-eGFP* and **(B)** *35S::HaRxL94b-eGFP*. White arrowheads indicate an HaRxL94b-eGFP punctum in the absence of an MRC-RFP punctum. GFP signal is shown in yellow with RFP signal in magenta. Scale bars, 20 µm.

3.3.5 ChEC153 and HaRxL94b co-localise with TaCSP41a

As introduced in Section 1.10.1, TaCSP41a is a protein with RNA-binding and -cleaving capabilities implicated in regulation of plastid transcription/translation (Bollenbach et al., 2009). The wheat CSP41a ortholog has been identified as a putative susceptibility factor to yellow rust (*Pst*), and when transiently expressed in *N. benthamiana* was determined to be chloroplastic and punctate (Corredor-Moreno et al., 2022).

Based on the distinctive resemblance of this localisation to that of the putative effector proteins ChEC153 and HaRxL94b, I sought to co-express the putative effectors with this susceptibility factor. An *Agrobacterium* strain carrying a Gateway construct for *35S::TaCSP41a-eGFP* generated by Dr Andrey Korolev was kindly provided to me by the Saunders lab (Corredor-Moreno et al., 2022). I co-expressed this construct alongside either *35S::ChEC153-mCherry* or *35S::HaRxL94b-mCherry* transiently in *N. benthamiana*. From my observations, both putative effector proteins appear to co-localise with TaCSP41a (Figure 3.9).

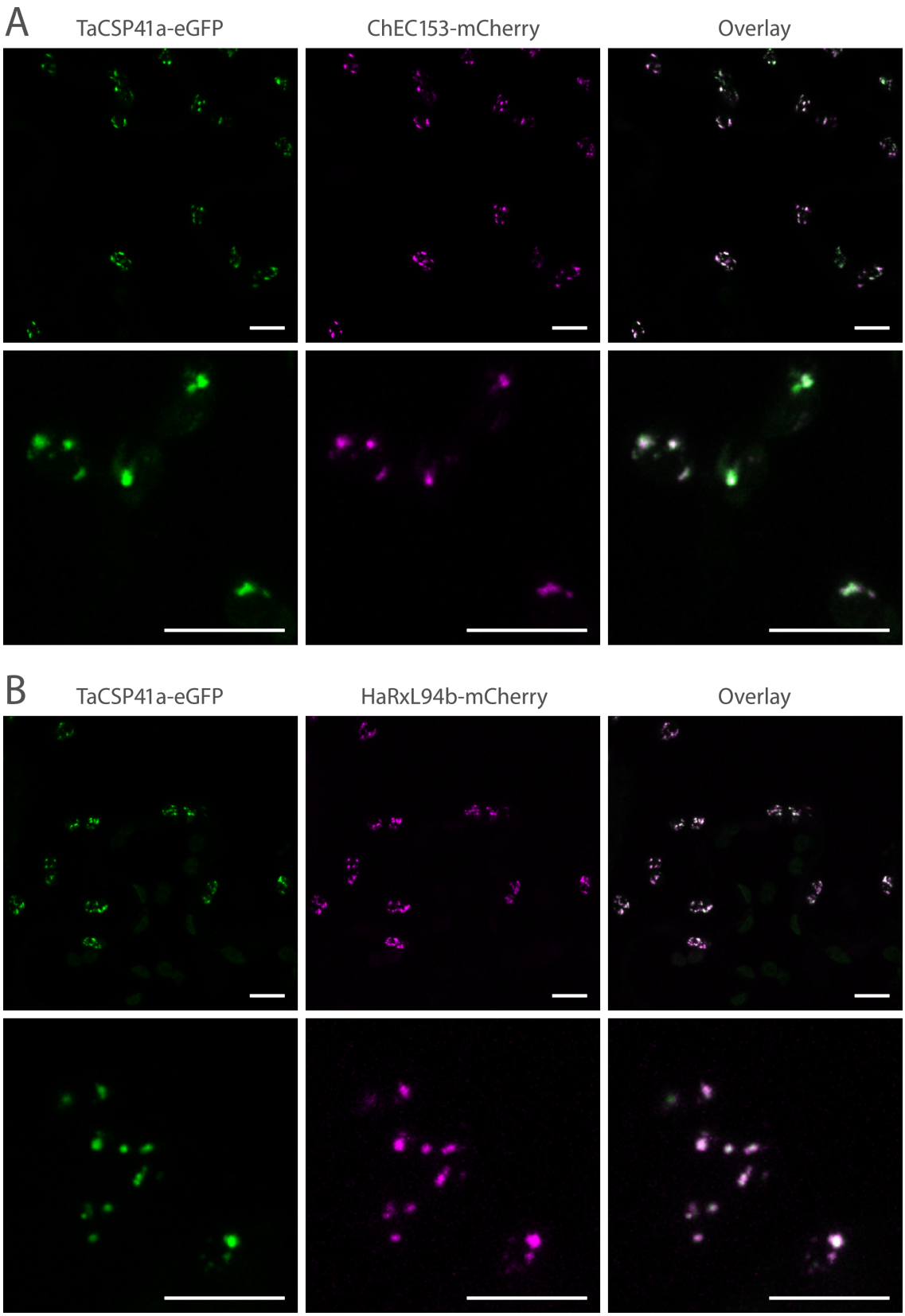


Figure 3.9: ChEC153 and HaRxL94b co-localise with TaCSP41a. Confocal images showing *Agrobacterium*-mediated transient expression in *N. benthamiana* of *35S::TaCSP41a-eGFP* co-expressed with **(A)** *35S::ChEC153-mCherry* and **(B)** *35S::HaRxL94b-mCherry*. Images are maximum intensity projections of z-stacks. GFP signal is shown in green with mCherry signal in magenta. Scale bars, 10 µm.

Based on the co-localisation of ChEC153 and HaRxL94b with both MRC and TaCSP41a, I co-expressed *35S::MRC-RFP* and *35S::TaCSP41a-eGFP* together in *N. benthamiana* to evaluate whether these two punctate chloroplastic proteins would co-localise with one another. I observed that MRC and TaCSP41a do appear to co-localise with one another (**Figure 3.10**). This may indicate that TaCSP41a localises to the site of starch granule initiation.

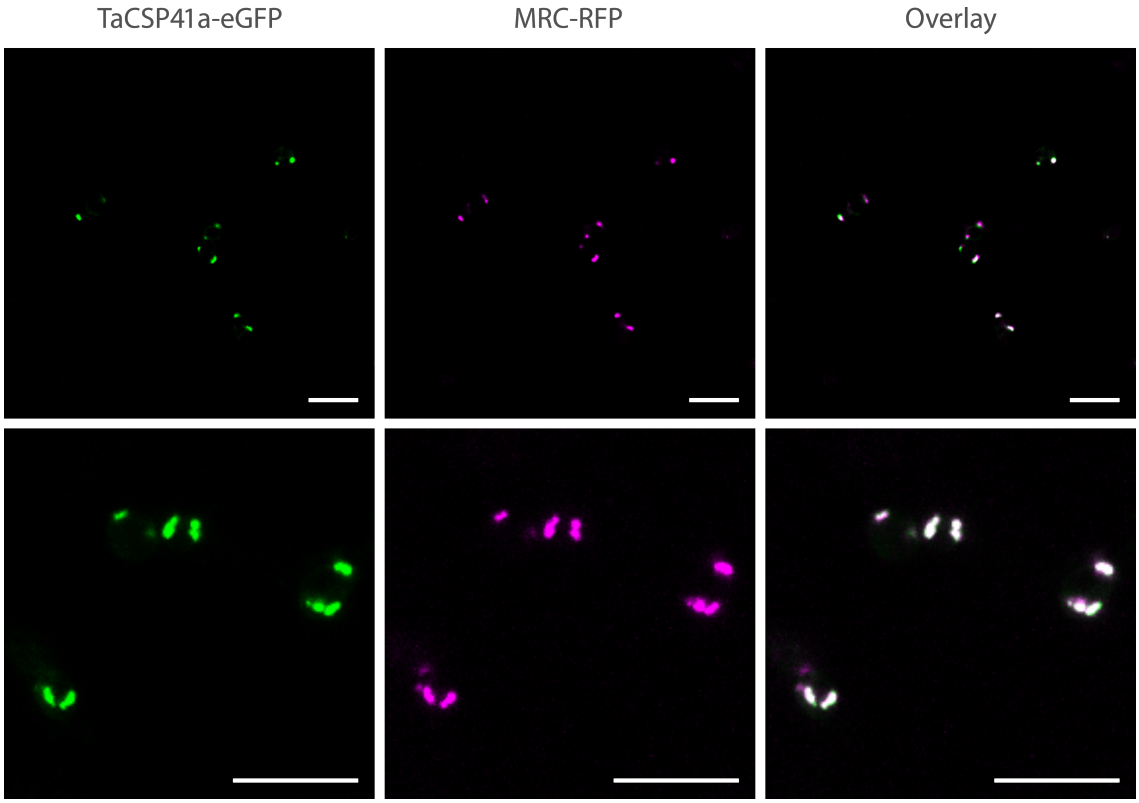


Figure 3.10: MRC co-localises with TaCSP41a. Confocal images showing *35S::MRC-RFP* co-expressed with *35S::TaCSP41a-eGFP* via *Agrobacterium*-mediated transient expression in *N. benthamiana*. Images are maximum intensity projections of z-stacks. GFP signal is shown in green with RFP signal in magenta. Scale bars, 10 µm.

3.3.6 The localisation of ChEC153 and HaRxL94b is partly associated with nucleoids

Given the similarity in localisation between the effectors and TaCSP41a, the link between TaCSP41a and nucleotide binding, and the fact that nucleoids are punctate structures within the chloroplast, I sought to investigate whether the effector localisation is reflective of nucleoids. To do this, the first 88 amino acids of the Arabidopsis PLASTID ENVELOPE DNA-BINDING protein ($AtPEND_{1-88}$) was used to mark the nucleoids (Terasawa and Sato, 2005). For this purpose, a pB7RWG2 35S::$AtPEND_{1-88}$-RFP construct was provided to me by the Seung lab, and used for transient co-expression of $PEND_{1-88}$-RFP alongside either ChEC153-eGFP or HaRxL94b-eGFP in *N. benthamiana*.

As shown by representative confocal images in **Figure 3.11**, I saw a partial co-localisation of the putative effectors and $PEND_{1-88}$-RFP. In some plastids, I observed a clear lack of co-localisation between the putative effector puncta and $PEND_{1-88}$ puncta, while in others the puncta appear to overlap. Based on these results, I am not able to conclude that the putative effector localisation is solely representative of nucleoids. The co-localisation of ChEC153 and HaRxL94b with MRC, in contrast, was far clearer, being seen consistently in all experiments carried out. Therefore, I conclude that there is only a partial co-localisation between the putative effectors and $PEND_{1-88}$.

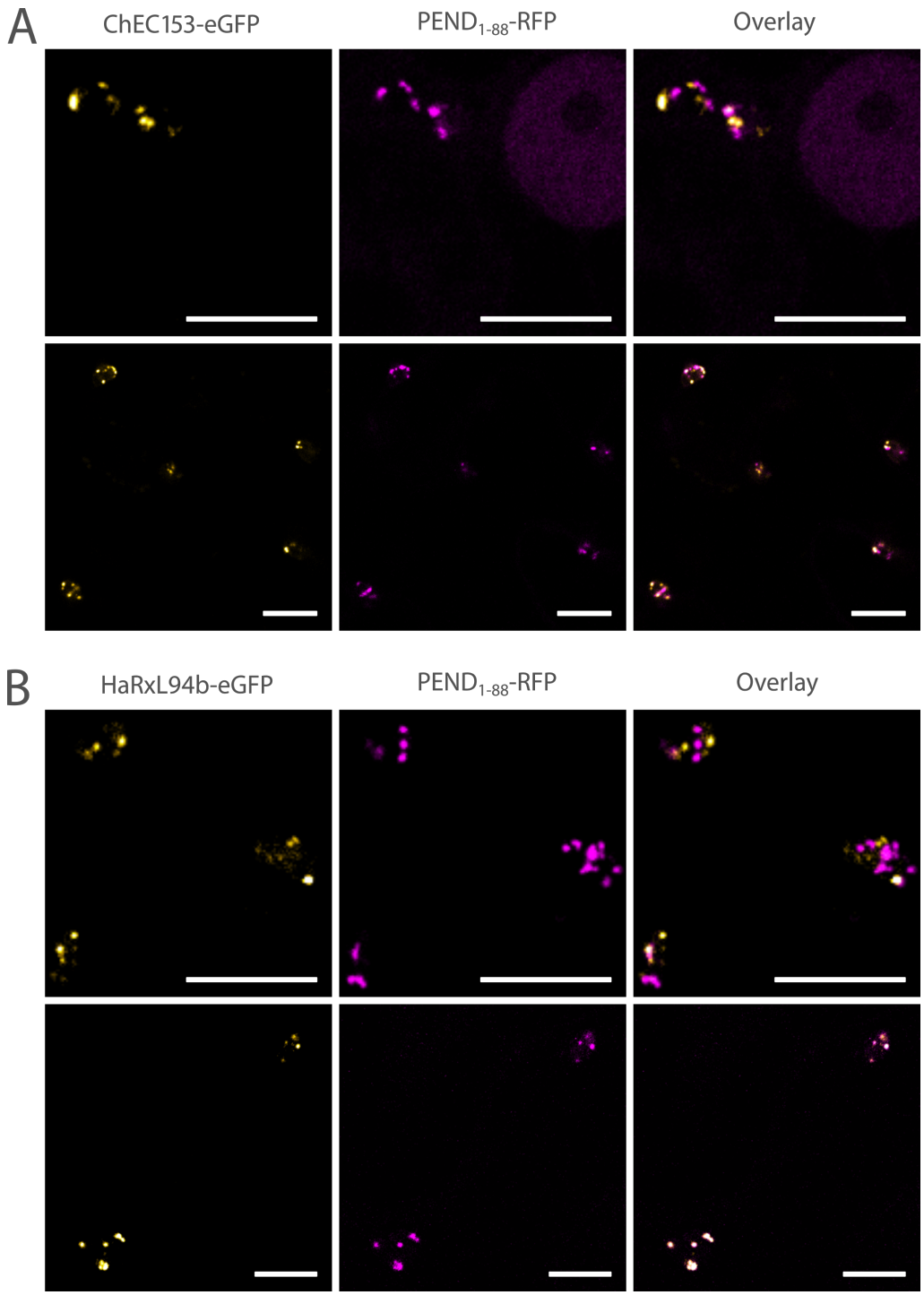


Figure 3.11: ChEC153 and HaRxL94b do not consistently co-localise with $PEND_{1-88}$. Confocal images taken during *Agrobacterium*-mediated transient expression in *N. benthamiana* of *35S::$PEND_{1-88}$-RFP* co-expressed with **(A)** *35S::ChEC153-eGFP* and **(B)** *35S::HaRxL94b-eGFP*. Images are maximum intensity projections of z-stacks. GFP signal is shown in yellow with RFP signal in magenta. Scale bars, 10 µm.

3.3.7 Transient expression of *ChEC153* induces formation of small starch granules

I wondered whether the effectors were directly targeting the starch granules of the chloroplast. GBSS, which is responsible for amylose synthesis, is tightly associated with the starch granule (Mu-Forster et al., 1996), and therefore when tagged with a fluorescent protein can be used as a marker for the starch granules within the chloroplast (Ichikawa et al., 2023). I used transient co-expression in *N. benthamiana* to look at the localisation of ChEC153 and HaRxL94b in relation to NbGBSSa-RFP, evaluating the positioning of the effector puncta in relation to the starch granules. The NbGBSSa sequence used is inactivated through an E490Q mutation in the catalytic domain ensuring that it can be used as a marker without unwanted effects on the *N. benthamiana* starch granules. Confocal images in **Figure 3.12 A** show the localisation of GBSS-RFP when expressed alone in *N. benthamiana*. In these images, the RFP signal marks the surface of several large starch granules within each chloroplast. **Figure 3.12 B** and **C** show confocal images of chloroplasts expressing both *GBSS-RFP* and either *ChEC153-eGFP* or *HaRxL94b-eGFP*, respectively. While the starch granules occupy the majority of the space within the chloroplasts imaged, ChEC153 and HaRxL94b puncta appear sometimes to be slightly spatially separated from the large starch granules. Thus, the effectors do not appear to be solely associated with these large starch granules. Unexpected puncta of GBSS-RFP were observed in the presence of the effectors, particularly ChEC153, indicated with white arrowheads in **Figure 3.12 B** and **C**. These GBSS-RFP puncta appear considerably smaller than the usual starch granules observed during expression of *GBSS-RFP* alone. Where clear GBSS-RFP puncta were observed, they generally co-localised with the effector puncta. It is possible that the effectors are associating with or inducing the formation of small starch granules, or else perturbing the normal localisation of the GBSS protein to the starch granule.

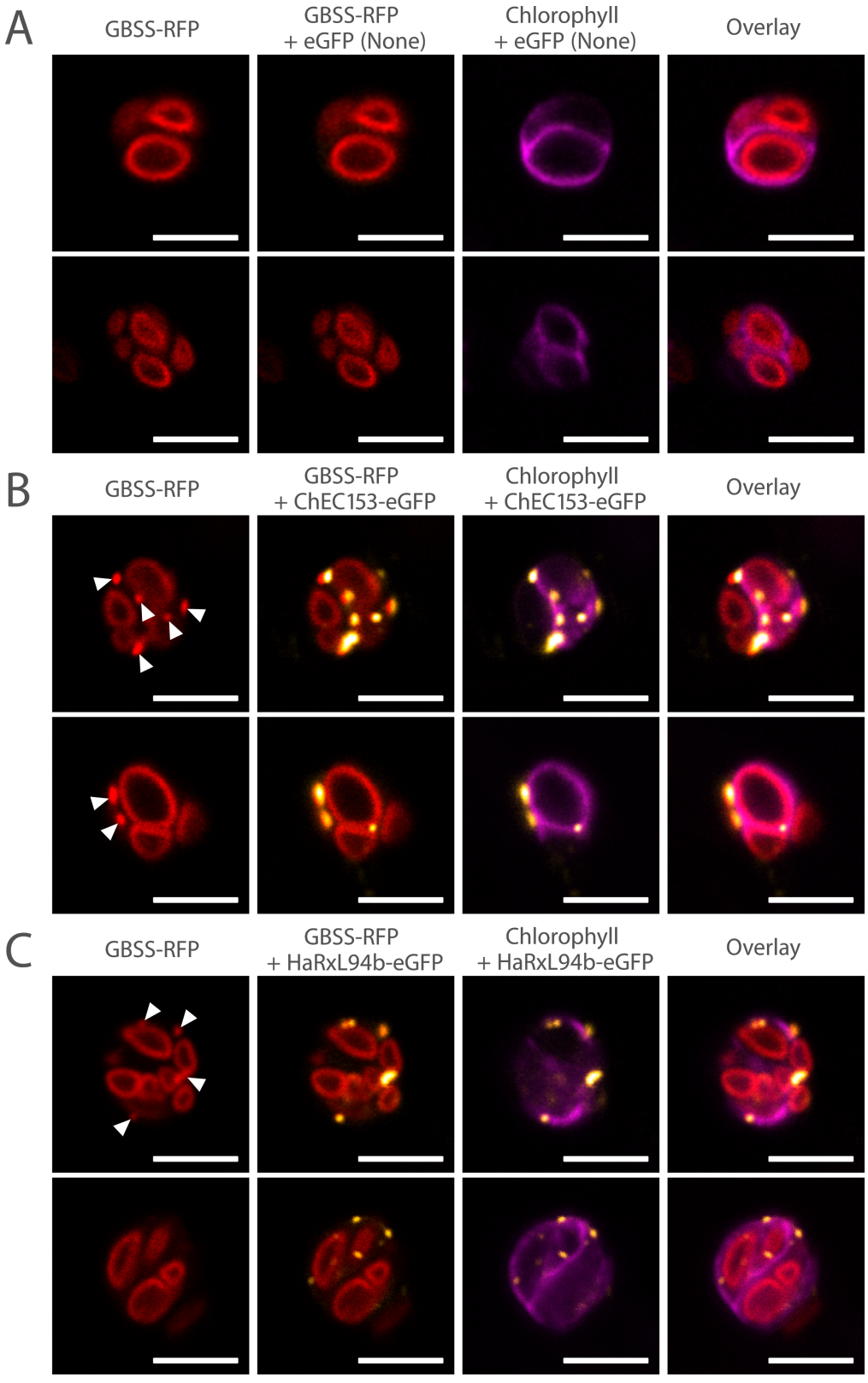


Figure 3.12: **Localisation of ChEC153 and HaRxL94b alongside starch marker GBSS.** Confocal images showing *Agrobacterium*-mediated transient expression in *N. benthamiana*. **(A)** *35S::GBSS-RFP* expressed alone; **(B)** *35S::GBSS-RFP* co-expressed with *35S::ChEC153-eGFP*; and **(C)** *35S::GBSS-RFP* co-expressed with *35S::HaRxL94b-eGFP*. Each panel shows images of two individual chloroplasts. White arrowheads in the RFP-only channel images indicate GBSS-RFP puncta. GBSS-RFP is shown in red, with ChEC153-eGFP and HaRxL94b-eGFP in yellow, and chlorophyll autofluorescence in magenta where shown. Scale bars, 5 µm.

In order to establish whether the putative effectors promoted the formation of these GBSS-RFP puncta, I repeated co-expression experiments to quantify the number of GBSS-RFP puncta per chloroplast in the presence and absence of ChEC153 and HaRxL94b, with the resulting quantification data presented in Figure 3.13.

Three independent replicates of the experiment were carried out, with 10-11 confocal z-stack images collected for each combination of *GBSS-RFP* alone, *GBSS-RFP* + *ChEC153-eGFP* and *GBSS-RFP* + *HaRxL94b-eGFP* in each replicate. From these images, the GBSS-RFP puncta in five chloroplasts expressing the expected fluorophores were counted. To ensure unbiased chloroplast selection, I used a grid system to select a quincunx of chloroplasts from each image for puncta quantification. A median bootstrapping method (Johnston and Faulkner, 2021) was applied to determine whether visible differences in the puncta frequency histograms are statistically significant.

Comparing GBSS-RFP puncta number in the presence and absence of ChEC153, a significant difference was observed, with expression of *ChEC153-eGFP* resulting in an increased GBSS-RFP puncta number (Figure 3.13 A). In contrast, comparing the puncta number in the presence and absence of HaRxL94b revealed no significant change in the number of GBSS-RFP puncta per chloroplast (Figure 3.13 B). Thus, co-expression of *ChEC153-eGFP* specifically promotes the formation of small GBSS-RFP puncta. This may represent induction of small starch granules or else a perturbation of the usual GBSS localisation by the putative *C. higginsianum* effector. Given that in some cases GBSS-RFP puncta were observed in the absence of either effector, it is likely that they do not represent an aberrant localisation of the GBSS-RFP protein and instead show the induced formation of small starch granules in the presence of ChEC153.

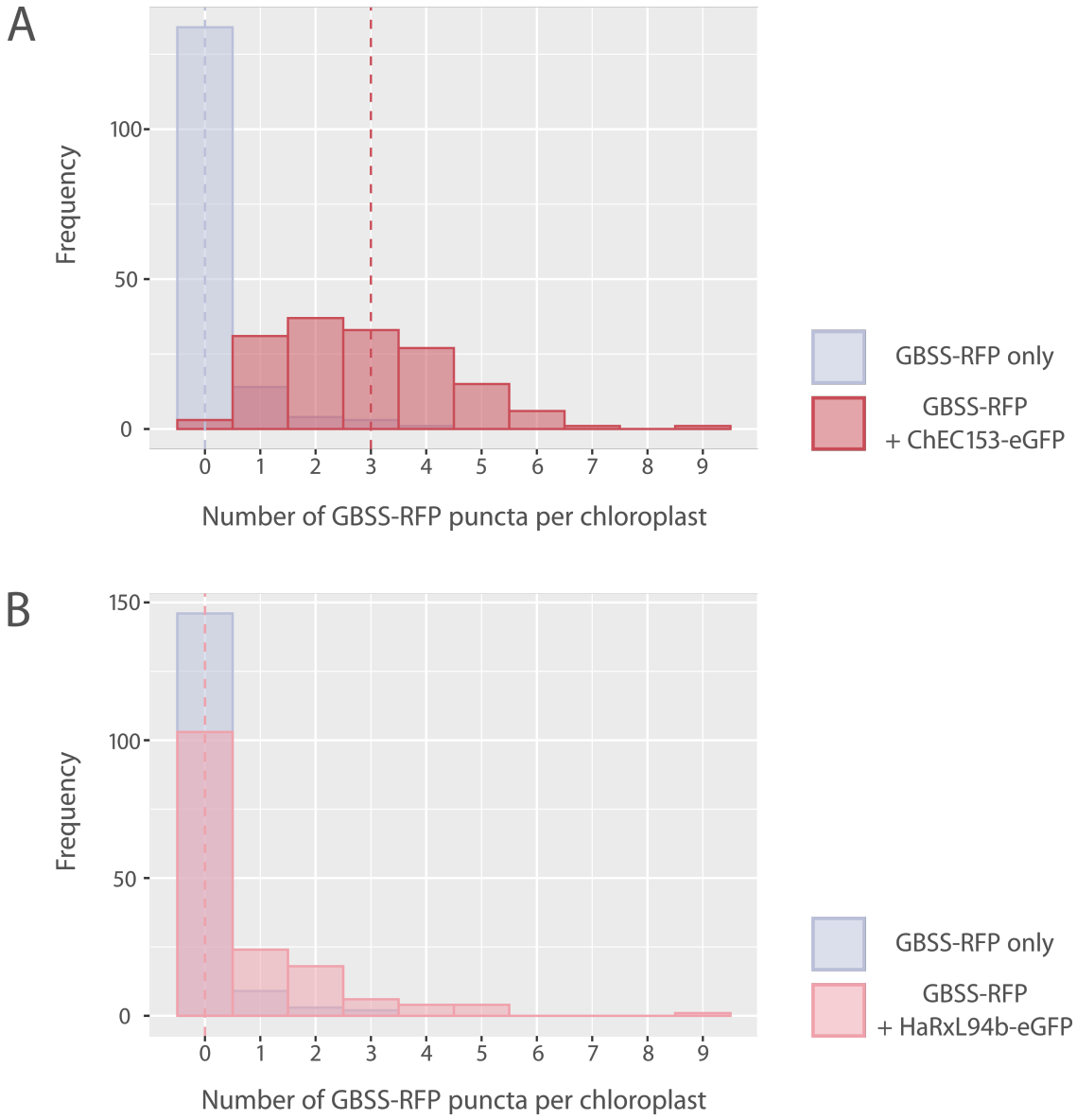


Figure 3.13: ChEC153 induces the formation of GBSS-RFP puncta. Quantification of GBSS-RFP puncta observed during *Agrobacterium*-mediated transient expression in *N. benthamiana*. **(A)** Histogram showing the quantification of the number of GBSS-RFP puncta per chloroplast for tissues expressing *35S::GBSS-RFP* only (blue) or *35S::GBSS-RFP* + *35S::ChEC153-eGFP* (red). **(B)** Histogram showing the quantification of the number of GBSS-RFP puncta per chloroplast for tissues expressing *35S::GBSS-RFP* only (blue) or *35S::GBSS-RFP* + *35S::HaRxL94b-eGFP* (pink). Data represent three pooled independent biological replicates, with 11 separate z-stack images collected for each, and the puncta from 5 chloroplasts from each image quantified where possible. Dashed lines within histograms represent the median puncta number per chloroplast for each sample.

3.3.8 ChEC153-eGFP does not appear to be strongly associated with starch granules

The co-localisation of ChEC153-eGFP with GBSS-RFP puncta seen in **Figure 3.12** raises the question whether the putative effector interacts directly with starch granules. In order to test this, I used biochemical fractionation to see whether ChEC153-eGFP co-purifies with starch granules. I purified starch from rosettes of Col-0 and from an Arabidopsis line expressing *ChEC153-eGFP* (*ChEC153-eGFP* 2-9, one of two *35S::ChEC153-eGFP* lines generated by Dr Joanna Jennings). Purified starch was boiled in 1× loading buffer for SDS-PAGE and Western blotting with both anti-GFP and anti-GBSS (Seung et al., 2015) antibodies. As GBSS is embedded within the starch granule, it was used as a positive control. Alongside these starch-bound protein samples, whole leaf extracts were also isolated from Arabidopsis rosettes of lines Col-0, and *ChEC153-eGFP* lines 2-5 and 2-9, to test whether *ChEC153-eGFP* was detectable. Samples of both the soluble and insoluble fractions were retained, and the remainder of the soluble fraction used for immunoprecipitation with anti-GFP beads. Proteins were separated by SDS-PAGE and probed with an anti-GFP antibody. Western blots from these two experiments (separation of proteins from rosettes (**Figure 3.14 A**), and from purified starch (**Figure 3.14 B**)) are presented together in **Figure 3.14.**

GBSS was detected very strongly in both starch samples, but no anti-GFP signal was seen in either starch sample. This suggests that ChEC153-eGFP is not tightly associated with the starch granule in the way that GBSS – an integral protein embedded within and bound to the granule – is. ChEC153-eGFP is enriched in the immunoprecipitation (IP) samples and also detected in both the soluble and insoluble fractions of the protein extractions from *ChEC153*-expressing plants. ChEC153-eGFP is expected to be approximately 84.67 kDa but is detected here as two bands between 70 and 100 kDa. This shift in protein size may be due to post-translational modifications. There was an enrichment of ChEC153-eGFP in the insoluble fraction relative to the soluble fraction, which may suggest an inefficiency in the solubilisation. To summarise, while ChEC153-eGFP can be detected in leaves of Arabidopsis lines expressing the putative effector in both soluble and insoluble fractions, it could not be detected in samples of purified starch.

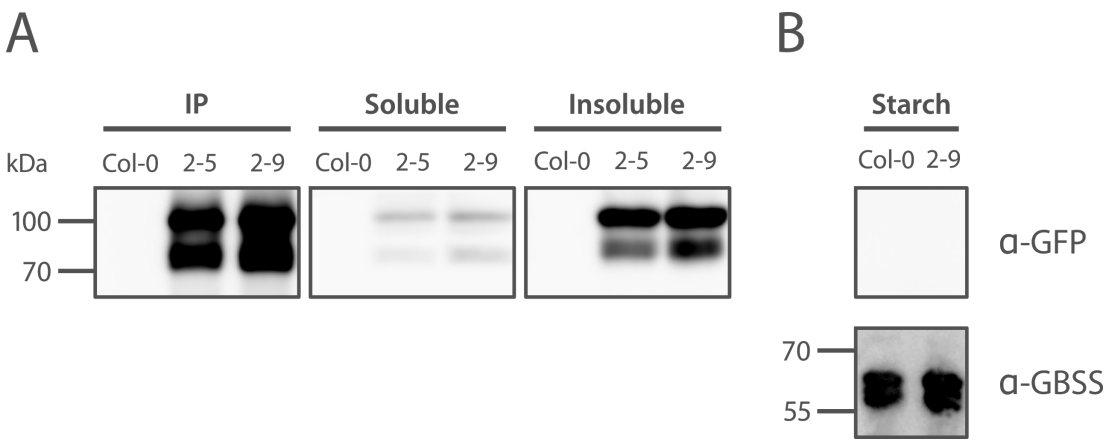


Figure 3.14: **ChEC153-eGFP is not detected in starch isolated from putative effector-expressing Arabidopsis rosettes. (A)** Western blot analysis of soluble (input) and insoluble (pellet) fractions of proteins extracted from Arabidopsis rosettes of Col-0 and effector-expressing (2-5 and 2-9) lines. Immunoprecipitation (IP) of the soluble fraction enriches the presence of the putative effector (IP, input and pellet samples loaded equally). **(B)** Starch extracted from Col-0 and *ChEC153-eGFP*-expressing line 2-9 was separately probed for the presence of ChEC153-eGFP. Sample proteins were separated by SDS-PAGE prior to being probed with α-GFP antibodies or α-GBSS antibodies as indicated.

3.3.9 The localisations of ChEC153 and HaRxL94b do not require MRC, SS4, or starch

Based on the co-localisation of ChEC153 and HaRxL94b with MRC, I thought it possible that these putative effectors target MRC specifically, and that this targeting determines the punctate effector localisation. By extension, given that MRC co-localises and interacts with SS4 (Vandromme et al., 2019, Chen, 2022), SS4 presented an additional target to screen for. Further, I was interested in whether the putative effector puncta were starch-dependent structures based on their observed co-localisation with the GBSS-RFP puncta seen in **Figure 3.12**. To determine whether the localisation of the putative effectors to puncta requires the presence of MRC, SS4 or starch, I used particle bombardment to transform cells of Arabidopsis lines Col-0, *mrc*, *ss4* and *pgm* (an essentially starch-free mutant) with vectors for expression of either ChEC153-eGFP or HaRxL94b-eGFP alongside an ER-localised RFP marker for transformation. If these proteins or starch were required for targeting of the putative effector to sites of starch granule initiation, the punctate localisation of ChEC153-eGFP and HaRxL94b-eGFP would be disrupted in these mutant lines.

I imaged cells expressing the putative effector constructs using confocal microscopy, with effector-eGFP puncta forming in all lines (**Figure 3.15**). Additional faint, diffuse GFP signal in the nucleus typical of HaRxL94b-eGFP expression in *N. benthamiana* was also seen in several cells for the HaRxL94b-eGFP localisation in these Arabidopsis lines (**Figure 3.15**, white arrowhead). Thus, I have

demonstrated that the formation of ChEC153-eGFP or HaRxL94b-eGFP puncta in the chloroplast is not dependent on the presence of either MRC, SS4, or starch.

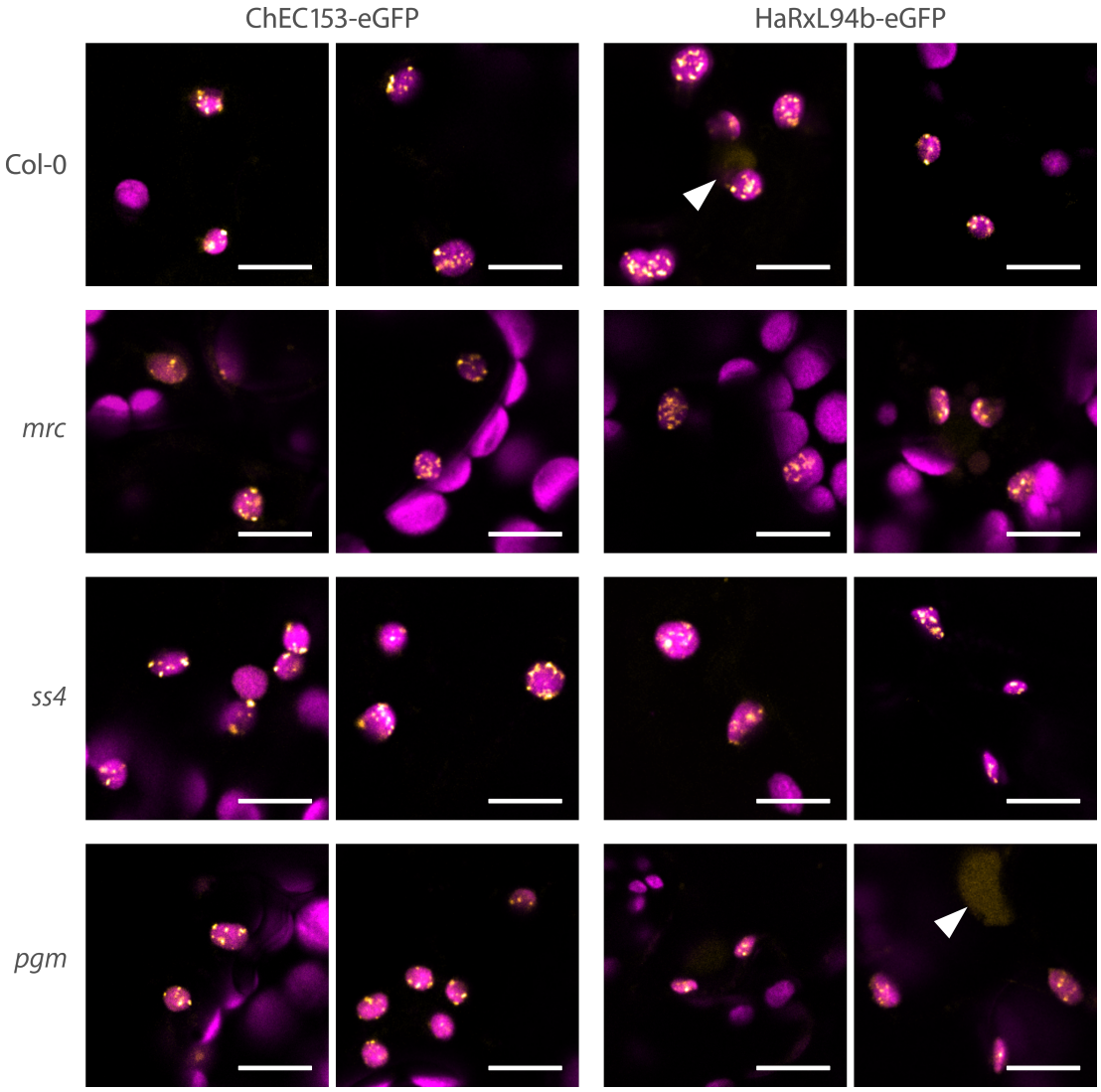


Figure 3.15: Localisations of ChEC153 and HaRxL94b in Arabidopsis lines Col-0, *mrc*, *ss4,* and *pgm*. Particle bombardment was used to transform leaves from each line with a vector for the expression of either effector protein: *35S::ChEC153-eGFP* and *35S::HaRxL94b-eGFP*. GFP signal is shown in yellow, with chlorophyll autofluorescence in magenta. White arrowheads indicate nuclear localisation of HaRxL94b-eGFP. Scale bars, 10 µm.

3.3.10 Expression of *ChEC153-eGFP* in Arabidopsis and *N. benthamiana* does not have a clear impact on the total starch or sugar content of leaves

We expect that *C. higginsianum* takes up carbon as soluble sugars, and so I hypothesise that the pathogen may benefit from a lowered relative starch content with concomitant elevation of soluble sugar levels. I therefore sought to assay for any impact of heterologous *ChEC153*

expression on the starch and sugar content of leaves, to see whether this could correlate the presence of ChEC153 with any potential benefit to the pathogen in terms of carbon availability.

3.3.10.1 Quantification of total starch content

In order to look for an effect of *ChEC153* expression on the plant in terms of the starch content, I extracted starch from Arabidopsis rosettes of Col-0 and *ChEC153-eGFP* lines, and from *N. benthamiana* leaf samples during transient expression of either a *p19*-only control or *ChEC153-eGFP* + *p19*.

Starch content of Arabidopsis stable lines expressing *ChEC153-eGFP*

I measured the starch content of Arabidopsis lines Col-0 and two independent *35S::ChEC153-eGFP* stable lines. I harvested whole rosettes of plants and extracted their starch, which I quantified by an enzymatic assay, with results presented in **Figure 3.16.** As starch synthesis and starch turnover rates are heavily dependent on photoperiod, I quantified starch from plants grown in two different growth conditions: a 12-hour photoperiod (**A**), or 16-hour photoperiod (**B**). I observed no statistically significant differences between Col-0 and *ChEC153-eGFP* expressing Arabidopsis rosettes in terms of their total starch content at the end-of-day or end-of-night under either growth condition.

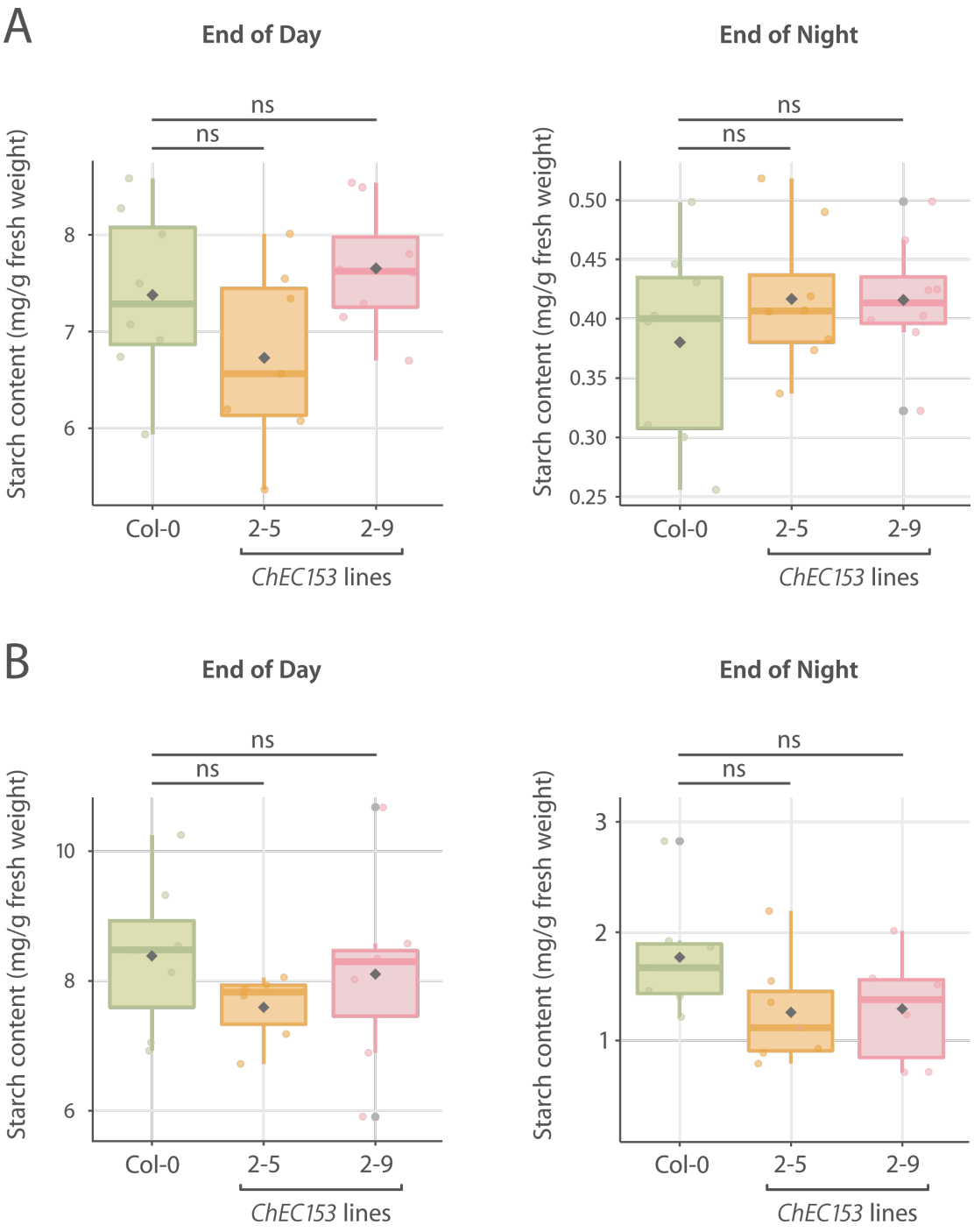


Figure 3.16: **Starch content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines.** Starch (mg/g fresh weight) measured at the end of the day and end of the night. **(A)** Plants grown standard growth conditions (12 h: 12 h, light: dark). **(B)** Plants grown under high light (400 µmol photons $m^{-2}s^{-1}$) and long day (16 h: 8 h, light: dark) growth conditions. Statistical significance determined by unpaired Welch two sample t-tests; ns $p > 0.05$ (not significant).

Starch content of *N. benthamiana* leaf tissue transiently expressing *ChEC153-eGFP*

Despite seeing no difference in the starch content of Arabidopsis leaves expressing *ChEC153*, I also sought to quantify the starch content of *N. benthamiana* leaf tissue during *Agrobacterium*-mediated transient expression of *ChEC153*. This would allow me to see whether an impact of ChEC153 on starch content would be evident in this pulse-chase context in which the induced formation of GBSS-RFP puncta was observed, or whether the same result would be shared between the two plant systems. The end-of-day starch content of *N. benthamiana* leaf tissue transiently expressing *ChEC153-eGFP* is not statistically significantly different to a *p19*-only control (**Figure 3.17**), in agreement with the lack of difference seen in Arabidopsis samples (**Figure 3.16 A** and **B**).

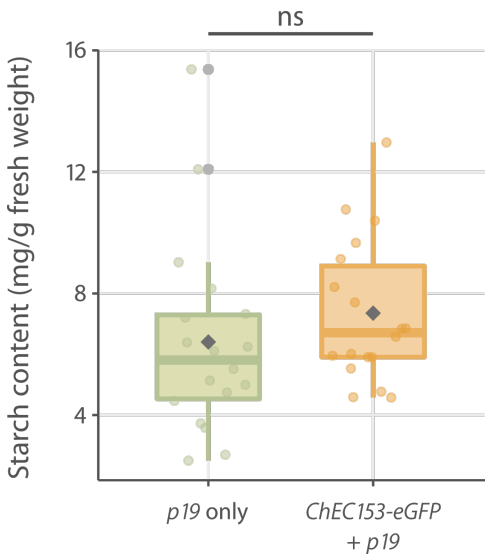


Figure 3.17: **End of day total starch content of *N. benthamiana* leaf tissue transiently expressing *p19* only or *35S::ChEC153-eGFP + p19*.** Samples were collected three days post infiltration with *Agrobacterium*. Control samples: *p19* at OD_{600nm} = 0.5; ChEC153-eGFP samples: *p19* at OD_{600nm} = 0.2 and *35S::ChEC153-eGFP* at OD_{600nm} = 0.3. Difference in total starch content (mg/g fresh weight) evaluated by Welch two sample t-test and deemed to be not significant (ns), $p > 0.05$.

3.3.10.2 Quantification of soluble sugars

Soluble sugar content of Arabidopsis stable lines expressing *ChEC153-eGFP*

In addition to the starch quantification, I also quantified sugar content of the Arabidopsis lines from the same samples as were used to quantify starch. The results presented here are from plants grown in high light, long day (18-hour photoperiod) conditions. The trends seen for the quantification of soluble sugars in Arabidopsis lines expressing *ChEC153-eGFP* relative to Col-0 remain consistent in either growth condition. No significant differences in the glucose, fructose or sucrose contents were observed between effector-expressing lines and the wild-type (**Figure 3.18 A**, **B** and **C**). The measured values for end-of-day glucose content in *ChEC153-eGFP* line 2-5 were all negative, and therefore deemed not to be reliable. Differences between these data and the corresponding Col-0 data were therefore not determined (**Figure 3.18 A**). In separate experiments using plants grown with a 12-hour photoperiod, end-of-day glucose values remained low, but were positive. For these data, no differences between the glucose content of Col-0 and *ChEC153*-expressing plants were seen (**Figure 3.19**). I conclude that the expression of *ChEC153-eGFP* in Arabidopsis does not have an impact on the total sugar levels.

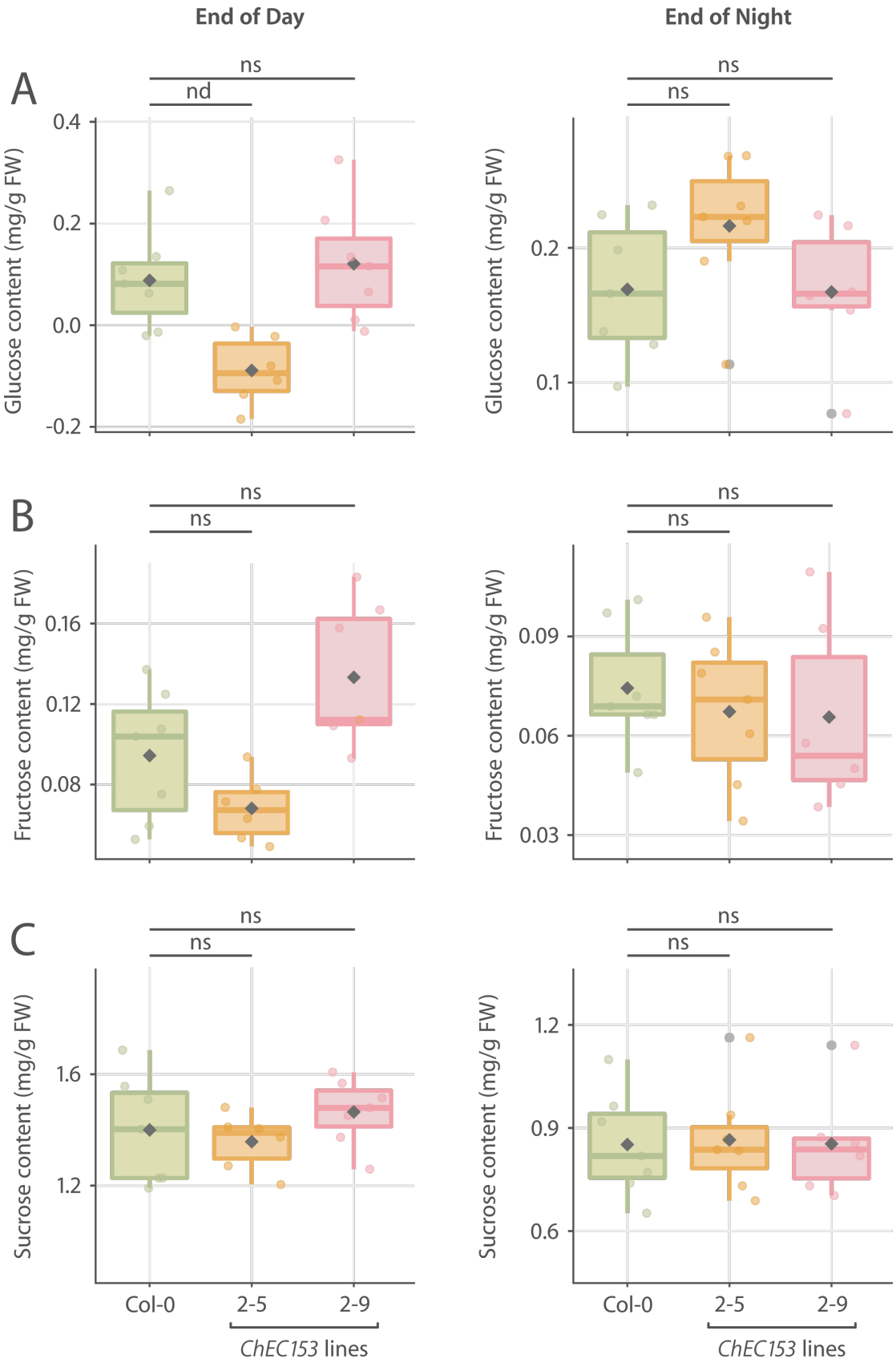


Figure 3.18: Soluble sugar content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines. Soluble sugars (mg/g fresh weight) measured at the end of the day and end of the night. **(A)** Glucose content. **(B)** Fructose content. **(C)** Sucrose content. Plants grown under high light (400 µmol photons $m^{-2}s^{-1}$) and long day (16 h: 8 h, light: dark) growth conditions. Statistical significance determined by unpaired Welch two sample t-tests; ns $p > 0.05$ (not significant); nd (not determined).

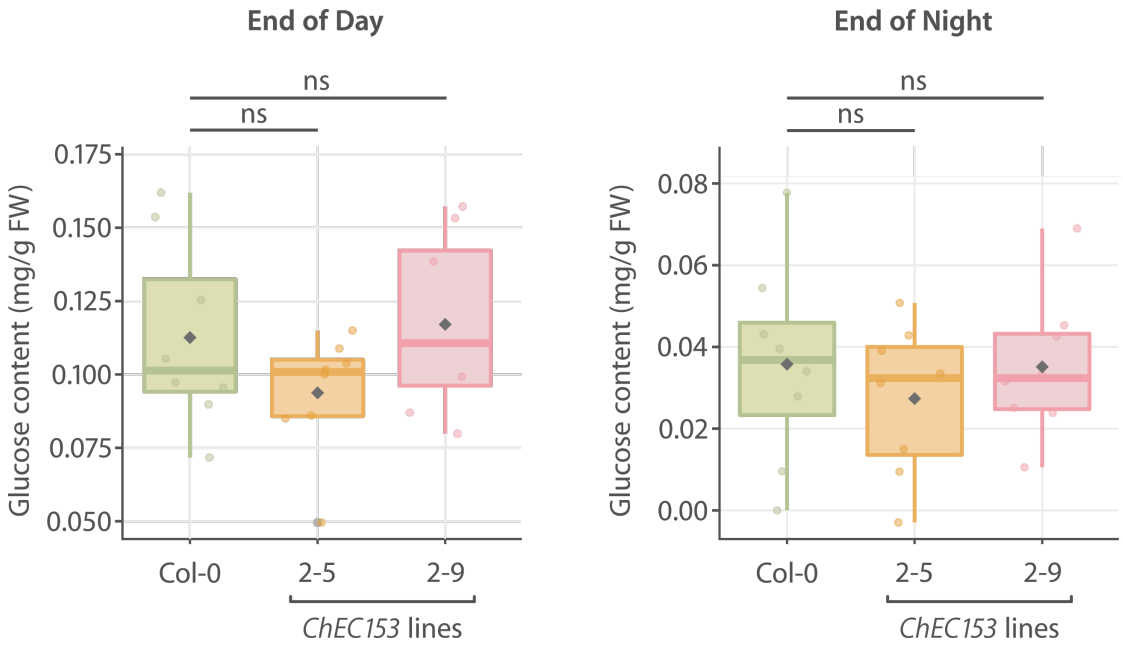


Figure 3.19: **Glucose content of Arabidopsis Col-0 and *35S::ChEC153-eGFP*-expressing lines grown with a 12-hour photoperiod.** Glucose content (mg/g fresh weight) measured at the end of the day and end of the night. Plants grown under 12 h: 12 h, light: dark. Statistical significance determined by unpaired Welch two sample t-tests; ns $p > 0.05$ (not significant).

Soluble sugar content of *N. benthamiana* leaf tissue transiently expressing *ChEC153-eGFP*

I quantified the soluble sugars glucose, fructose, and sucrose from *N. benthamiana* leaf tissue transiently expressing either only the *p19* silencing suppressor, or *ChEC153-eGFP* in addition to *p19*. The results of this experiment are presented in **Figure 3.20.** I conclude that there is no clear impact of *ChEC153* expression on the levels of soluble sugars in *N. benthamiana*, as no significant differences in soluble sugar content were observed between control and *ChEC153*-expressing leaf tissue.

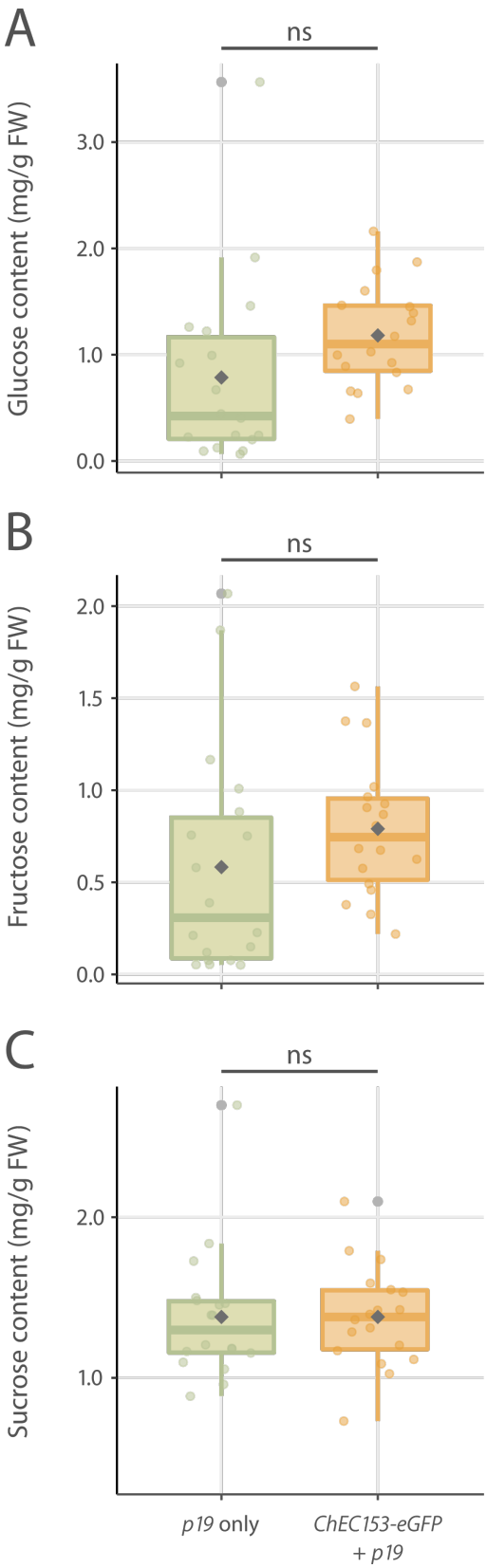


Figure 3.20: **End of day soluble sugar content of *N. benthamiana* leaf tissue transiently expressing *p19* only or *35S::ChEC153-eGFP + p19*.** Soluble sugar content (mg/g fresh weight (FW)) for samples collected three days post infiltration with *Agrobacterium*. Control samples: *p19* at OD_{600nm} = 0.5; ChEC153-eGFP samples: *p19* at OD_{600nm} = 0.2 and *35S::ChEC153-eGFP* at OD_{600nm} = 0.3. **(A)** Glucose content. **(B)** Fructose content. **(C)** Sucrose content. Differences in sugar content were evaluated by Welch two sample t-tests and deemed to be not significant (ns).

From the data presented in **Figure 3.16** to **Figure 3.20**, I conclude that there is no impact of *ChEC153* expression on the total starch or soluble sugar contents in stable Arabidopsis lines or upon transient expression in *N. benthamiana*. It should be noted that these experiments consisting of *in planta* heterologous expression of a single effector gene will not necessarily be reflective of what would be seen in an infection context. It is possible that during infection, ChEC153 acts in combination with other effector proteins to manipulate carbon availability. Further, any changes may occur in a more localised manner or on a smaller scale than could be observed using this approach.

3.3.11 The starch granules of Arabidopsis leaves expressing *ChEC153-eGFP* are slightly larger than those of the wild-type

That there is no clear correlation between the expression of *ChEC153* and total starch content in Arabidopsis does not rule out a potential impact of the effector on granule initiation. In some genotypes with perturbed starch granule initiation, such as *mrc* Arabidopsis mutants, the total starch content of leaves remains comparable to the wild-type. While the number of starch granules per chloroplast in these lines is decreased, the size of starch granules is increased (Vandromme et al., 2019, Seung et al., 2018). Therefore, I analysed the size of starch granules in *ChEC153*-expressing Arabidopsis leaves to see whether expression of *ChEC153* influences the granule size relative to Col-0. I also aimed to investigate whether a secondary population of small starch granules would be apparent in the effector-expressing lines, correlating with the GBSS-RFP puncta observed during transient expression of *ChEC153* in *N. benthamiana* (**Figure 3.12**). I hypothesised that smaller starch granules, corresponding to a larger surface area to volume ratio might benefit the rapid degradation of starch, rendering the host carbon more available to the pathogen.

3.3.11.1 Particle size analysis of purified starch

In order to evaluate the size of the starch granules in Arabidopsis lines expressing *ChEC153-eGFP*, I purified starch from rosettes of these lines alongside Col-0 and *mrc*, pooling a total of 60 rosettes per genotype. I carried out particle size analysis by resistive pulse sensing with a Coulter counter to obtain the granule size distributions of these lines, which are presented in **Figure 3.21**.

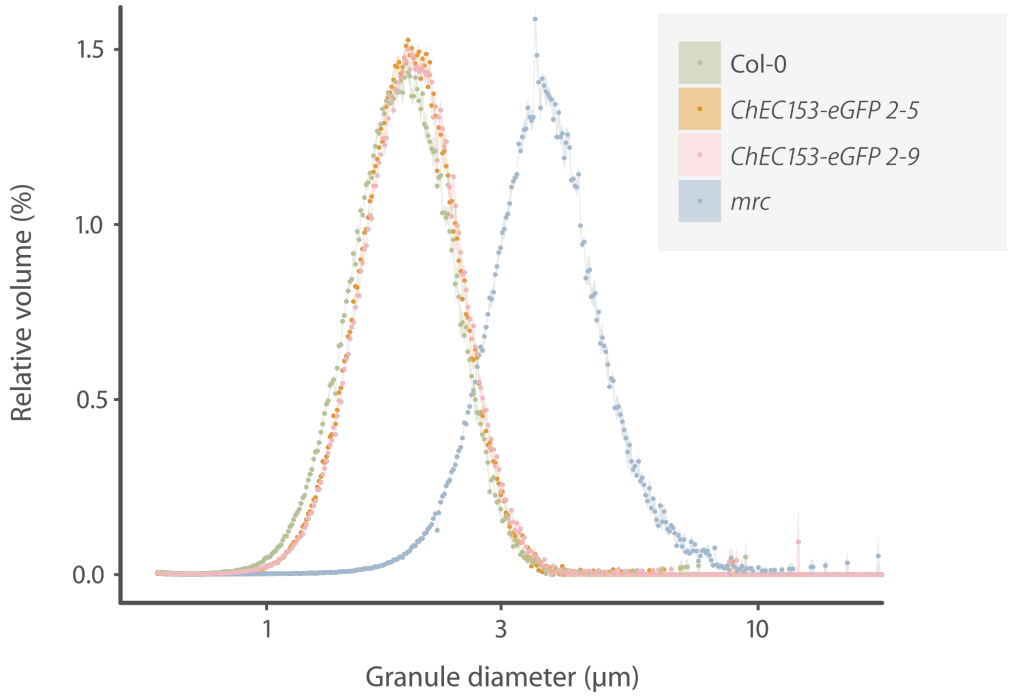


Figure 3.21: Size distributions showing the diameters of starch granules purified from Col-0, *ChEC153-eGFP*-expressing lines, and *mrc*. Samples were harvested at the end of the day, pooling 60 rosettes for each genotype. Data are presented as relative percentage volume against granule diameter (µm), with points representing mean averages of three technical replicates for each sample. Shaded regions represent standard error.

Average starch granule diameters for each sample were extracted from the raw data by lognormal curve fitting and used for statistical analysis. The means of the three technical replicates for each line were used to estimate the average granule diameter. These data are presented in Table 3.2.

Table 3.2: Average starch granule diameters of starch purified from Arabidopsis lines Col-0, *ChEC153*-expressing lines, and *mrc*. Average granule diameters (µm) presented are means of three technical replicates, with measurements extracted from raw particle size data using unimodal, lognormal curve fitting analyses.

Line	Average granule diameter (µm)	Standard error
Col-0	1.958	0.002
ChEC153 2-5	2.016	0.003
ChEC153 2-9	2.040	0.003
mrc	3.782	0.008

I statistically evaluated the differences in mean starch granule diameters for each line using a linear model. A linear model for the average granule diameter for three technical replicates for each line was subjected to a one-way ANOVA (*F* value: 39045, p-value of *F* statistic: $< 2\times10^{-16}$) followed by pairwise multiple comparisons using the Tukey method. The mean granule diameters of all lines were found to be statistically significantly different to one another ($p < 0.05$).

A small difference in granule size is seen between the two *ChEC153*-expressing lines. While this difference between mean average granule sizes was statistically significant ($p = 0.0234$), it represents only an average increase in mean granule diameter between *ChEC153* lines 2-5 and 2-9 of 24 nm (a 1.19% increase). This difference is likely only significant due to measurements being technical replicates, and the high accuracy of the Coulter counter in measuring particle size (as reflected by the small standard error for each line). The granules of the *ChEC153*-expressing lines appear slightly larger than the wild-type on average, with increases of 58 and 82 nm (or 2.96% and 4.19%) on average for lines 2-5 and 2-9, respectively ($p = 0.0001$ and $p < 0.0001$ in pairwise comparisons to Col-0). Whether differences in starch granule diameters on this scale would be of biological relevance is not clear. As expected, and previously reported (Seung et al., 2018, Vandromme et al., 2019), the *mrc* mutant has considerably larger starch granules than the wild-type, with the average granule diameter increasing by 1824 nm (a 93.16% increase, $p < 0.0001$). In comparison to the scale of starch granule size phenotypes that are generally deemed of interest in the field, the size differences seen between Col-0 and *ChEC153*-expressing lines seen here are small. However, I conclude that ChEC153 slightly increases the size of starch granules in Arabidopsis leaves.

Using a 30 µm aperture, the Coulter counter has a lower limit of size detection of 0.6 µm. This limits interference of electronic noise (which drastically increases when particle diameter is less than 2% of the aperture diameter) and may minimise unwanted quantification of small debris particles contaminating samples. However, it is possible that should the GBSS-RFP puncta seen in Figure 3.12 C in the presence of ChEC153-eGFP represent small starch granules below this detection threshold, they would not be detectable using this method. In order to check whether the GBSS-RFP puncta are likely to be within the range of detection by the particle size analyser, I measured the maximum visible length of a randomly selected subset of GBSS-RFP puncta in my microscopy images, with measurements shown in Figure 3.22. The mean maximum diameter of GBSS-RFP puncta I measured was 0.617 µm. I therefore conclude that the sizes of puncta observed from my images are indicative of starch granules which may not be detected by the particle size analyser, as their size is around the lower detection threshold. Measurement of starch granules in this range using the Coulter counter is impeded by the technical limitations of the machine using the 30 µm aperture.

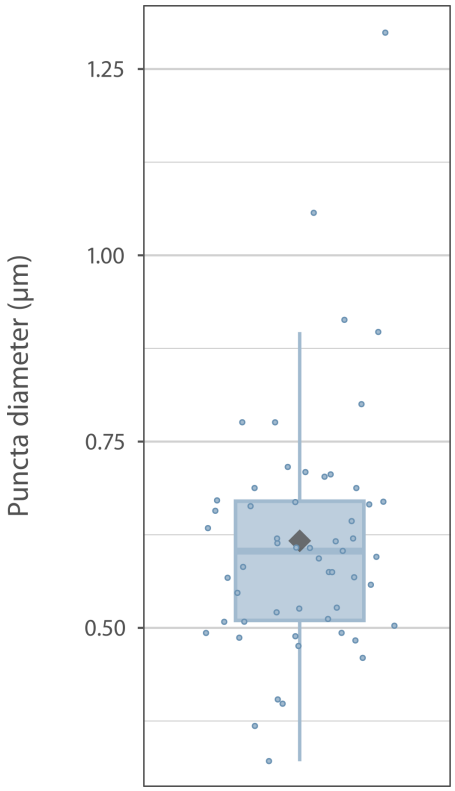


Figure 3.22: GBSS-RFP puncta diameter (µm) during co-expression of *35S::GBSS-RFP* and *35S::ChEC153-eGFP* in *N. benthamiana*. Puncta diameters were measured manually in ImageJ at the maximum visible diameter from confocal microscopy images. Datapoints represent individual puncta. Mean puncta diameter shown as a grey diamond, $\bar{x}$ = 0.617 µm.

3.3.11.2 Analysis of purified starch using scanning electron microscopy

In order to probe the purified starch for granules of diameters below the particle size analyser detection threshold, I sought to look at the purified starch using microscopy. Samples of starch purified from the Col-0, *ChEC153 2-5*, *ChEC153 2-9*, and *mrc* lines were used for scanning electron microscopy (SEM) by Dr David Seung. Representative SEM images are presented in **Figure 3.23**. From these images, as expected the *mrc* starch granules are considerably larger than the wild-type, but the *ChEC153*-expressing line starch granules are essentially indistinguishable from the Col-0 control. These results complement those seen in the particle size analysis in **Figure 3.21**, with the small increase in granule diameter suggested for *ChEC153*-expressing lines likely too small to be obvious from micrographs. No population of small starch granules is apparent in the *ChEC153*-expressing lines using this methodology of purification and SEM analysis.

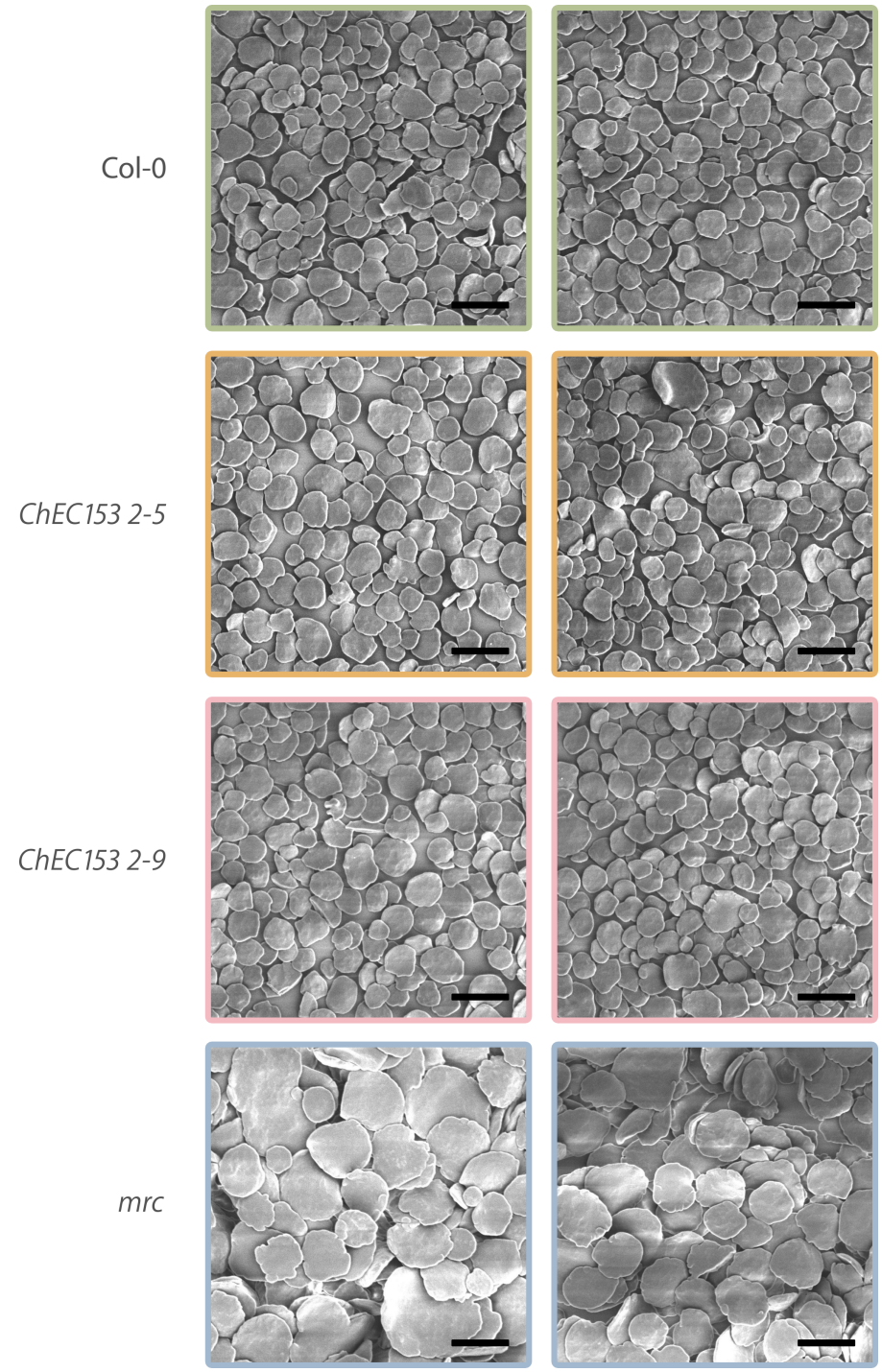


Figure 3.23: Scanning electron micrographs of starch purified from Col-0, *ChEC153*-expressing lines 2-5 and 2-9, and *mrc*. Scale bars, 5 µm. Images were acquired by Dr David Seung.

3.4 DISCUSSION

3.4.1 What is the nature of the punctate localisation of ChEC153 and HaRxL94b?

Despite deriving from relatively distantly related pathogens and lacking any considerable sequence similarity or shared domain annotations, ChEC153 and HaRxL94b showed uncannily similar localisations upon transient expression with C-terminal fluorescent tags (**Figure 1.5** and **Figure 1.6**), localising to several discrete puncta within each chloroplast. A number of proteins are known to localise to chloroplastic puncta, and these may reflect spatially and functionally disparate sub-chloroplastic localisations. Co-localisation experiments confirmed the shared localisation of ChEC153 and HaRxL94b to the same set of puncta (**Figure 3.3**). Based on this observation, it appears possible that they are targeting the same host proteins or processes. Further co-localisation experiments confirm that ChEC153 and HaRxL94b co-localise with the key starch granule initiation protein MRC (**Figure 3.8**). From these observations, it appears that ChEC153 and HaRxL94b may have evolved convergently to both target MRC, and likely the same aspect of starch granule initiation.

Both ChEC153 and HaRxL94b are observed to co-localise with the wheat susceptibility factor and RNA-binding protein, TaCSP41a (**Figure 3.9**), raising the question as to whether CSP41a localises to the site of starch granule initiation. It is possible that multiple processes are carried out at these punctate sub-chloroplastic regions, and that while CSP41a has not been implicated in starch granule initiation, the granule initial sites may also host RNA processing machinery. Further, the association of TaCSP41a with susceptibility to *Pst* (Corredor-Moreno et al., 2022) further points to potential involvement of these puncta in infection-related processes.

While the putative effector proteins reliably co-localise with one another, MRC, and TaCSP41a, they were not consistently observed to co-localise with the nucleoid marker $PEND_{1-88}$ (**Figure 3.10**). This partial co-localisation indicates that the putative effector localisations may not be exclusively reflective of nucleoids. Furthermore, current evidence suggests that MRC itself is not nucleoid-associated (Seung et al., 2018). MFP1, another protein critical for normal starch granule initiation and displaying a punctate localisation within the chloroplast (Seung et al., 2018), is associated with thylakoid membranes and has DNA-binding capacity, leading to the suggestion that it is associated with nucleoids (Jeong et al., 2003). The partial co-localisation of the putative effectors with $PEND_{1-88}$ raises questions regarding any potential link between nucleoids and the sites of starch granule initiation, which is currently unclear. It is possible that proteins may play multiple roles relating to both of two distinct punctate processes. Definitive experiments co-expressing MRC and $PEND_{1-88}$

would be valuable in seeking to elucidate any link between granule initial puncta and nucleoid puncta.

It is likely that there are multiple sets of distinct, punctate localisations within the plant chloroplast. Wang and colleagues explore this concept in their 2023 paper looking at punctate chloroplastic proteins of *Chlamydomonas* (Wang et al., 2023c). They identify 11 sub-chloroplastic punctate localisations which they deem to be discrete based on the number and size of puncta observed during confocal imaging of various fluorescently-tagged *Chlamydomonas* proteins. My investigation of the punctate localisation of ChEC153 and HaRxL94b focuses on the presence or absence of co-localisation of fluorescently-tagged proteins, rather than inferences from puncta number, size or location. It is unclear whether subsets of sub-chloroplastic puncta may be transient in nature. From my data, I conclude that the putative effectors ChEC153 and HaRxL94b localise to punctate sites of starch granule initiation, which may partially or transiently correspond to nucleoids.

Starch granule initiation appears to be a dynamic process and involves a number of proteins. While here the localisation of the putative effectors is examined in the context of co-localisation with MRC, how this relates to the rest of the starch granule initial remains unclear. MRC interacts with SS4 and PTST2, and PTST2 also interacts with MFP1, which itself is believed to associate with the thylakoid membrane. The concentration of glucan-elongating activity (for example of SS4, the correct localisation/function of which appears to be facilitated by MRC) and glucan binding activity (for example conferred by PTST2) into a given location such as a punctum or granule initial may aggregate glucans efficiently in one place and promote the formation of crystalline starch granules. Interestingly, the localisations of PTST2 and MFP1 at chloroplastic puncta appear distinct to those seen for MRC and SS4 (as well as ChEC153 and HaRxL94b), in that the PTST2 and MFP1 puncta are generally seen to be more numerous and less discrete. Further, analysis of Arabidopsis protein extract fractions suggests a separation, with MRC found in the stromal fraction, PTST2 detected in both stromal and thylakoid fractions, and MFP1 in the thylakoid fraction (Seung et al., 2018). Protein extracts from *ChEC153*-expressing plants were crudely fractionated into soluble and insoluble fractions, with ChEC153 being detected in both (**Figure 3.14**). While this may represent an incomplete extraction resulting in protein retention in the pellet, it is also possible that ChEC153 has some association, perhaps dynamically, with the thylakoid membrane.

ChEC153 and HaRxL94b do not appear to interact directly with starch. No clear carbohydrate binding modules were predicted in either putative effector. Despite a co-localisation between the effector puncta and GBSS-RFP puncta (**Figure 3.12**), the punctate localisation of neither protein relies on the presence of starch, as seen by their localisation in the starch-free *pgm* mutant (**Figure 3.15**). Further, ChEC153-eGFP was not detected in starch purified from *ChEC153*-expressing

Arabidopsis rosettes, despite the protein separately being detected in protein extracts from this line (Figure 3.14). If ChEC153 does form an association with starch, it is not sufficient to retain the protein throughout the starch-extraction process. Therefore, it seems unlikely that either putative effector functions by directly binding starch.

The targeting of ChEC153 and HaRxL94b to the starch granule initial does not require MRC or SS4, as evidenced by their localisation in *mrc* and *ss4* lines (Figure 3.15). It is possible that the putative effectors may interact with both MRC and SS4, and thus rely upon the presence of only either one of these proteins for correct localisation. This may be determined by looking at the localisation of ChEC153 in a double mutant *ss4mrc* line. However, in the absence of MRC, SS4 typically displays a less discrete, more patchy or diffuse localisation than when co-expressed with MRC (Chen, 2022). Thus, in *mrc* mutant lines, SS4 is likely not to localise to such discrete puncta, and so the observation of discrete ChEC153 and HaRxL94b puncta in this line could suggest that the effector puncta formation may not require either MRC or SS4.

Alongside ChEC153 and HaRxL94b, several other putative effector proteins have been identified to localise to chloroplastic puncta, including one from *Pst* (Andac et al., 2020) and one from *Melampsora lini* (Petre et al., 2016a). Whether these other putative effector proteins are localising to the same puncta as ChEC153 and HaRxL94b has not been explored, but this could exemplify similar targeting in a number of pathogen species possibly highlighting some importance of these effectors and/or effector targets in infection by filamentous plant pathogens with hemibiotrophic or biotrophic lifestyles.

3.4.2 Does ChEC153 impact total starch content or starch granule size?

Despite the implication of ChEC153 in starch granule initiation and starch synthesis from its localisation, the total quantities of starch and soluble sugars in leaves remained unchanged in the presence of ChEC153 (Figure 3.16 to Figure 3.20). Thus, it seems that any impact on starch and carbon availability is less obvious than an absolute change in the total starch or sugar content. Any theoretical impacts of ChEC153 on starch or sugar levels may be transitory, infection-specific or highly localised, or may pertain to an alteration in the structure of starch granules or their relative size and number within the chloroplast.

According to particle size analysis of purified starch, the presence of ChEC153 results in a small increase in starch granule size while the overall starch granule size distribution remains indistinguishable from the wild-type (Figure 3.21 and Table 3.2). Whether an increase in starch granule size on this scale is of biological relevance remains to be determined, and the frequency of these small starch granules relative to the typical granules is unknown. To remain consistent with my observation that the total starch content of these lines is unchanged, the relative number

of larger starch granules per chloroplast may decrease slightly in the presence of ChEC153. In this way, the same quantity of substrate may be differentially distributed into small and large granules in the presence of ChEC153. This could be investigated by microscopic analysis of Periodic Acid-Schiff-stained sectioned leaf tissues, allowing an estimation of the number of starch granules within each chloroplast, and potentially the observation of these presumed small starch granules. Similar phenomena are apparent in Arabidopsis mutant lines such as *mrc* where total starch content remains relatively similar to the wild-type, but the starch is contrastingly distributed between fewer, larger starch granules (Vandromme et al., 2019, Seung et al., 2018). The overall morphology of starch granules purified from *ChEC153*-expressing Arabidopsis was indistinguishable from the wild-type using SEM analysis (**Figure 3.23**).

3.4.3 ChEC153-induced formation of GBSS puncta: formation of small starch granules or mislocalisation of the GBSS protein?

I observed an induced formation of GBSS-RFP puncta in the presence of ChEC153-eGFP during transient co-expression in *N. benthamiana* (**Figure 3.12**). It is most likely that this represents the induced formation of a subset of small starch granules by ChEC153. Critically, GBSS is known to be unstable when not bound to a starch granule (Seung et al., 2015, Smith et al., 2004), rendering it an effective starch marker, and so its appearance in these discrete puncta likely indicates underlying glucans/small starch granules. Based on my observations of ChEC153-induced GBSS-RFP puncta, I hypothesise that they represent a small, secondary population of starch granules at around the lower detection limit of the particle size analyser. Whether such small starch granules would be efficiently purified using the Percoll cushion density gradient method I applied is unclear – as they appear to be considerably smaller (~0.6 µm diameter, **Figure 3.22**) than the typical Arabidopsis starch granules (~2 µm diameter), the population of small granules may have been separated and discarded from the bulk of larger "normal" granules in the purification process. Further, whether these GBSS-RFP puncta represent insoluble starch granules, or soluble aggregates of glucans not yet developed into insoluble semi-crystalline structures, is not clear. If they are not insoluble, these small "granules" could have been lost in the purification process for that reason. This may explain the lack of observable small starch granules in my particle size analysis and SEM images (**Figure 3.21** and **Figure 3.23**). Further, the number of these small starch granules in the tissues from which I purified starch may be comparatively low: ChEC153-induced GBSS-RFP puncta were observed in epidermal cells in which the putative effector was clearly visible, but other cell layers were not observed their starch may be unchanged. If the number of small starch granules is considerably lower than the "normal" starch granules, in the particle size and SEM analyses, they may be lost against this background of "normal" starch. Further, I found the Arabidopsis lines expressing *35S::ChEC153-eGFP* to show expression of the putative effector in

many, but not every, epidermal cell – in some cells eGFP signal was not apparent, which may further limit the proportion of these small starch granules.

To verify the nature of these small starch granules, I could use a staining method to complement the GBSS-RFP genetic marker used here. I could carry out Periodic Acid-Schiff staining of starch in semi-thin sections of tissue and use light microscopy to collect images from which to quantify the starch granule size and the number of starch granules per chloroplast. Alternatively, Lugol staining of starch in epidermal peels may offer sufficient resolution to look at the starch granules with light microscopy having first checked for consistent *ChEC153-eGFP* expression with confocal microscopy. This would have the advantage of allowing me to collect z-stacks of images from which it may be easier to verify the size of granules. Further, a recent publication from Ichikawa and colleagues highlights the use of fluorescein to stain starch in live cell imaging (Ichikawa et al., 2023). This stain allowed the authors to visualise starch granules in Arabidopsis transgenic leaf tissue alongside GBSS-RFP, as well as visualising the starch granules of *N. benthamiana* leaf tissue (Ichikawa et al., 2023), and thereby could also be used to evaluate the starch granules of *N. benthamiana* tissue transiently expressing *ChEC153-RFP* to complement my observation with GBSS-RFP that small starch granules are being induced. Further, if I were to generate Arabidopsis transgenic lines expressing *ChEC153-RFP*, or to transiently express *ChEC153-RFP* in Arabidopsis, I would be able to assay for the presence of induced small starch granules in the host using fluorescein staining. However, as highlighted above, the GBSS-RFP starch granule marker is compelling evidence for the induced formation of small starch granules by ChEC153. These effector-induced small starch granules could represent an interesting starch phenotype for an infection context.

I hypothesise that the high surface area to volume ratio of small starch granules may be beneficial to pathogens seeking to acquire starch from the host. An increased surface area would increase the efficiency of enzymatic degradation of the starch granules, potentially releasing accessible carbon more rapidly than would be seen in the breakdown of a larger starch granule. In this way, while the total starch and sugar content is unaffected in the presence of ChEC153 (Figure 3.16 to Figure 3.20), the distribution of this starch into a subset of very small starch granules (indicated by the presence of GBSS-RFP puncta) may result in the accessibility of this comparable amount of carbon being increased in an infection context.

Whether the induced GBSS-RFP puncta definitively represent the formation of small starch granules, or a mislocalisation of GBSS to puncta with no underlying glucan chains remains to be fully elucidated, but I view the former to be more likely. These GBSS-RFP puncta were observed at three days post infiltration in the *N. benthamiana* transient expression system, and the fate of these puncta during constant, rather than pulse-chase, *ChEC153*-expression is unknown. The

GBSS-RFP puncta themselves may be of a transient nature. Isolation of starch granules from *N. benthamiana* leaf tissue at the point of GBSS-RFP puncta observation, or expression of *GBSS-RFP* alongside *ChEC153-eGFP* may seek to clarify whether this phenomenon is host- or system-specific.

A putative mislocalisation or sequestration of the GBSS protein by ChEC153 would likely affect the structure of starch granules formed. GBSS is specifically required for amylose biosynthesis, and both its action and localisation are additionally dependent on the presence of PTST1 which appears to direct GBSS to the starch granule. Disruption of the GBSS localisation by loss of PTST1 results in the formation of amylose-free starch (Seung et al., 2015). Thus, it stands to reason that any perturbation of the localisation of GBSS by ChEC153 may impact the amylose/amylopectin ratio of starch granules being formed. Such an alteration, presumably lowering the amylose content in favour of amylopectin, might result in starch of a more accessible form for degradation. The branched nature of amylopectin corresponds to an increased number of sites from which the starch can be enzymatically degraded, relative to the linear amylose molecule. Starch degradation in leaves takes place mostly from the action of β-amylases on the non-reducing ends of chains, which are more numerous in a branched molecule. Thus, if a pathogen were relying on starch degradation as a route to increase accessible carbon, a shift toward higher amylopectin content may accelerate this. It is also possible that perturbed amylose/amylopectin ratios would alter the crystallinity of the starch granules.

Whether induced formation of small starch granules or alterations to GBSS localisation by ChEC153 occurs over timescales relevant to infection for an impact on virulence has not been evaluated here. The modification of granules through targeting of starch granule initiation or structural changes is unlikely to be immediate. Instead, a day-night cycle may be required for the effect of a disrupting protein to become apparent. For example, a hypothetical perturbation of GBSS by ChEC153 starting at the end of one day will likely have little effect on the starch granules until the following day as starch synthesis begins again after a round of degradation. However, expression of *ChEC153* during early *C. higginsianum* infection (expression seen at 22 and 40 hours post inoculation), indicated in published transcriptomic data (O'Connell et al., 2012), may leave sufficient time for starch granules to be altered before the pathogen completes its lifecycle. The time taken for *C. higginsianum* to progress to the necrotic phase of infection is highly variable, but O'Connell and colleagues estimate the necrotic phase to begin at around 60 hours post inoculation (O'Connell et al., 2012). Thus, based on these growth conditions, I can estimate around 38 hours between the initial expression of *ChEC153* and the transition to necrotrophy. This is likely gives sufficient time for starch perturbation following ChEC153 production and translocation into the host prior to a the necrotrophic stage of infection.

While the nature of the observed GBSS-RFP puncta has not been directly confirmed here, I hypothesise that they represent small starch granules. This would be consistent with their co-localisation with ChEC153 at the sites of starch granule initiation – it is not likely that a theoretical mislocalisation of GBSS would otherwise correspond so strikingly with the localisation of MRC.

3.4.4 Chapter conclusions

In conclusion, I have found that putative effector proteins ChEC153 and HaRxL94b localise to starch granule initiation sites alongside key starch granule initiation protein MRC, and that ChEC153 induces the formation of a subset of small starch granules. Interestingly, the induction of GBSS-RFP puncta was specific to the presence of ChEC153 and was not seen in the presence of HaRxL94b. This supports the concept of a specific impact of ChEC153. *H. arabidopsidis* is genetically intractable and cannot be cultured axenically. For these reasons, the remainder of the work within this project focusses on characterisation of the *C. higginsianum* putative effector ChEC153 over HaRxL94b. Whether and how this targeting of starch granule initiation by ChEC153 facilitates the colonisation of Arabidopsis by *C. higginsianum*, with ChEC153 acting as a bona fide effector protein, forms the basis of the following Chapter.

Chapter 4

Exploring the role of ChEC153 in infection

4.1 INTRODUCTION

4.1.1 Infection assays

In order to evaluate the pathogenicity of a pathogen, or the susceptibility of a host, infection assays are commonly employed. Observing the rate of infection allows for comparison of pathogen strains and host genotypes. For *C. higginsianum*, infection assays typically involve either observation of macroscopic lesions, of microscopic markers of infection progression, or quantification of fungal biomass, each of which can be used as a metric of the susceptibility of the host or virulence of the pathogen. To inform on the role of putative infection-related genes, these assays have been applied to elucidate the impact of transgenic expression of pathogen genes in the host (Ohtsu et al., 2023, Caillaud et al., 2013, Niu et al., 2016, Bos et al., 2010), and the overexpression or removal of endogenous genes from the host or pathogen (Ohtsu et al., 2023). To this end, genetic manipulation of either or both members of the pathosystem is a great advantage.

4.1.2 Genetic transformation of *C. higginsianum*

As mentioned in Chapter 1 Section 1.3.3, the genetic tractability of *C. higginsianum* serves as a boon in the study of this phytopathogen. Like other filamentous fungi (de Groot et al., 1998), *C. higginsianum* is amenable to *Agrobacterium tumefaciens*-mediated transformation. This was first demonstrated by the transformation of *C. higginsianum* (then identified as *C. destructivum*, but later determined to be a separate species (Damm et al., 2014)) with a vector for GFP expression (O'Connell et al., 2004). Transfer DNA (T-DNA) can therefore be introduced into *C. higginsianum* conidia, either to integrate randomly or, given stretches of homology, specifically into the fungal genome. Another approach for transformation of *C. higginsianum* is polyethylene glycol-mediated protoplast transformation (Yang et al., 2018). Genetic transformation of the fungus can be applied to investigate the role of endogenous proteins through generation of overexpression or knockout mutants, and strains with tagged proteins enable confocal microscopy experiments or biochemical analyses. For example, the *GFP*-expressing *C. higginsianum* strain generated by O'Connell and colleagues was recently used by Tang and colleagues to visualise the infection process to facilitate single-cell transcriptomic analyses of the host response (O'Connell et al., 2004, Tang et al., 2023). Fluorescent tagging of endogenous fungal proteins can allow subcellular localisation studies (Kleemann et al., 2012), including by complementation of knockout strains (Schmidpeter et al., 2017).

4.1.2.1 Targeted *C. higginsianum* transformation

To study the role of *C. higginsianum* genes, ATMT has been applied for the generation of knockout strains disrupting or lacking endogenous genes; this is achieved by either targeted or untargeted approaches. Targeted genetic manipulation can be employed to generate putative effector knockout strains (Korn et al., 2015, Tsushima et al., 2021, Campo et al., 2016). Further, the generation of multigene knockouts is now possible through application of *URA3*-based selection marker recycling, as demonstrated in *C. orbiculare* (Kumakura et al., 2019) and, more recently, *C. higginsianum* (Yonehara et al., 2023). These advances will aid the characterisation of effector genes despite possible functional redundancy.

The efficiency of targeted integrations can be facilitated by the use of *C. higginsianum* strains deficient in non-homologous end-joining. Non-homologous end joining (NHEJ) is a key pathway for the repair of DNA double-strand breaks in eukaryotes. In addition to NHEJ, DNA can be repaired by homologous recombination (HR). For targeted transformation of the *C. higginsianum* genome, HR can be exploited to direct the integration of a T-DNA sequence flanked by a left and right border (LB and RB) to a given genomic location. In order to increase the rates of specific integrations by directed homologous recombination, strains lacking components of the NHEJ machinery can be employed. Ku70 and Ku80 form a Ku heterodimer which is a critical component of NHEJ machinery by binding DNA double-strand breaks (Ramsden and Gellert, 1998). Loss of either component of the Ku heterodimer is sufficient to disrupt NHEJ.

Efficient targeted transformation of *C. higginsianum* using NHEJ mutant strains has been demonstrated, with Ushimaru and colleagues using a *ΔChKu70* mutant (Ushimaru et al., 2010). The authors demonstrate a correlation between the length of the regions of homology to the genome for targeting and the efficiency of integration. Around 1000 bp of homology in total (two regions of 500 bp each) or more was sufficient for a high (>96%) rate of homologous integration in the *ΔChKu70* background, while the wild-type background displayed a far lower (<1%) rate (Ushimaru et al., 2010). These stretches of homology for targeted integrations are fairly long compared to those that would be required for HR in fungi with highly efficient HR systems, such as *S. cerevisiae*. However, they are short enough to facilitate straightforward cloning from amplified or synthesised parts to vastly increase the efficiency of desired integrations and therefore remove much of the burden of transformant screening. Further, *ΔChKu80* mutants have also been used to generate effector-candidate knockouts (Korn et al., 2015). Similarly, *Ku* disruption has been demonstrated to increase targeting of transformation in other phytopathogens, including *Aspergillus* and *Magnaporthe* spp. (Tadashi et al., 2006, Villalba et al., 2008, Kershaw and Talbot, 2009).

4.1.2.2 Untargeted *C. higginsianum* transformation

Untargeted transformation of *C. higginsianum* can be employed for expression of exogenous genes (O'Connell et al., 2004), or overexpression of endogenous genes (Schmidpeter et al., 2017). The random integration of such expression cassettes presents a chance of unintentionally disrupting additional endogenous gene expression. As such, these approaches may benefit from identification and targeting of genomic safe harbour sites - regions of the genome into which to integrate expression cassettes in a targeted manner in order to reduce the likelihood of affecting genes for essential cellular processes. A number of such safe harbour sites have been identified and used for targeted integrations into other filamentous fungi such as the *URA3* and *SUCCINATE DEHYDROGENASE 1* (*SDI1*) loci (Kilaru et al., 2015). Further, the use of *SDI1* as a selection marker has been demonstrated in a number of fungi, as a single amino acid substitution in SDI1 can confer carboxin-resistance (Guo et al., 2016, Foster et al., 2018, Shima et al., 2009). Analogous approaches in *C. higginsianum* would facilitate transgenic expression or over-expression of endogenous genes.

The untargeted integration of cassettes for gene disruption is a powerful tool in forward genetics approaches to isolate mutants with interesting phenotypes and to track these integrations back to the causative gene disruption (Huser et al., 2009, Korn et al., 2015). Random insertional mutagenesis with large-scale screening has been employed to identify *C. higginsianum* genes involved in pathogenicity (Huser et al., 2009, Liu et al., 2013, Korn et al., 2015).

4.1.2.3 Selection of *C. higginsianum* transformants

Selection methods are required to obtain transformant colonies of *C. higginsianum.* As mentioned in Section 4.1.2.1, the use of *URA3* as a selection marker in *C. higginsianum* has recently been demonstrated (Yonehara et al., 2023). *URA3* encodes the enzyme OROTIDINE-5′ PHOSPHATE DECARBOXYLASE, which is required for synthesis of pyrimidine ribonucleotides, and can therefore be used as an auxotrophic marker. This strategy exploits the targeted deletion of the *URA3 C. higginsianum* homolog *CH63R_12904* (Yonehara et al., 2023) in *ura3* strains which are unable to grow unless supplied with exogenous uracil/uridine in the media. These strains can then be used as recipients for vectors containing a cassette for *URA3* expression, enabling positive selection of transformants on uracil/uridine deficient media. Loss of the *URA3* marker can be achieved using 5-fluoroorotic acid, allowing multiple rounds of transformation using *URA3* for selection. Prior to the development of *ura3 C. higginsianum* strains for *URA3* marker recycling (Yonehara et al., 2023), selection of positive *C. higginsianum* transformants has been generally reliant on initial selection using antibiotics. Common antibiotics used for fungal transformant selection include nourseothricin and hygromycin B. Hygromycin resistance is conferred by the expression of *HPH* gene (also referred to within this work as *HygR*) encoding HYGROMYCIN B PHOSPHOTRANSFERASE.

Isolation of single conidium transformants is employed to ensure a homogenous strain is acquired, and PCR-based screening methods and sequencing are also used to validate transformants. As ATMT of the fungus may result in multiple insertion events, even when targeted to specific loci, Southern blotting is also often employed in order to determine the copy number of T-DNA inserts, and complementation is applied to verify an identified gene of interest as the cause of any interesting mutant phenotypes.

4.1.3 Chapter aims

The work presented in this chapter aims to explore the role of ChEC153 in infection. I aimed to use *C. higginsianum* infection assays to evaluate the impact of overexpression of *ChEC153* in the host, as well as removal of *ChEC153* from the pathogen via targeted gene knockout. Further, I sought to identify Arabidopsis interactors of ChEC153 to inform on host proteins or processes being targeted by the putative effector during infection.

4.2 MATERIALS & METHODS

4.2.1 Culturing of *C. higginsianum*

C. higginsianum cultures were maintained on Mathur's media agar plates (2% agar, 20 mM glucose, 10 mM $MgSO_4$, 20 mM KH_2PO_4, 2.8 g/L oxoid mycological peptone), grown at room temperature in the dark and sub-cultured every two weeks by transferring an agar plug from the growing edge of a two-week-old plate to the centre of a fresh plate.

Conidia were harvested by suspension in 2.5 mL sterile water by gently scraping the plate surface with a sterile cotton bud and filtering the resulting suspension through cotton wool prior to calculation of the concentration of conidial suspensions using a haemocytometer. Glycerol stocks were made by mixing conidial suspensions with glycerol to give a final glycerol concentration of 25%, flash frozen in liquid nitrogen and stored at -70 °C. Strains were recovered from glycerol stocks by plating a small portion of the stock in the centre of a Mathur's media agar plate.

4.2.2 *C. higginsianum* infection assays

C. higginsianum conidia were isolated from Mathur's media plates as described above, and the concentration of the conidial suspension measured with a haemocytometer. This initial suspension was used to prepare a conidial suspension of a known concentration: for *ChEC153*-overexpression

Arabidopsis infection assays, ~5×10^6 conidia/mL; for all other infection assays, 2×10^6 conidia/mL. Leaves were detached from approximately five-week-old soil-grown Arabidopsis plants and their petioles inserted into 2% water agar plates. Leaves were each drop-inoculated with one 3 µL droplet of the *C. higginsianum* conidial suspension on the adaxial side, avoiding the midrib of the leaves. Plates were sealed with two layers of parafilm and incubated (25 °C; 10 h:14 h, light: dark) for up to seven days, with developing lesions imaged between three- and seven-days post inoculation (dpi). Lesions were measured from these photographs by manual selection using ImageJ (Schindelin et al., 2012). Infection assays were independently carried out a minimum of three times, and data analysed by application of an ANOVA or t-test to a linear mixed effects model with independent replicates of the experiment included as a random factor. A schematic illustrating *C. higginsianum* infection assays is presented in Figure 4.1.

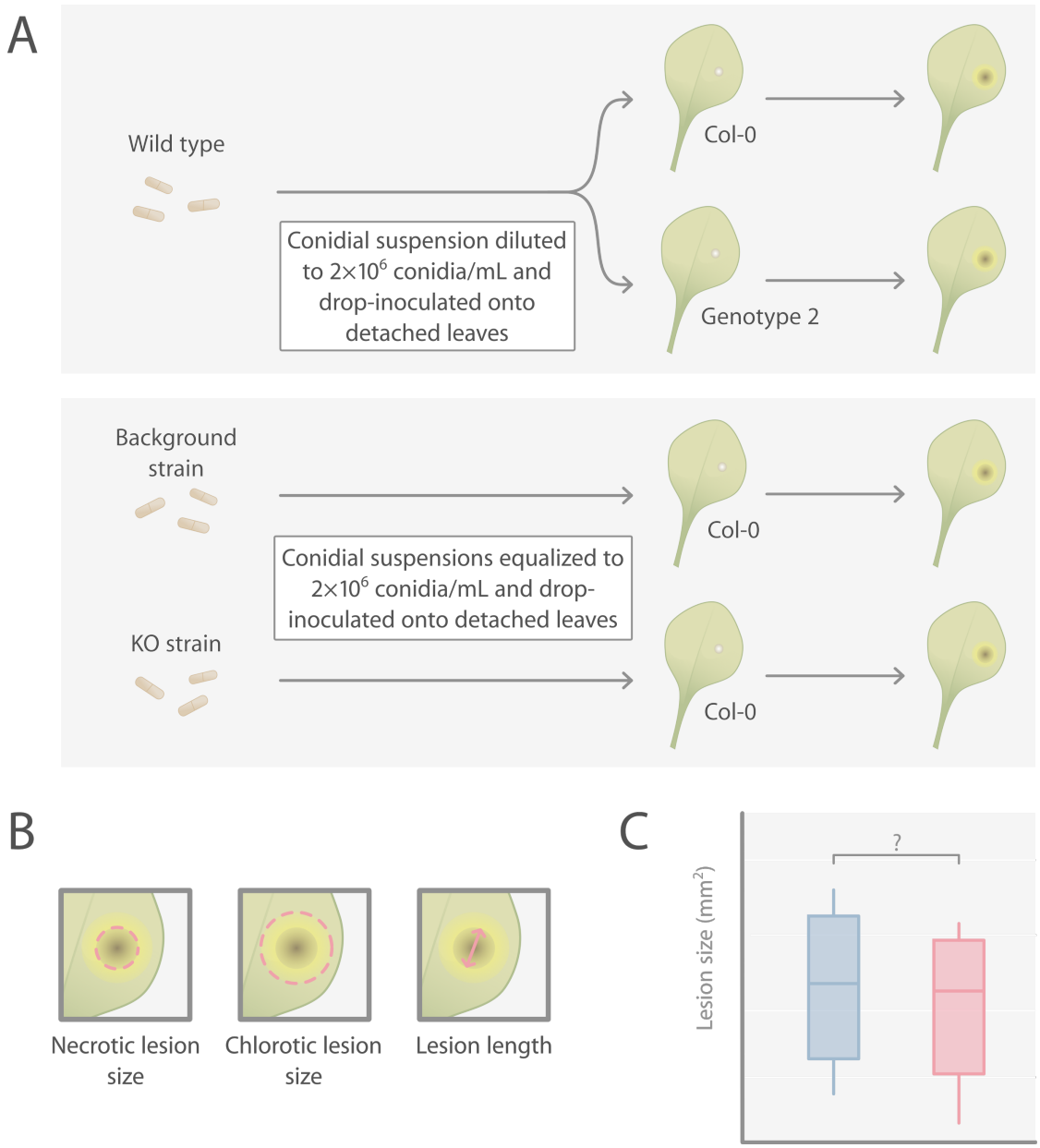


Figure 4.1: **Schematic illustrating *C. higginsianum* infection assays. (A)** Infection assay to compare: the susceptibility of an Arabidopsis line ("Genotype 2") to the Col-0 wild-type (top panel), and the virulence of a *C. higginsianum* knockout (KO) strain to the background strain from which it was generated (bottom panel). *C. higginsianum* conidia are harvested from two-week-old Mathur's medium plates by suspension in sterile water. Conidia are counted using a haemocytometer and dilutions made to give 2×10^6 conidia/mL for each strain. 3 µL droplets of each conidial suspension are placed on detached leaves (taken from four-to-five-week-old Arabidopsis plants) on 2% water agar plates. Plates are sealed with parafilm and incubated for up to seven days, with developing lesions being imaged on days four, five and six. **(B)** Lesion areas are measured in ImageJ, with necrotic and chlorotic lesion areas indicated in pink. Additionally, lesion lengths may be measured at the maximum lesion diameter visible. **(C)** Statistical comparison of quantified lesion sizes to evaluate rates of infection.

4.2.3 Generation of *C. higginsianum* gene knockout strains

4.2.3.1 Cloning of fungal gene knockout plasmids

To generate plasmids for fungal transformation (**Figure 4.4**), Golden Gate cloning methods were employed as described in the general materials and methods **Chapter 2 Section 2.5**, according to the standard system outlined in Patron et al., 2015. Details of plasmids and primers referred to are provided in **Appendix 1**. Specifically, a fungal transformation vector backbone available in the Faulkner lab (fv *HygR,* **Figure 4.4 A**) was adapted to remove the hygromycin resistance gene (*hygromycin B phosphotransferase, HPH, HygR*) to allow its use for selection in directed integration of a *HygR* expression cassette into the *C. higginsianum* genome to replace *ChEC153*. Targeted integration of the *HygR* expression cassette was achieved by flanking homologous regions to the genomic sites directly upstream and downstream of *ChEC153*. The original fv *HygR* backbone was based on a vector from the O'Connell lab for use generating a GFP-expressing *C. higginsianum* strain (pGFP-HPH, O'Connell et al., 2004).

The fv *HygR* construct was cloned by Dr Joanna Jennings from six Golden Gate compatible parts synthesised by GENEWIZ. Four of these parts were used in the construction of the new Golden Gate Level 1 fungal vector backbone fv *ΔHygR*, the synthesised parts omitted being the *HygR* CDS and promoter. Further, one of the parts (**Figure 4.2 A**, part A) was adapted to give the desired overhang (CGCT) for the Golden Gate reaction to close the backbone in the absence of the *HygR* CDS and promoter parts. To achieve this, part A was amplified with Phusion polymerase using primers 015 and 016. The resulting amplicon was purified following agarose gel electrophoresis (part A′). All fv backbone parts were digested with BsaI and purified following agarose gel electrophoresis to provide parts for Golden Gate assembly due to the synthesised parts being provided in pUC57-Kan backbones with kanamycin resistance, and the bacterial antibiotic selection for the final plasmid also being *KanR*. This eliminated the chances of undigested backbone parts being mistaken for the correctly assembled fv *ΔHygR* backbone. Thus, fv *ΔHygR* (**Figure 4.3 B**), was assembled in a BsaI-driven Golden Gate reaction and selected using kanamycin resistance and blue/white selection.

Regions of the *C. higginsianum* genome flanking the *ChEC153* CDS were amplified from wild-type *C. higginsianum* strain IMI 349061 gDNA provided by Dr Joanna Jennings using Phusion polymerase. Primers 029 and 030 were used for amplification of a 881 bp region upstream of *ChEC153* (5′ HR; chromosome 8: 4,301,441 – 4,300,561), and primers 031 and 032 for amplification of a 977 bp region downstream of *ChEC153* (3′ HR; chromosome 8: 4,298,712 – 4,297,750) according to the ASM167251v1 genome assembly of strain IMI 349063 (Zampounis et al., 2016). The PCR products were purified following agarose gel electrophoresis and cloned into Level 0

backbones pAGM1251 and pICH53399 to generate constructs L0M *ChEC153* 5′ HR and L0M *ChEC153* 3′ HR, respectively. These parts are designed to occupy positions 1 and 5 of the cloning system illustrated in Figure 4.2 B and C.

To drive the expression of *HygR* gene in the final KO constructs, two promoters were used: a cos1gpdA promoter which had previously been used in the Faulkner lab to drive expression of *GFP* and *mCherry* in transgenic *C. higginsianum* strains, and the gpdA promoter of the original fv *HygR* backbone. The cos1gpdA promoter was amplified using primers 023 and 024, and the gpdA promoter using primers 017 and 018. Both PCR products were cloned into the Level 0 receiver vector pAGM1276 (to occupy position 2 of the cloning scheme represented in Figure 4.2 B and C), to produce constructs L0M cos1gpdA P and L0M gpdA P respectively. The *HygR* CDS was amplified from the *HygR* synthesised part using primers to clone it to a Golden Gate Level 0 vector for the CDS position, position 3 (primers 019 and 025, cloning the PCR product into Level 0 backbone pICH41308 to generate L0M *HygR* CDS). A Level 0 construct for the trpC terminator with the required Golden Gate overhangs was generated from the PCR product of primers 021 and 022 amplifying from a pre-existing trpC terminator part construct (pICSL60013, TSL SynBio) cloned into the receiver vector pICH53388 to occupy position 4 in the cloning scheme represented in Figure 4.2 B and C (L0M trpC T).

Due to the need to flank the expression cassette with parts (homologous flanking regions) in the Golden Gate Level 1 format, the usual cloning position syntax was not applicable, but standard cloning positions with the usual defined overhangs were repurposed for these parts. As labelled in Figure 4.2 B and C: Position 0 = backbone, position 1 = P5Uf, position 2 = NTAG, position 3 = CDS, position 4 = 3UTR, position 5 = TERM (see Appendix 1 Tab. B).

The Level 0 parts detailed above were then assembled into the fv *ΔHygR* backbone via Golden Gate cloning with BsaI to produce the Level 1 fungal vectors fv *ChEC153* KO 1 and fv *ChEC153* KO 2, with the latter illustrated in Figure 4.2 B and C.

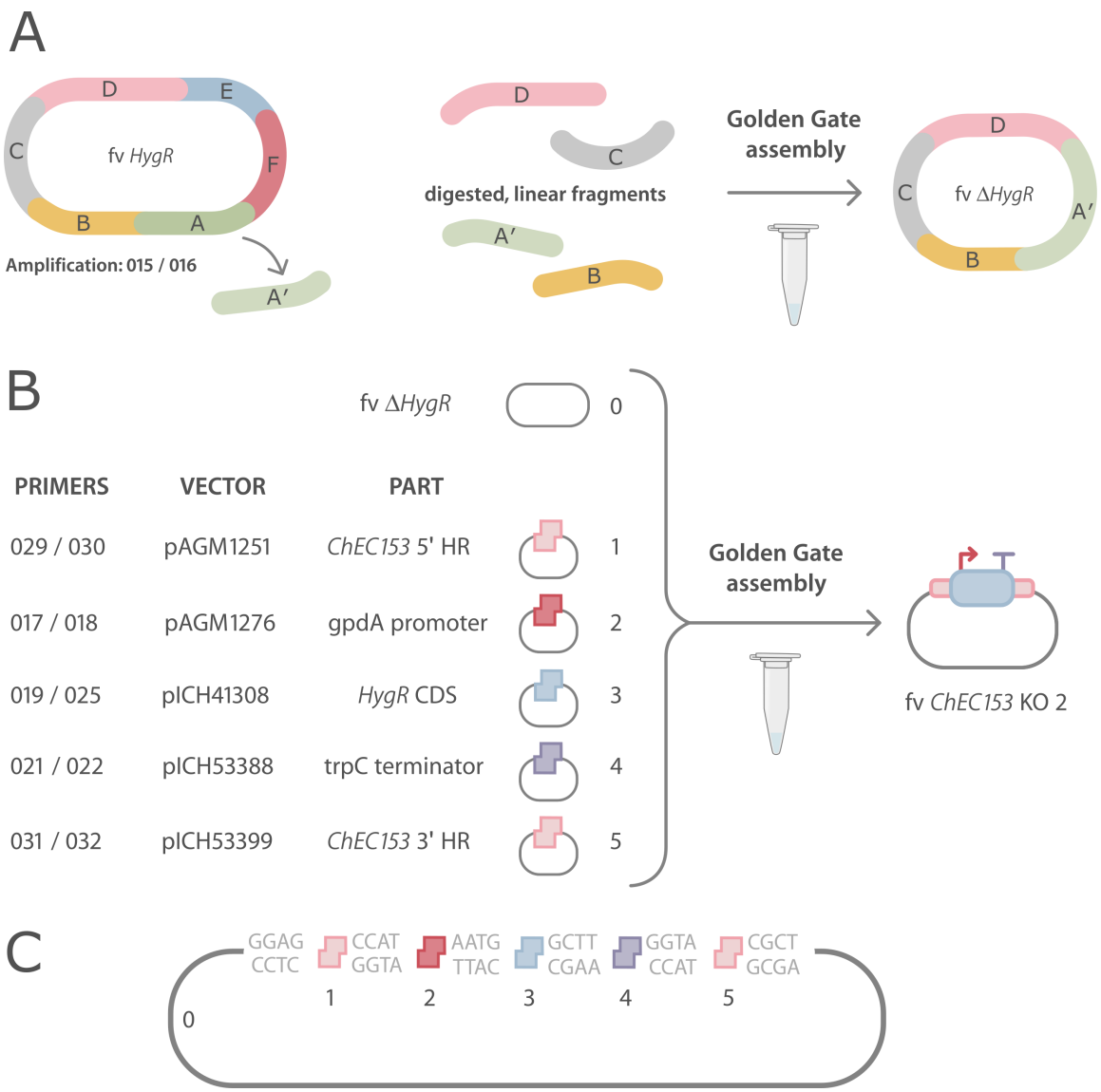


Figure 4.2: **Schematic for Golden Gate cloning of fungal gene-knockout constructs. (A)** Cloning of fv *ΔHygR* backbone for KO construct generation. Two parts (E and F) from fv *HygR* were omitted, and part A was amplified with primers 015 and 016 to introduce a new overhang (A′) to allow fusion with part D in the assembly reaction. **(B)** Assembly of fv *ChEC153* KO 2 from five Level 0 parts and the fv *ΔHygR* backbone (Fv part position 0) via a BsaI-driven Golden Gate reaction. A cassette for *HygR* expression (positions 2/3/4 containing gpdA promoter, CDS, and trpC terminator) flanked by regions of homology to the genomic sequence upstream (5′ HR) and downstream (3′ HR) of *ChEC153* in the *C. higginsianum* genome. Fusion sites between parts are indicated in **(C)**.

4.2.3.2 *C. higginsianum* transformation

Agrobacterium strain AGL1 carrying a fungal knockout (KO) plasmid (fv *ChEC153* KO 1 or fv *ChEC153* KO 2) was grown in 10 mL LB + rifampicin (50 µg/mL), carbenicillin (100 µg/mL) and kanamycin (50 µg/mL) for ~16 hours at 28 °C with shaking to obtain a dense culture. *Agrobacteria* were pelleted by centrifugation and resuspended in glycerol induction (GI) broth (0.1 M NH_4Cl, 5

mM $MgSO_4 \cdot 7H_2O$, 1 mM KCl, 70 µM $CaCl_2 \cdot H_2O$, 9 µM $FeSO_4 \cdot 7H_2O$, 10 mM glucose, 50 mM glycerol, 50 mM MES/NaOH pH 5.2) + 200 µM acetosyringone. The OD_{600nm} of this suspension was measured and a dilution made with GI broth + 200 µM acetosyringone to give a final OD_{600nm} = 0.4. *ΔChKu80 C. higginsianum* conidia (Strain CY6021 derived from MAFF 305635, Korn et al., 2015) were harvested from two-week-old Mathur's media plates by suspension in 2 mL GI broth + 200 µM acetosyringone. The concentration of this suspension was calculated using a haemocytometer, and a dilution made in GI broth + 200 µM acetosyringone to give 1×10^7 conidia/mL.

Sterile mixed cellulose ester membrane filter discs (GE Whatman, 7141-114) were placed onto GI agar plates (GI broth with 1.5% agar, 5 mM $NaH_2PO_4 \cdot H_2O$ and 200 µM acetosyringone) and any air bubbles removed with sterile spreaders. *Agrobacteria* and conidial suspensions were then mixed at a 1:1 ratio and 200 µL aliquots of this suspension were placed onto the filter discs and spread across the surface using a sterile spreader. GI plates were incubated at room temperature for two days of co-cultivation before the filter discs were inverted and transferred to potato dextrose agar (PDA; Formedium, PDA0102; 41 g/L) + 100 µg/mL hygromycin, 50 µg/mL cefotaxime and 50 µg/mL spectinomycin plates. PDA plates were then incubated for a further three days at room temperature before filter discs were removed. Plates were re-sealed and incubated until the appearance of fungal colonies, which were then sub-cultured on fresh PDA + hygromycin plates for two rounds of growth to confirm hygromycin resistance. Resistant strains were diluted to single conidia in sterile dH_2O by counting conidia with a haemocytometer, and re-plated on PDA + hygromycin. Following selection, strains were maintained on Mathur's media plates and stored as glycerol stocks (see Section 4.2.1).

4.2.3.3 Extraction of *C. higginsianum* genomic DNA

To generate fungal material for genomic DNA isolation, two approximately 1 cm ø discs were cut from the edge of *C. higginsianum* plates and transferred to 100 mL potato dextrose broth (PDB; Formedium, PDB0102; 12 g/L) which was incubated at 27 °C with shaking at 100 rpm for seven days. Mycelium was harvested by filtering cultures through two layers of Miracloth (Millipore, 475855), and rinsing with 100 mL sterile dH_2O. Fungal material was briefly dried and flattened in sterile filter papers before grinding in liquid nitrogen with a pestle and mortar. Ground material was collected in 2 mL microfuge tubes in approximately 100 mg aliquots and stored at -70 °C prior to genomic DNA extraction. For genomic DNA extraction, a DNeasy Plant Mini Kit (QIAGEN, 69104) was used according to the manufacturer's instructions, with elution in 50 µL pre-warmed Buffer AE.

4.2.3.4 Screening of *C. higginsianum* transformant colonies

Potential transformant colonies were screened for the presence/absence of *ChEC153* by PCR using fungal material scraped from the surface of the plate with a sterile pipette tip as template with primers 035/036 or 037/038 using Phusion polymerase as described in Section 2.5.1. Promising *C. higginsianum* strains identified were diluted to single conidia, selected on PDA + hygromycin plates and then maintained on Mathur's media plates for further screening, including isolation of genomic DNA for PCR analysis. PCR analysis from genomic DNA was carried out as described for amplification from fungal colony material, using 0.5 µL of extracted gDNA (~100 ng/µL) as the template.

4.2.4 Lugol staining of infected Arabidopsis leaves

Lugol solution (iodine/potassium iodide solution) was used to stain starch in infected leaves. Infected Arabidopsis leaves were first cleared by washing and incubating several times in 80% ethanol. Leaves were subsequently washed twice in sterile dH_2O to remove ethanol, before being incubated submerged in 100% Lugol solution (Sigma, L6146) for >10 hours. Following staining, samples were rinsed twice in sterile dH_2O to remove excess Lugol, and photographs taken.

4.2.5 Immunoprecipitation-mass spectrometry (IP-MS)

For mass spectrometry, proteins were extracted as described in Section 2.8 for three Col-0 and three *ChEC153-eGFP*-expressing Arabidopsis samples, and IP was carried out with 50 µL GFP-Trap® Magnetic Agarose beads (ChromoTek, gtma). Following IP, proteins were recovered from beads by boiling with 30 µL of extraction buffer and 10 µL of 5× loading buffer. 1 µL of each sample was subject to SDS-PAGE and Western blotting to check for the presence of the ChEC153-eGFP before proceeding with mass spectrometry sample preparation.

The remaining volume of immunoprecipitated samples were loaded on 10% SDS polyacrylamide gels (1.5 mm thickness) and run approximately 6 mm into the gel. Gel slices containing protein were excised and processed for trypsin digest using a method adapted from Shevchenko *et al*. (Shevchenko et al., 2006). Gel pieces were de-stained in 30% EtOH for 30 minutes and washed with 50 mM triethylammonium bicarbonate buffer (TEAB)/50% acetonitrile (ACN) before being incubated at 55 °C with 10 mM DTT in 50 mM TEAB for 30 minutes. DTT was removed and replaced with IAA solution (30 mM iodoacetamide in 50 mM TEAB), then incubated for 30 minutes at room temperature in the dark. IAA was removed and samples washed with 50 mM TEAB/50% ACN, then with 50 mM TEAB. Buffer was removed from gel slices, which were then cut into approximately 1 mm × 1 mm pieces and transferred to protein LoBind tubes. Samples were washed with 50 mM

TEAB/50% ACN then 100% ACN before being submitted to the JIC proteomics department for trypsin digestion and mass spectrometry.

The JIC Proteomics Platform carried out trypsin digestion and mass spectrometry of samples using the following protocols, written/provided by Dr Carlo de Oliveira Martins (JIC Proteomics Platform). After washing and dehydration with acetonitrile, the gels were soaked with 50 mM TEAB containing 10 ng/µL Sequencing Grade Trypsin (Promega) and incubated at 40 °C for 8 hours. The extracted peptide solution was dried down, and the peptides dissolved in 3% acetonitrile and 0.1% TFA (trifluoroacetic acid). Aliquots were analysed by nanoLC-MS/MS on an Orbitrap Eclipse™ Tribrid™ mass spectrometer coupled to an UltiMate® 3000 RSLCnano LC system (Thermo Fisher Scientific, Hemel Hempstead, UK). The samples were loaded and trapped using a pre-column with 0.1% TFA at 15 µL per minute for 4 minutes. The trap column was then switched in-line with the analytical column (nanoEase M/Z column, HSS C18 T3, 100 Å, 1.8 µm; Waters, Wilmslow, UK) for separation using the following gradient of solvents A (water, 0.1% formic acid) and B (80% acetonitrile, 0.1% formic acid) at a flow rate of 0.2 µL per minute : 0-3 min 3% B (parallel to trapping); 4-10 min linear increase B to 9%; 10-70 min increase B to 40%; 70-90 min increase B to 60%; followed by a ramp to 99% B and re-equilibration to 3% B. Data were acquired with the following mass spectrometer settings in positive ion mode: MS1/Orbitrap (OT): resolution 120K, profile mode, mass range m/z 300-1800, normalised Automated Gain Control (AGC) target 100%, maximum fill time 50 ms; MS2/Ion Trap (IT): data dependent analysis was performed using parallel Collision-Induced Dissociation (CID) and Higher-energy C-trap Dissociation (HCD) fragmentation with the following parameters: top20 in IT turbo mode, centroid mode, isolation window 1.0 Da, charge states 2-5, threshold 1.0e4, CE = 33, AGC target 1e4, max. inject time 35 ms, dynamic exclusion 1 count, 15 s exclusion, exclusion mass window ±10 ppm.

Raw mass spectrometry files were used to generate peaklists using MSConvert v2.0 from Proteowizard (Chambers et al., 2012). The final search was performed using the in-house Mascot Server 2.7.0.1 (Matrixscience, London, UK) on the *Arabidopsis thaliana* protein database (arabidopsis.org, 35,386 entries) together with a custom database containing the target sequence of interest and the MaxQuant contaminants database (250 entries). A precursor tolerance of 6 ppm and a fragment tolerance of 0.6 Da was used for this search, with the enzyme set to trypsin with a maximum of two allowed missed cleavages. Oxidation (M), deamidation (N/Q) and acetylation (protein N-terminus) were set as variable modifications and carbamidomethylation (CAM) of cysteine as fixed modification. The Mascot search results were imported into Scaffold 4.11.0 (www.proteomesoftware.com) for visualisation. SAINTexpress (Teo et al., 2014) was used to examine the exclusive unique peptide counts from Scaffold. Enriched prey proteins were defined as those with a Bayesian false discovery rate (BFDR) in SAINTexpress of < 0.25 (25%).

4.2.6 Arabidopsis genotyping: genomic DNA extraction and genotyping PCRs

Genotyping of T-DNA insertion lines was carried out by PCR analysis using genomic DNA as a template. Genomic DNA was extracted from flash-frozen leaves homogenised to a fine powder using a Geno/Grinder® (SPEX Sample Prep). Samples were kept on ice while 200 µL of DNA extraction buffer (0.2 M Tris-HCl pH 8, 0.4 M LiCl, 25 mM EDTA, 1% SDS) was added, and samples vortexed thoroughly before being centrifuged at 4 °C for 10 minutes (15,800 ×g). The supernatant from each sample was added to 120 µL of ice-cold 100% isopropanol and mixed gently by pipetting. Samples were centrifuged at 4 °C for 15 minutes (15,800 ×g) to pellet the DNA, and the supernatant from each sample was discarded. Pellets of DNA were washed briefly in 200 µL of 70% EtOH before being dried at 37 °C for 15 minutes. DNA was then resuspended in 50 µL dH_2O. Genotyping PCRs were carried out using 1 µL of genomic DNA as the template in reactions using 0.05 µL GoTaq® G2 Flexi DNA Polymerase (Promega, M780B) and 2 µL 5× Green GoTaq® Flexi Buffer (Promega, M891A). Reactions included 0.5 µL of either primer (10 µM), 0.2 µL 10 mM dNTPs, and 0.8 µL of 25 mM $MgCl_2$, and were made up to a final reaction volume of 10 µL using sterile dH_2O. Primers used to test for the presence or absence of the T-DNA insert are listed in Appendix 1 Tab. E under the heading “genotyping primers”. PCR products were separated by agarose gel electrophoresis on a 1% agarose gel stained with EtBr.

4.3 RESULTS

4.3.1 Host expression of *ChEC153* increases susceptibility to *C. higginsianum*

Bona fide effector proteins are expected to contribute to the virulence of the pathogen that produces them. Should ChEC153 have a clear role in infection, overexpression of the effector in the host may induce an increased susceptibility to infection. In order to investigate whether ChEC153 has a role in infection, I first carried out *C. higginsianum* infection assays on Arabidopsis lines expressing *ChEC153-eGFP* alongside Col-0 controls. These transgenic lines have no visible morphological differences compared to the wild-type when grown under long- or short-day conditions.

I measured resulting necrotic lesion areas at four and five dpi, with data presented in **Figure 4.3**. Infection assay data were analysed using an ANOVA on a linear mixed effects model, with the individual replicate of the experiment included as a random factor (plotted as differently coloured points). Arabidopsis lines expressing *35S::ChEC153-eGFP* appear slightly but significantly more susceptible to *C. higginsianum* infection than the wild-type in terms of the size of necrotic lesions at four days post inoculation (ANOVA: $F = 7.9401$, $df = 2$, $p < 0.001$; Tukey HSD ChEC153 2-5 vs Col-0: $p = 0.002$; Tukey HSD ChEC153 2-9 vs Col-0: $p = 0.008$). This increase in susceptibility is only significant for one *ChEC153-eGFP* line (2-9) at five dpi (ANOVA: $F = 5.3696$, $df = 2$, $p = 0.005$; Tukey HSD ChEC153 2-5 vs Col-0: $p = 0.155$; Tukey HSD ChEC153 2-9 vs Col-0: $p = 0.004$), and by six dpi neither line shows significantly larger lesion sizes than the Col-0 control (**Appendix 2 Fig. A** and **Appendix 2 Tab. A**). Similar trends were seen for chlorotic lesion sizes, with a significant increase in lesion size relative to Col-0 seen for both *ChEC153-eGFP* lines only at four dpi (ANOVA: $F = 6.4088$, $df = 2$, $p = 0.002$; Tukey HSD ChEC153 2-5 vs Col-0: $p = 0.007$; Tukey HSD ChEC153 2-9 vs Col-0: $p = 0.011$; **Appendix 2 Tab. A**). These results support the notion that ChEC153 contributes to the virulence of the fungus or the susceptibility of the host early during the infection process.

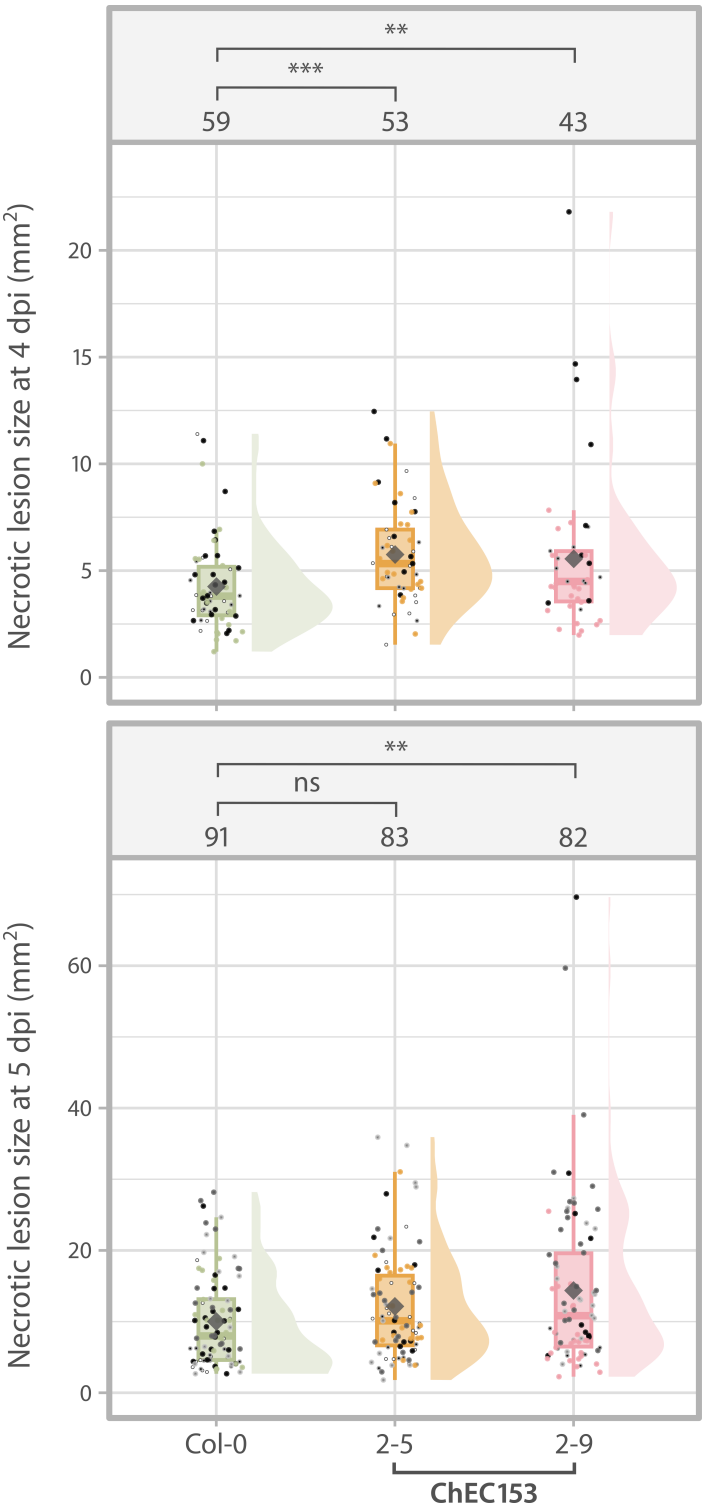


Figure 4.3: *35S::ChEC153-eGFP* Arabidopsis lines show increased susceptibility to *C. higginsianum*. Necrotic lesion areas measured (mm^2) at four and five dpi for two independent *35S::ChEC153-eGFP* lines: 2-5 and 2-9. Infection assay data for each individual replicate of the experiment are plotted as differently coloured points. Data were analysed using an ANOVA on a linear mixed effects model, with the individual replicate of the experiment included as a random factor, and Tukey HSD post hoc tests. Asterisks denote significant differences between Col-0 and *ChEC153* lines: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$ (not significant). Values of n are shown above each plot, and mean lesion sizes for each are represented by grey diamonds. Full statistical test results shown in Appendix 2 Tab. A.

4.3.2 *ΔChEC153 C. higginsianum* strains do not have altered virulence

4.3.2.1 Generation of *ChEC153*-knockout *C. higginsianum* strains

To gain more insight into the role of ChEC153, and to complement infection assays in which *ChEC153* is overexpressed in the host, I sought to generate *C. higginsianum* strains in which *ChEC153* was absent. In order to do this, I generated a fungal transformation vector based on the backbone already available in the Faulkner lab (fv *HygR*, **Figure 4.4 A**; itself based on a transformation vector from the lab of Prof. Richard O'Connell (O'Connell et al., 2004) and made Golden Gate-compatible by Dr Joanna Jennings in the Faulkner lab), but in which the hygromycin resistance cassette was removed (new backbone without hygromycin resistance cassette: fv *ΔHygR*, **Figure 4.4 B**). This allowed me to assemble vectors containing a hygromycin resistance cassette flanked by regions of approximately 900 bp homologous to the sequences directly upstream and downstream of *ChEC153* within the *C. higginsianum* genome. I generated two such vectors, differing only by the promoter used to drive the expression of the hygromycin resistance gene: fv *ChEC153* KO 1 with a cos1gpdA promoter, and fv *ChEC153* KO 2 with a gpdA promoter (**Figure 4.4 C** and **D**). The gpdA promoter is the promoter from the *Aspergillus nidulans GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE* gene (Punt et al., 1990, Punt et al., 1991), with the full gpdA promoter being used in fv *ChEC153* KO 2. In contrast, the fv *ChEC153* KO 1 cos1gpdA promoter, which was derived from the pGFP-HPH vector from O'Connell and colleagues, contains a cos1 sequence and ~400 bp fragment of the gpdA promoter (O'Connell et al., 2004, Jennings, 2021). Both constructs were used here in case only one resulted in the production of transformant colonies due to potential *HygR* expression level differences. Further details of the cloning of these constructs are provided in **Section 4.2.3.1**.

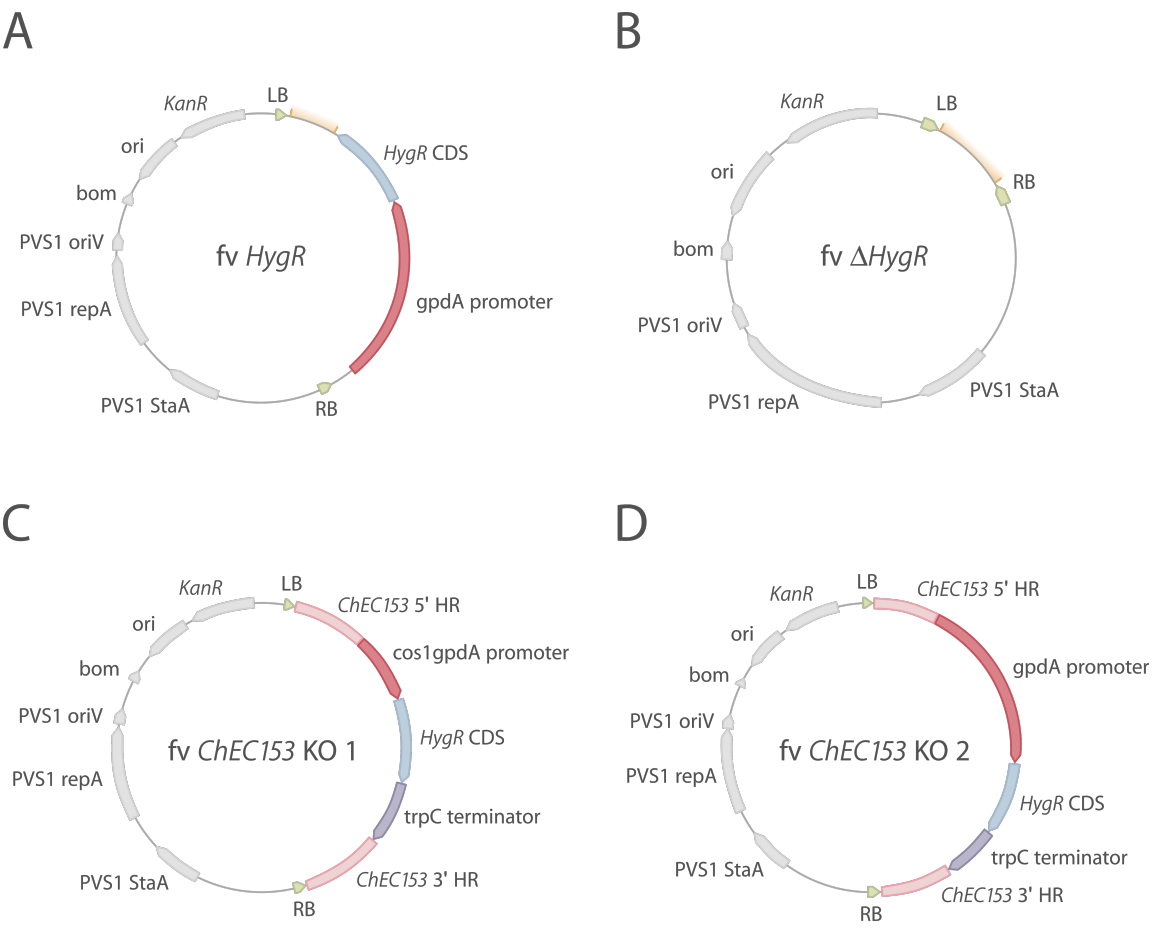


Figure 4.4: Vectors used for generation of fungal putative effector-knockout strains. **(A)** Original fungal vector (fv) for fungal transformation (fv *HygR*), assembled by Joanna Jennings from synthesised parts based on a vector from the O'Connell lab. Parts were made Golden Gate-compatible such that the final backbone contains convergent BsaI sites (orange lines) for Level 1 vector assembly to replace a *lacZ* cassette for blue-white selection (sequence marked in orange). The vector contains a kanamycin resistance gene for bacterial selection, and hygromycin resistance gene within the left and right borders (LB and RB) for fungal selection. **(B)** Altered fungal transformation backbone (fv *ΔHygR*) generated in this work to remove the hygromycin resistance cassette from within the left and right borders. The sequence is otherwise unchanged relative to fv *HygR*. **(C)** and **(D)** Level 1 vectors for fungal effector knockout containing homologous regions (HRs) to the sequences directly 5' and 3' of *ChEC153* in the *C. higginsianum* genome, flanking a cassette for hygromycin resistance driven either by a cos1gpdA promoter (fv *ChEC153* KO 1) or gpdA promoter (fv *ChEC153* KO 2).

I used ATMT to transform *C. higginsianum* with the effector-knockout vectors (fv *ChEC153* KO 1 and fv *ChEC153* KO 2), allowing for directed homologous recombination to replace the effector gene with the cassette for hygromycin resistance. In order to increase the likelihood of efficient targeting and reduce the colony screening requirements, I used a non-homologous end-joining deficient strain *ΔChKu80* generated from MAFF 305635 and kindly provided by Prof. Christian Koch (Korn et al., 2015). Conidial suspensions of putative transformants which grew on media containing 100 µg/mL hygromycin were diluted and re-plated to obtain single conidial isolates, which were

then sub-cultured. Candidate effector-knockout strains were grown in liquid culture and genomic DNA extracted to allow further validation by PCR. A schematic for the generation of *ChEC153*-knockout strains is depicted in **Figure 4.5**. Using this methodology, two independent knockout strains were isolated, hereafter termed *ΔChEC153-1* and *ΔChEC153-2*, generated from transformation of *ΔChKu80* with fv *ChEC153* KO 1 and fv *ChEC153* KO 2, respectively. The PCR-based confirmation of the loss of *ChEC153* from these strains, and the integration of the hygromycin resistance cassette, are presented in **Figure 4.6**.

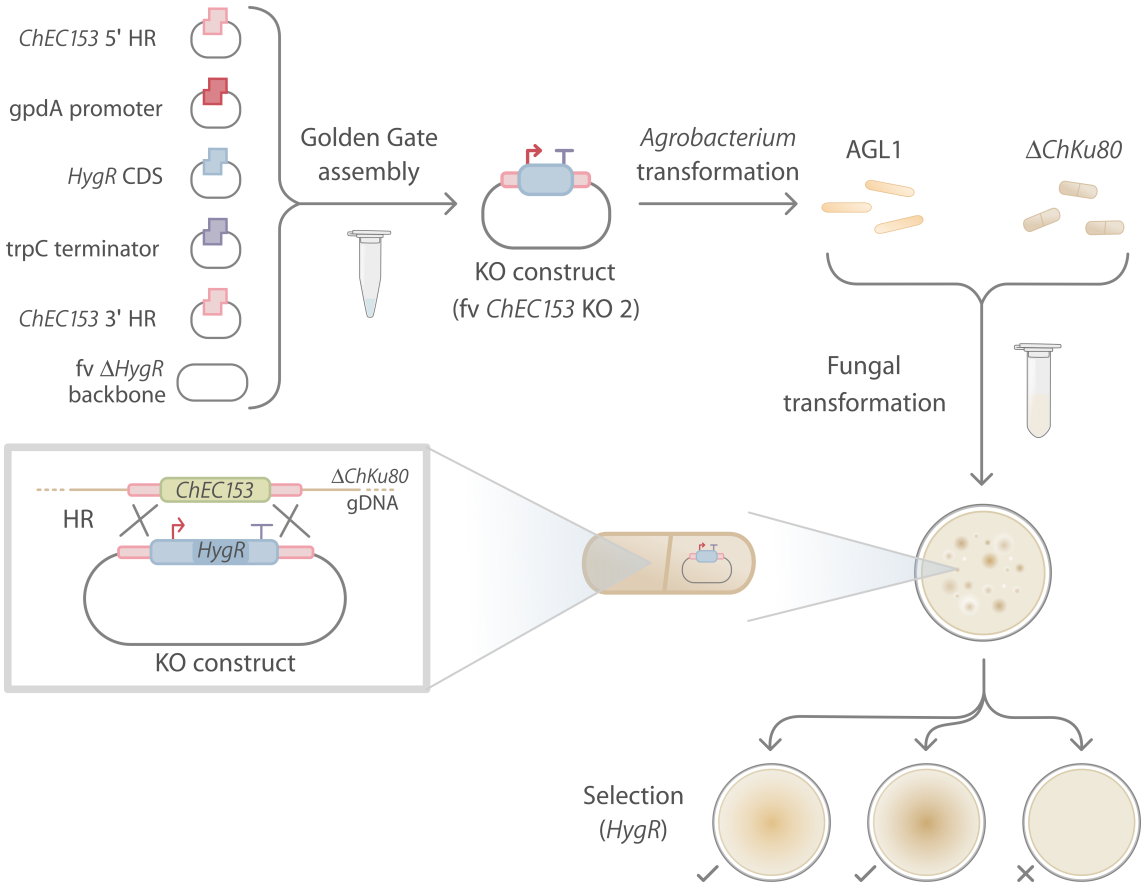


Figure 4.5: **Schematic depicting the generation of *C. higginsianum ChEC153*-knockout strains.** *ChEC153*-knockout via *Agrobacterium tumefaciens*-mediated transformation. A knockout (KO) construct was generated via Golden Gate cloning as previously described. This construct was transformed into the *Agrobacterium* strain AGL1, which was used to transform conidia of the *ΔChKu80* strain via T-DNA transfer. Co-cultivation was carried out for three days before selection on PDA + hygromycin plates and potential transformant strains isolated for further validation. Grey box: homologous recombination (HR) between the KO construct and *ΔChKu80* genomic DNA (gDNA) should result in replacement of *ChEC153* with the hygromycin resistance (*HygR*) cassette.

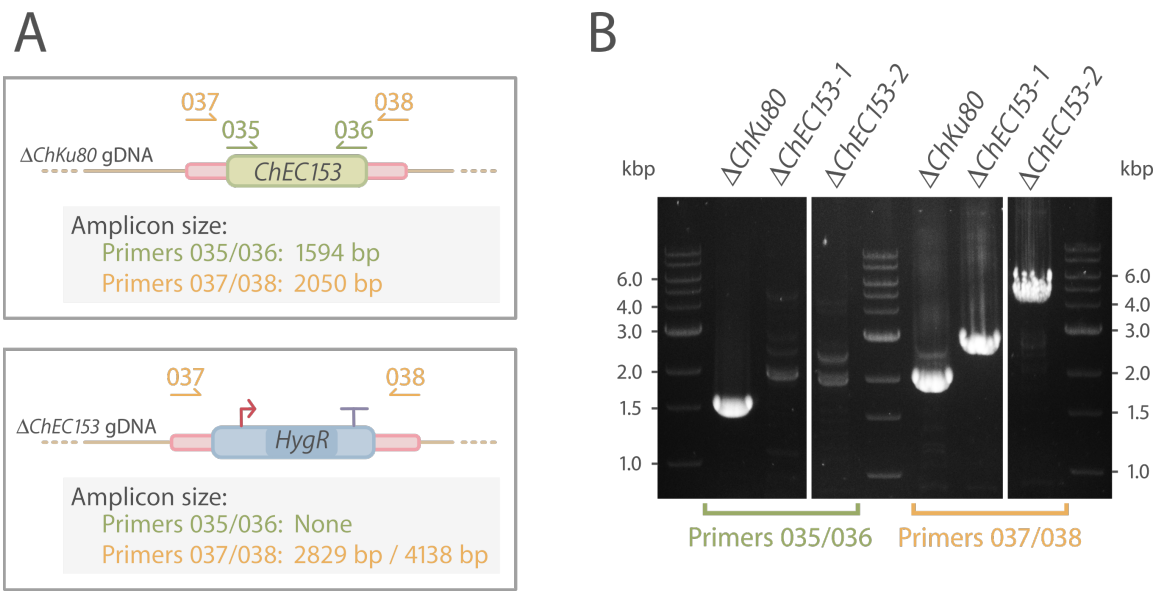


Figure 4.6: PCR-based verification of *ChEC153*-knockout strains from extracted genomic DNA. (A) Schematic depicting the binding sites of primers used within the background strain genome (top) or the positive transformant genome (bottom). Expected PCR amplicon sizes are stated for each reaction for which amplicons are anticipated. (B) Results of PCR genotyping reactions using *ΔChKu80* background strain, *ΔChEC153-1* knockout strain, or *ΔChEC153-2* knockout strain gDNA as the PCR template. Amplification of the effector sequence is seen with the background strain, but not the knockout strains, as the template using primers 035 and 036 (left panel). Relative to the background strain, the knockout strains show a shift in product size amplified with primers 037 and 038 which bind the homologous flanking regions of *ChEC153*, indicative of the integration of the hygromycin resistance cassette (right panel). PCR products were separated on a 1% agarose gel stained with EtBr alongside a 1 kb plus ladder (NEB).

4.3.2.2 *In vitro* phenotypes of Δ*ChEC153* strains

To check for defects in vegetative growth of the C. *higginsianum ChEC153*-knockout strains generated, I plated 20 µL of 2×10^6 conidia/mL suspensions of each *ΔChEC153-1, ΔChEC153-2,* and the background strain *ΔChKu80* on non-selective Mathur's media plates and photographed the plates over time. While *ΔChEC153-2* appeared to grow at a similar rate to the *ΔChKu80* background strain, *ΔChEC153-1* showed a clear growth defect *in vitro* (**Figure 4.7 A**). Further, while *ΔChEC153-2* colonies appeared morphologically similar to *ΔChKu80, ΔChEC153-1* colony morphology appeared altered. The *ΔChEC153-1* colonies had a fluffier textural appearance possibly due to greater formation of aerial hyphae, and a less circular, more irregular overall shape than the background strain. Quantification of the colony size and statistical analysis of the growth rates (using an ANOVA on a linear model followed by a Tukey HSD post hoc test at each time-point) confirms this observation (**Figure 4.7 B**; full statistical results presented in **Appendix 2 Tab. B**), with *ΔChEC153-1* colonies being significantly smaller than background strain colonies at all time-points

recorded. *ΔChEC153-2* colonies were slightly larger than background colonies at four days after sub-culturing, but showed no significant differences in size at any of the other time-points, and I therefore conclude this strain to have a comparable growth rate to the background *ΔChKu80* strain.

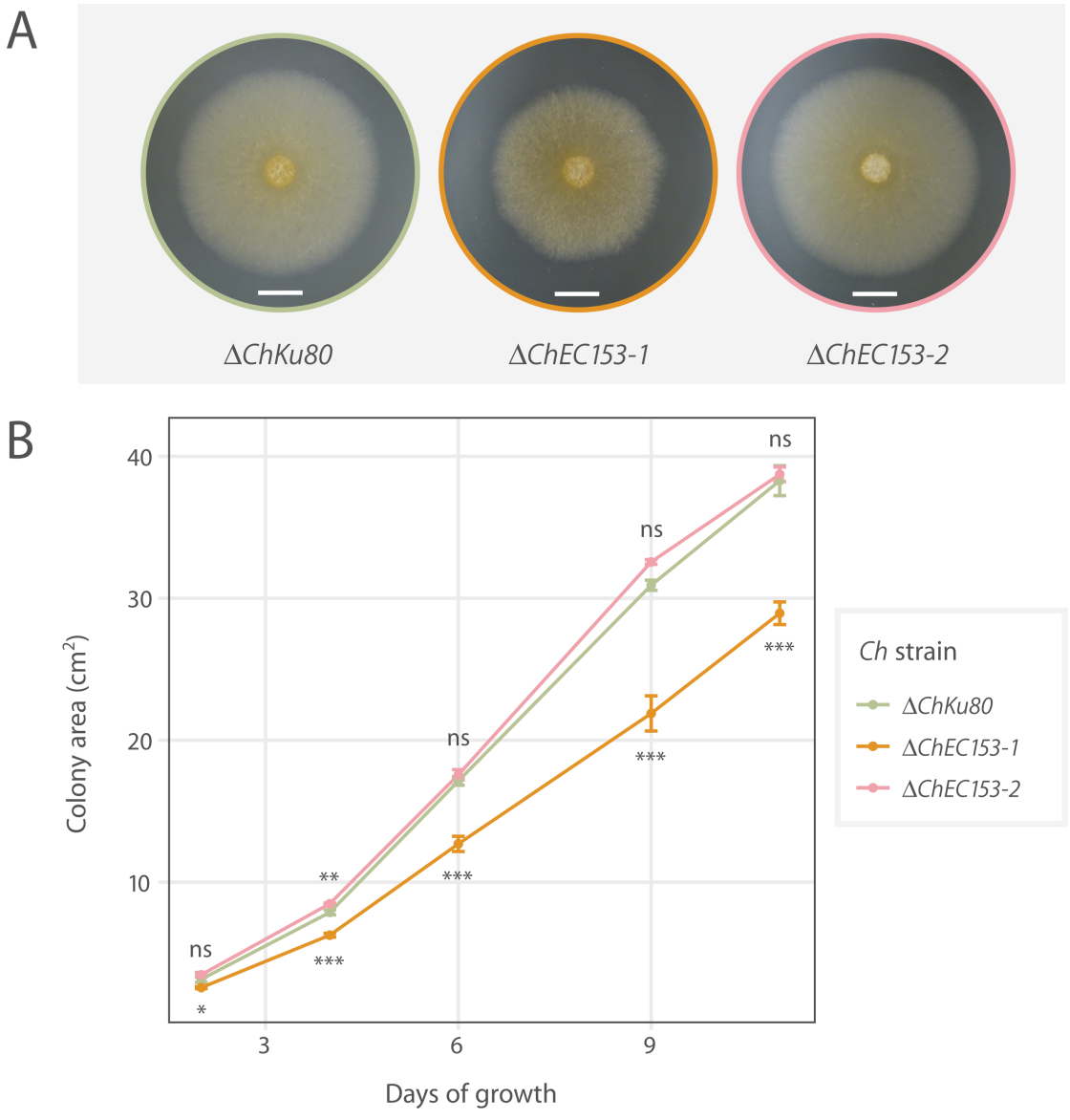


Figure 4.7: **The *in vitro* growth rate of strain *ΔChEC153-2* is no different to that of the background strain *ΔChKu80,* but *ΔChEC153-1* shows a slower rate of growth. (A)** Representative photographs showing one colony for each *C. higginsianum* strain after six days of growth following sub-culturing. Scale bars, 1 cm. **(B)** Colony area for each strain after two to eleven days of growth following sub-culturing. Data plotted are mean colony areas (cm^2) quantified from three plates, with error bars showing standard deviation. Asterisks denote significance of differences in colony size relative to *ΔChKu80:* *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = not significant, as determined by analysis of colony size data using an ANOVA on a linear mixed effects model, followed by a Tukey HSD post hoc test for each time-point, with full statistical results presented in **Appendix 2 Tab. B**.

To further verify the *in vitro* phenotype, I evaluated the sporulation rate of *ΔChEC153-2* with respect to *ΔChKu80*. I plated equal volumes of normalised conidial suspensions in the centre of Mathur's media plates. After ten days of growth, I harvested conidia from these plates in equal volumes of water and calculated the conidial concentrations of these suspensions using a haemocytometer. Data showing the conidial concentration attained from three plates for each strain are presented in **Figure 4.8**. Conidial concentration data were analysed using a t-test on a linear model ($p = 0.972$). I observed no difference in the concentration of conidial suspensions obtained between strains, indicating that there is no defect in conidiation in strain *ΔChEC153-2*. Thus, *ΔChEC153-2* represents a promising knockout strain with which to evaluate the importance of *ChEC153* on the virulence of *C. higginsianum* using Arabidopsis infection assays.

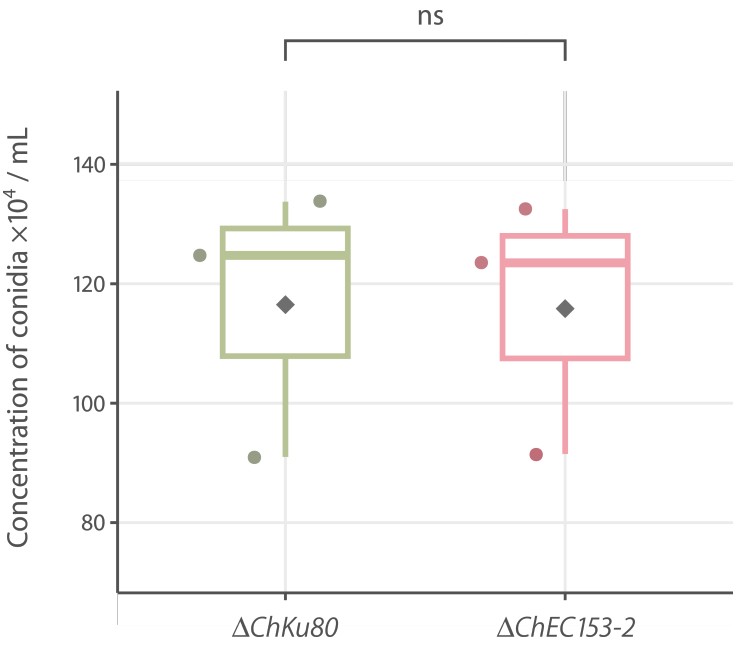


Figure 4.8: **The sporulation rate of *ΔChEC153-2* is no different to that of the background strain *ΔChKu80*.** The concentration of conidia (number of conidia $\times 10^4$/mL) obtained from Mathur's media agar plates of *C. higginsianum* strains *ΔChEC153-2* and *ΔChKu80* was calculated using a haemocytometer, with three plates being measured for each strain. Grey diamonds represent the mean concentrations of conidia. No significant difference (ns) is seen between conidial concentrations of the two strains, as determined by analysing the data using a t-test on a linear model.

4.3.2.3 *ΔChEC153* strains have no clear virulence phenotype on Arabidopsis, nor confer a macroscopic starch phenotype during infection

To evaluate whether the absence of the putative effector is sufficient to alter the virulence of *C. higginsianum* on Arabidopsis, I conducted infection assays using these *ChEC153*-knockout strains alongside the background strain *ΔChKu80*. I normalised conidial suspensions used for inoculations to the same concentration (2×10^6 conidia/mL), and conducted the experiment three independent times, with results for necrotic lesion quantification presented in **Figure 4.9**. Necrotic lesion size data at four and five dpi were analysed using an ANOVA on a linear mixed effects model, with the individual replicate of the experiment included as a random factor (plotted as differently coloured points). Significant differences between genotypes were observed at both time-points (4 dpi ANOVA: $F = 13.524$, $df = 2$, $p < 0.001$; 5 dpi ANOVA: $F = 22.642$, $df = 2$, $p < 0.001$). Analysis with Tukey HSD shows a significant decrease in the area of necrotic lesions produced by *ΔChEC153-1* (4 dpi: $p < 0.001$; 5 dpi: $p < 0.001$) but not *ΔChEC153-2* (4 dpi: $p = 0.877$; 5 dpi: $p = 0.997$) relative to *ΔChKu80*. I observed similar results when lesions were measured as chlorotic lesion areas and lesion lengths at four and five dpi (**Appendix 2 Fig. B**, statistical test results reported in **Appendix 2 Tab. C**). The lesion sizes resulting from infection at four- and five-days post inoculation only differed relative to the background strain for strain *ΔChEC153-1*, which as shown in **Figure 4.7**, already presents a growth defect outside of the infection context. The potential *in planta* loss-of-virulence phenotype seen for *ΔChEC153-1* cannot be disentangled from the *in vitro* growth defect. As no significant difference in lesion size was observed between the background strain and strain *ΔChEC153-2,* I cannot conclude there to be any impact on virulence of *C. higginsianum* losing the putative effector gene *ChEC153* under these conditions.

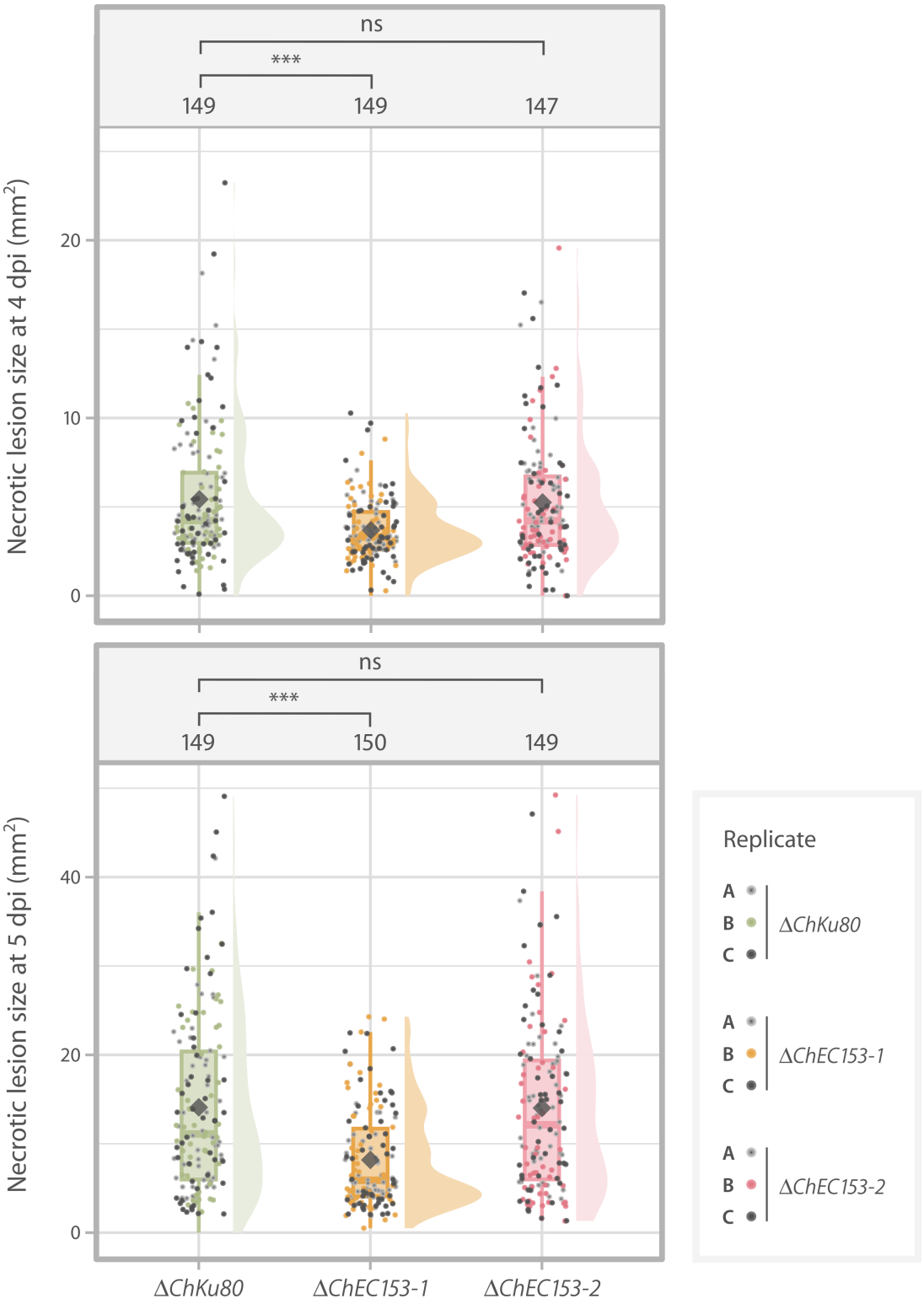


Figure 4.9: ***ΔChEC153*** **strains show no clear virulence phenotype *in planta* relative to *ΔChKu80* that cannot be accounted for by *in vitro* growth defects.** Necrotic lesion areas measured at four and five dpi, with each individual replicate of the experiment plotted as differently coloured datapoints. Data were analysed using an ANOVA on a linear mixed effects model, with the individual replicate of the experiment included as a random factor, and Tukey HSD post hoc tests. Asterisks denote significant differences between background and mutant genotypes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$ (not significant). Values of n are shown above each plot, and mean lesion sizes represented by grey diamonds. Full statistical test results shown in **Appendix 2 Tab. C**.

While no difference was seen in the total starch content of Arabidopsis lines or *N. benthamiana* tissue expressing *ChEC153*, I sought to look for localised changes in starch distribution on a macroscopic level during infection using Lugol to stain the leaf starch. I harvested samples of Col-0 and the *ChEC153-eGFP* line 2-9 at five-days post inoculation at the end of the photoperiod and cleared them in 80% EtOH. I then stained samples by incubating them submerged in 100% Lugol solution overnight and rinsed the samples with water. I took photographs of the samples at each stage of the staining process, as shown in Figure 4.10 A. Representative images of samples are shown in Figure 4.10 B, with images of the same sample approximately aligned such that they can be compared. I found that variability seen between samples of the same genotype made clear differences between genotype difficult to assess. Aside from the slightly larger lesion size seen in the *ChEC153*-expressing line indicative of its increased susceptibility, leading to larger areas of clearing of starch adjacent to the lesion (in the chlorotic ring surrounding each lesion), no differences between genotype were apparent.

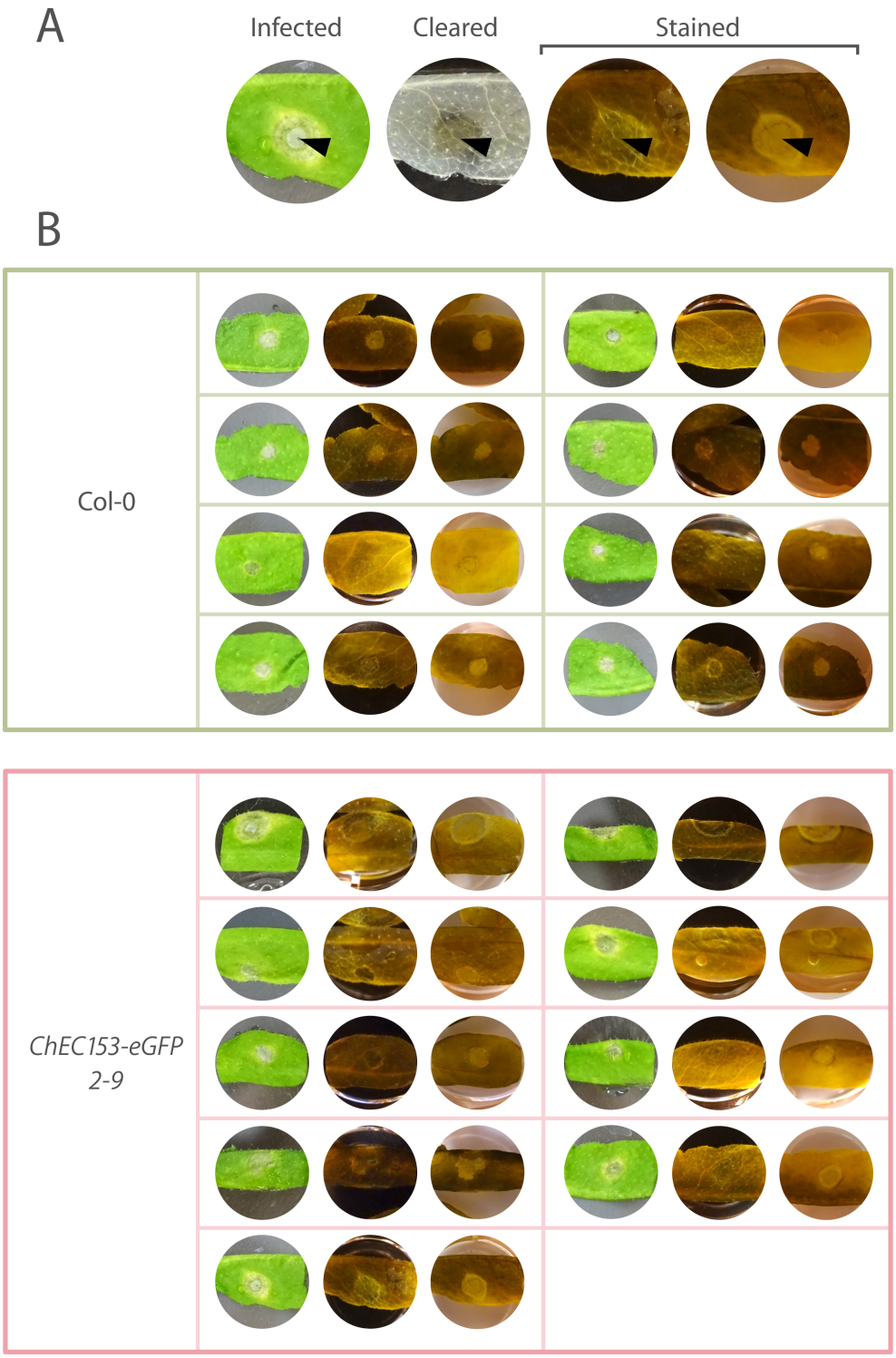


Figure 4.10: **Leaves of Arabidopsis lines expressing *ChEC153* do not have a clear macroscopic starch phenotype during infection compared to Col-0.** Attached leaves of each genotype (Col-0 and *ChEC153-eGFP* expressing line 2-9) were infected with the wild-type *C. higginsianum* strain IMI 349061, and leaf sections harvested at five-days post inoculation at the end of the photoperiod before staining with Lugol solution to visualise starch. **(A)** Example lesion following infection, clearing with ethanol, and staining with Lugol (against either a black or white background). Black arrowheads indicate the site of drop-inoculation. **(B)** Photographs are aligned show samples before and after staining.

Similarly, I carried out an analogous Lugol staining experiment using the *C. higginsianum ChEC153* knockout strains generated in this chapter. I collected leaf samples of Arabidopsis rosettes drop-inoculated with strains *ΔChKu80*, *ΔChEC153-1*, and *ΔChEC153-2* five days after inoculation. I cleared, stained, and photographed samples as described for infected *ChEC153*-expressing Arabidopsis lines. Images of these infected, Lugol-stained samples are presented in **Figure 4.11**. I again observed a large amount of variation between samples, and no clear difference between strains. Therefore, the starch phenotype during infection was not further analysed here.

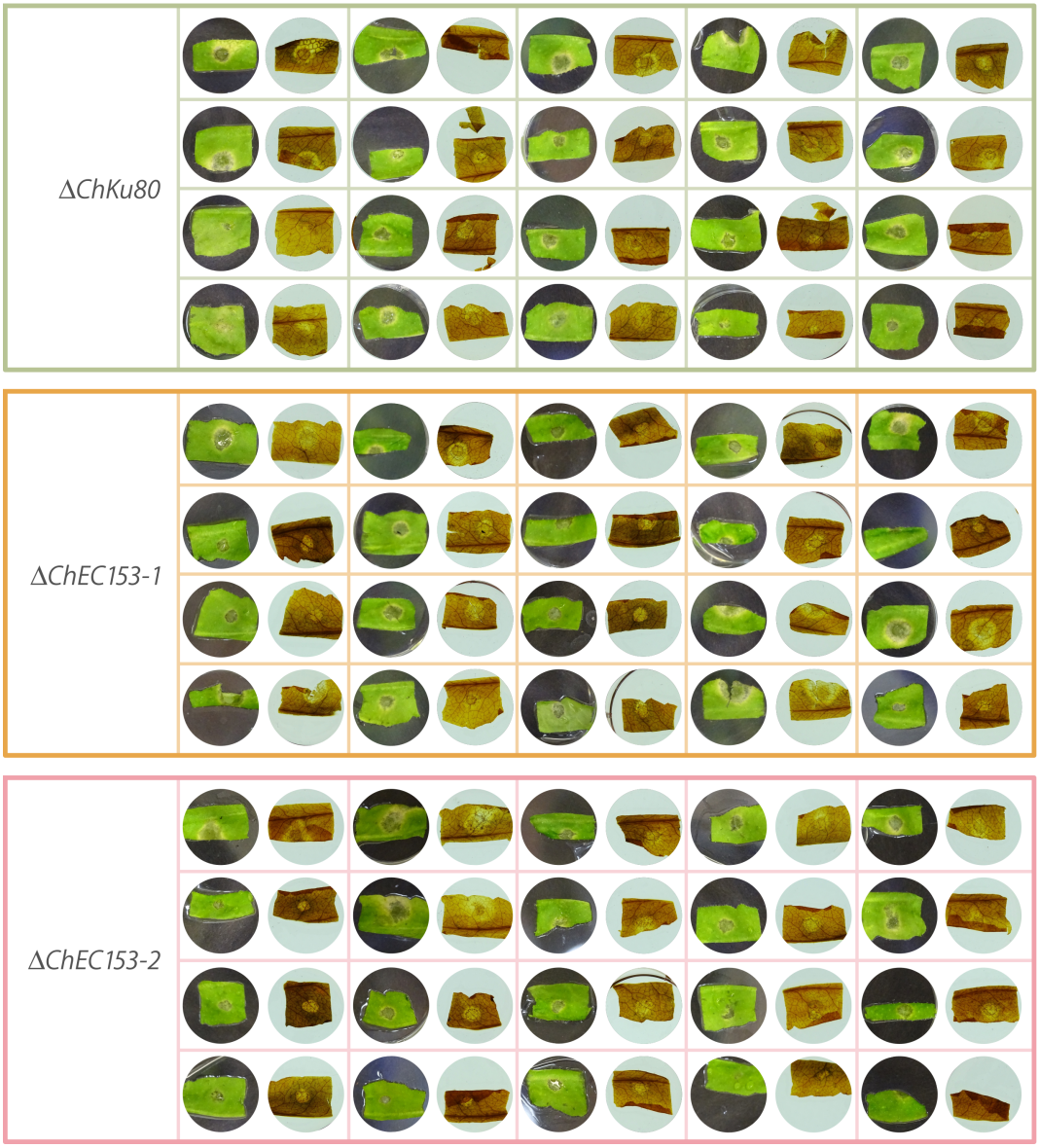


Figure 4.11: **Arabidopsis leaves infected with *ΔChEC153* strains have no clear macroscopic starch phenotype compared to *ΔChKu80* infection based on Lugol staining.** Attached Arabidopsis leaves infected with *C. higginsianum* strains *ΔChKu80*, *ΔChEC153-1*, and *ΔChEC153-2* and leaf sections harvested at five-days post inoculation at the end of the photoperiod. Photographs are aligned show each sample before (left-hand side) and after (right-hand side) staining.

4.3.3 *mrc* plants show increased resistance to *C. higginsianum*

Given the co-localisation of ChEC153 and MRC (Figure 3.8), I explored whether mutant Arabidopsis lines lacking *MRC* would have any differential susceptibility phenotype to *C. higginsianum* infection. It is known that some Arabidopsis starch mutants have altered resistance/susceptibility phenotypes (Engelsdorf et al., 2013, Engelsdorf et al., 2017), but this has not been investigated in starch granule initiation mutants. In order to test this, I inoculated Col-0 and *mrc* Arabidopsis leaves with the wild-type *C. higginsianum* strain IMI 349061 and measured the resulting lesions. The *mrc* mutant line used here has no clear morphological differences compared to the wild-type when grown under these conditions. The *mrc* mutant leaves showed a reduced necrotic lesion size compared to Col-0 leaves at both four- and five-days post inoculation with *C. higginsianum*, with data analysed using an ANOVA on a linear mixed effects model with the individual replicate of the experiment included as a random factor (4 dpi: $p < 0.001$; 5 dpi: $p = 0.010$). These data are presented in Figure 4.12, with the independent replicates of the experiment shown by differently coloured datapoints. Similarly, when measured as chlorotic lesion areas and lesion lengths, lesions on *mrc* leaves were generally smaller than seen on wild-type leaves (chlorotic lesion areas: 4 dpi: $p = 0.017$; 5 dpi: $p = 0.018$; lesion lengths: 4 dpi: $p = 0.021$; 5 dpi: $p = 0.285$; Appendix 2 Fig. C, statistical test results shown in Appendix 2 Tab. D). In summary, these results indicate a reduced susceptibility of Arabidopsis to *C. higginsianum* infection in the absence of MRC.

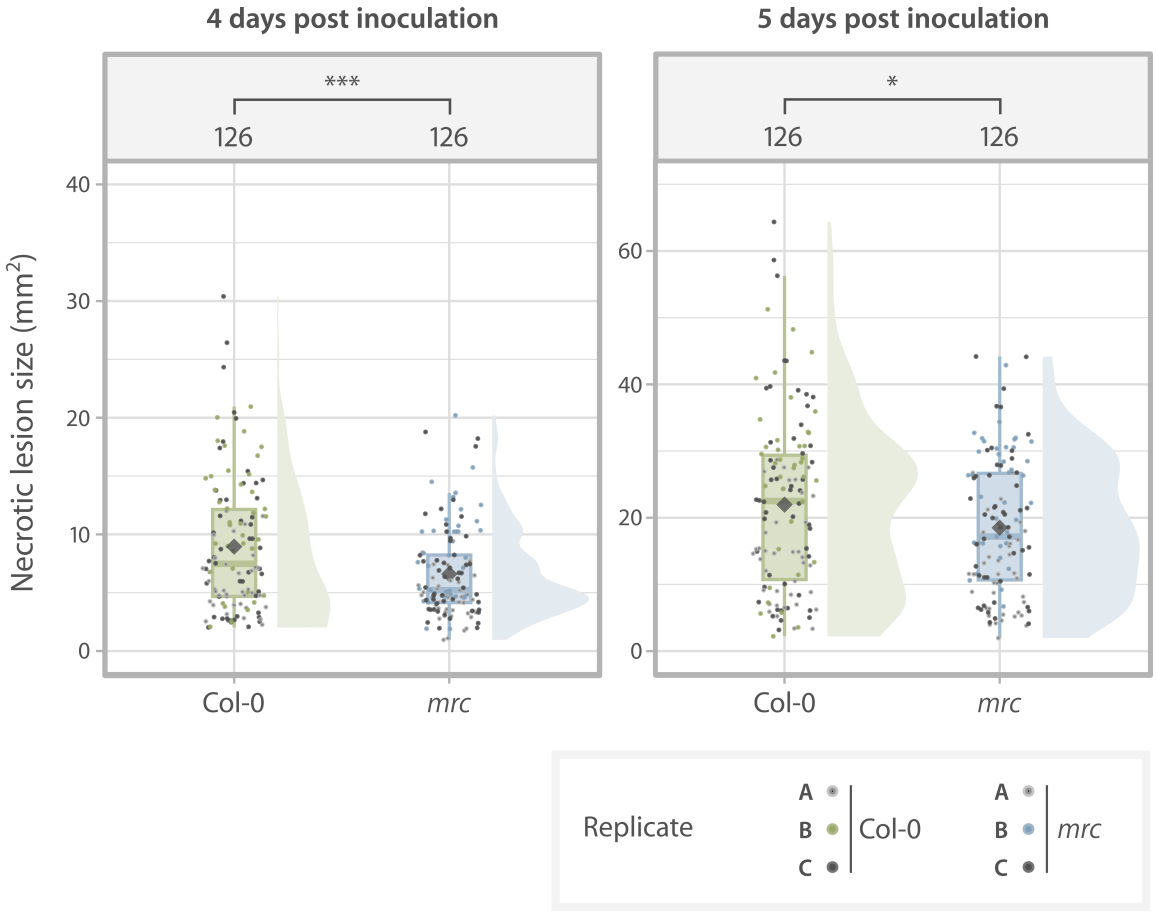


Figure 4.12: ***mrc* Arabidopsis mutants are more resistant to infection by *C. higginsianum* than Col-0.** Necrotic lesion areas (mm^2) four and five days after infection. Infection assay data for each individual replicate of the experiment are plotted as differently coloured points. Data were analysed using a t-test using Satterthwaite's method on a linear mixed effects model, with the individual replicate of the experiment included as a random factor. Asterisks denote significant differences between genotypes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Values of n are shown above each plot, and mean lesion sizes for each are represented by grey diamonds. Full statistical test results shown in **Appendix 2 Tab. D**.

4.3.4 IP-MS reveals potential interactors of ChEC153 in the Arabidopsis host

In order to identify potential host targets or interactors of ChEC153, I used immunoprecipitation-mass spectrometry (IP-MS) as an exploratory, untargeted approach. I extracted proteins of Arabidopsis leaves expressing *35S::ChEC153-eGFP* and Col-0 controls and immunoprecipitated proteins with α-GFP beads. I prepared samples for MS, which was then carried out by the JIC proteomics facility. The seven prey proteins identified as potential interactors of ChEC153 in Arabidopsis are listed in **Table 4.1**, several of which are known or expected to localise to the chloroplast.

Table 4.1: Candidate interactors of ChEC153 identified by IP-MS. Proteins of *ChEC153-eGFP*-expressing and Col-0 Arabidopsis lines were immunoprecipitated with α-GFP beads and subject to mass spectrometry. To identify candidates, exclusive unique peptide counts were used with a 25% (0.25) Bayesian false discovery rate (BFDR) cut-off in SAINTexpress. Eight prey proteins were identified as preferentially pulled down in *ChEC153-eGFP* samples relative to Col-0 (n=3), including the bait protein ChEC153-eGFP.

Prey	SAINT score	Fold change	BFDR	Protein description (TAIR)
AT5G56500.1	1	9.094	0	CHAPERONIN 60 BETA3 (CPNB3)
ChEC153-eGFP	1	31309.180	0	N/A
AT5G05470.1	0.740	8.363	0	EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT (EIF2-A2)
AT1G53120.1	0.666	7.666	0.087	RNA-BINDING S4 DOMAIN CONTAINING PROTEIN
AT5G23310.1	0.666	4.970	0.149	FE SUPEROXIDE DISMUTASE (FSD3)
AT1G63680.1	0.650	6.935	0.186	MURE, ALBINO OR PALE-GREEN 13 (APG13)
AT1G22060.1	0.620	7.200	0.213	SPORULATION-SPECIFIC PROTEIN
AT1G49670.1	0.605	6.950	0.237	NQR

In order to evaluate the likelihood that these candidate proteins pulled down by ChEC153 in IP-MS are true interactors of putative *C. higginsianum* effector, I first sought to elucidate their subcellular localisations. Promisingly, a number of these proteins are predicted to localise to the chloroplast using *in silico* localisation prediction tools, or are seen to localise to the chloroplast in the literature. To confirm the protein localisations, the coding sequences of the three most highly enriched IP-MS prey proteins identified: AT5G56500.1 (CHAPERONIN-60BETA3; CPNB3), AT5G05470.1

(EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT; EIF2-A2), and AT1G53120.1 (RNA-BINDING S4 DOMAIN-CONTAINING PROTEIN) were synthesised and I cloned these sequences for *Agrobacterium*-mediated transient expression in *N. benthamiana* with a C-terminal mCherry tag. Three days after *N. benthamiana* infiltration, I established the localisations of these ChEC153 interaction candidates using confocal microscopy.

4.3.5 Candidate ChEC153-interactor CPNB3 localises to the chloroplast, while EIF2-A2 localises to the cytoplasm

CPNB3 is expected to localise to the chloroplast according to protein sequence analysis using TargetP 2.0 (Armenteros et al., 2019a) which predicted a thylakoid luminal transfer peptide. Similarly, DeepLoc 1.0 predicts CPNB3 to be a soluble plastid-localised protein (Armenteros et al., 2017). Indeed, I found CPNB3 to localise to the chloroplasts, with a diffuse localisation typical of a stromal protein (**Figure 4.13 A**). This localisation was consistent upon co-expression with either putative effector (ChEC153 and HaRxL94b; **Figure 4.13 B** and **C**, respectively).

Contrastingly, no localisation for EIF2-A2 was predicted by TargetP 2.0 (Armenteros et al., 2019a), whereas DeepLoc 1.0 predicted this to be a soluble cytoplasmic protein (Armenteros et al., 2017). Confirming the DeepLoc 1.0 prediction, when I employed *Agrobacterium*-mediated transient expression to express EIF2-A2 in *N. benthamiana*, EIF2-A2 localised to the cytoplasm when expressed alone (**Figure 4.14 A**) and when co-expressed with either *ChEC153* or *HaRxL94b* (**Figure 4.14 B** and **C**).

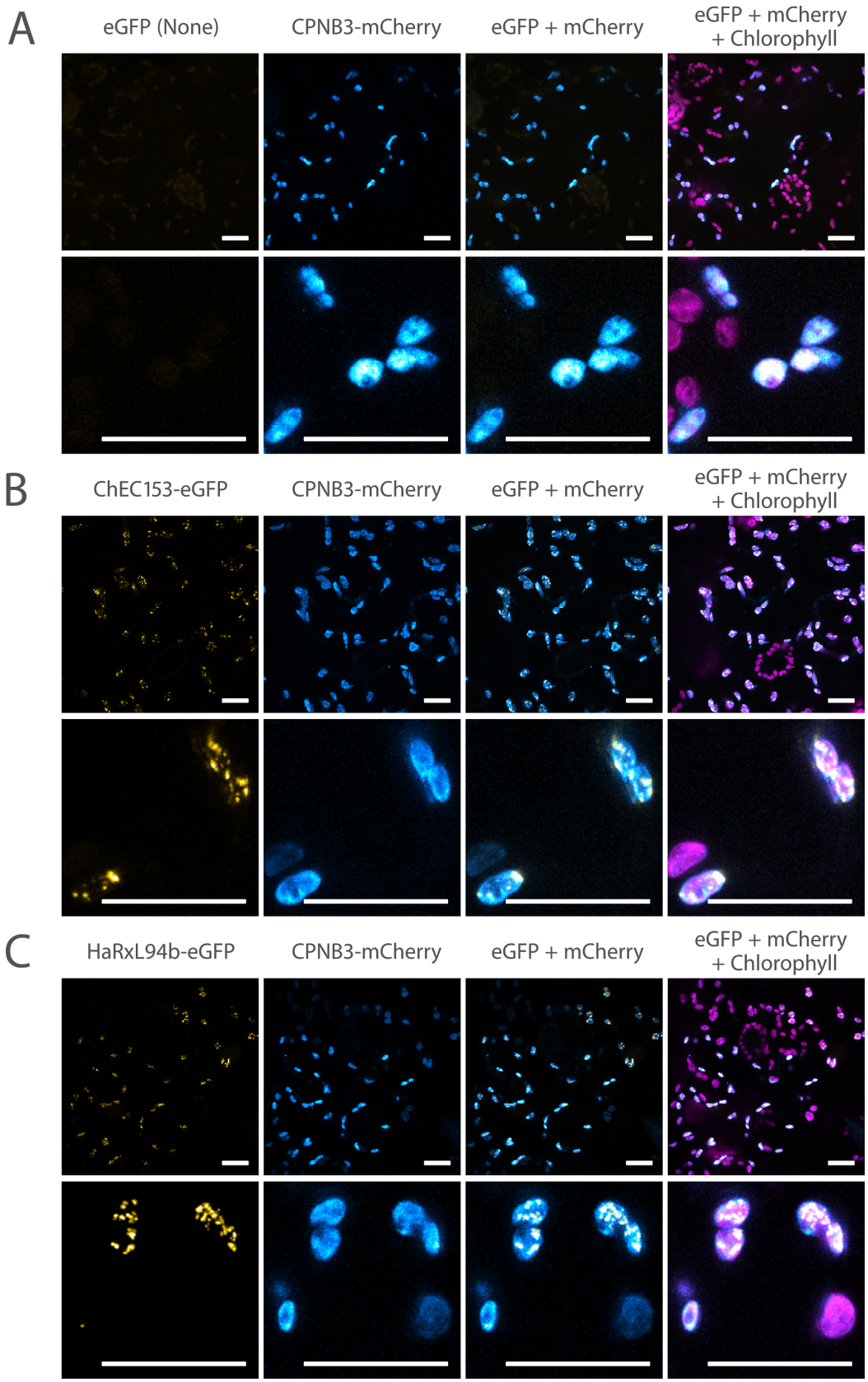


Figure 4.13: **AtCPNB3 localises to the chloroplast.** *Agrobacterium*-mediated transient expression of *35S::CPNB3-mCherry* in *N. benthamiana* **(A)** alone; **(B)** co-expressed with *35S::ChEC153-eGFP*; and **(C)** co-expressed with *35S::HaRxL94b-eGFP*. eGFP signal is shown in yellow, with mCherry in cyan and chlorophyll autofluorescence in magenta. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

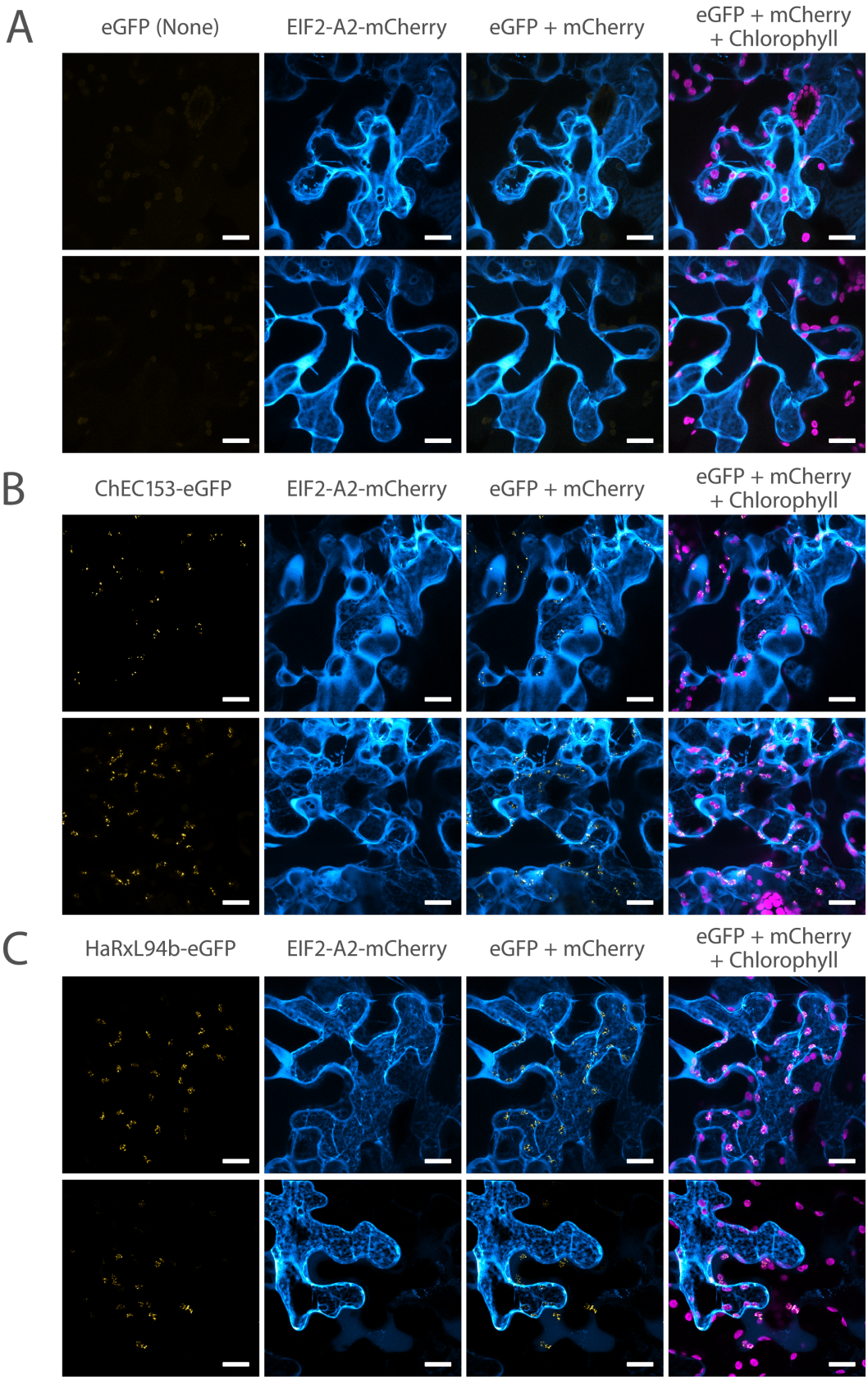


Figure 4.14: **AtEIF2-A2 localises to the cytoplasm.** *Agrobacterium*-mediated transient expression of *35S::EIF2-A2-mCherry* in *N. benthamiana* **(A)** alone; **(B)** co-expressed with *35S::ChEC153-eGFP*; and **(C)** co-expressed with *35S::HaRxL94b-eGFP*. eGFP signal is shown in yellow, with mCherry in cyan and chlorophyll autofluorescence in magenta. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

4.3.6 Candidate ChEC153-interactor CRBIC is recruited to ChEC153 puncta

The protein encoded at *AT1G53120.1* is as-yet unnamed, and simply described in TAIR as an RNA-binding S4 domain-containing protein. For brevity, I have named this protein for the purposes of this thesis as CHLOROPLASTIC RNA-BINDING INTERACTOR OF ChEC153 (CRBIC).

CRBIC is predicted by TargetP 2.0 to localise to the mitochondria, with a chloroplast transfer peptide being predicted with lower confidence (mitochondrial transfer peptide: 0.6005; chloroplast transfer peptide: 0.3886; Armenteros et al., 2019a). Conversely, DeepLoc 1.0 predicted CRBIC to be a soluble plastid-localised protein, returning a higher likelihood for plastid localisation (0.6155) than mitochondrial localisation (0.3834; Armenteros et al., 2017). Similarly to CPNB3 (**Figure 4.13 A**), I found that CRBIC also appeared to localise to the chloroplast stroma when transiently expressed alone in *N. benthamiana* (**Figure 4.15 A**). However, and strikingly, upon co-expression with *ChEC153-eGFP*, the localisation of CRBIC shifted from diffuse to punctate (**Figure 4.15 B**). These induced CRBIC-mCherry puncta co-localise with the ChEC153-eGFP puncta. Thus, ChEC153 appears to recruit CRBIC to puncta within the chloroplast, confirming that CRBIC interacts with ChEC153. Co-expression with HaRxL94b did not appear to alter the localisation of CRBIC (**Figure 4.15 C**). The shift in CRBIC localisation seen with ChEC153 is therefore somewhat specific, and may reflect the fact that this interaction candidate was pulled down by ChEC153 in IP-MS. This may also imply that the two putative effectors ChEC153 and HaRxL94b may have different targets within the host, despite both co-localising with MRC (**Figure 3.8**).

Based on the co-localisation of ChEC153 with both MRC (**Figure 3.8**) and TaCSP41a (**Figure 3.9**), and the recruitment of CRBIC to chloroplastic puncta in the presence of ChEC153, I tested whether MRC and/or TaCSP41a are also able to alter the localisation of CRBIC during transient co-expression in *N. benthamiana*. As shown in **Figure 4.16**, I found that while *CRBIC* co-expressed with *eGFP* targeted to the chloroplast by the N-terminal fusion of the chloroplast transit peptide of the small subunit of RuBisCO (*35S::cTP-eGFP*) remains diffuse (**Figure 4.16 A**), either *MRC* or *TaCSP41a* co-expression is sufficient to recruit CRBIC to puncta within the chloroplast (**Figure 4.16 B** and **C**).

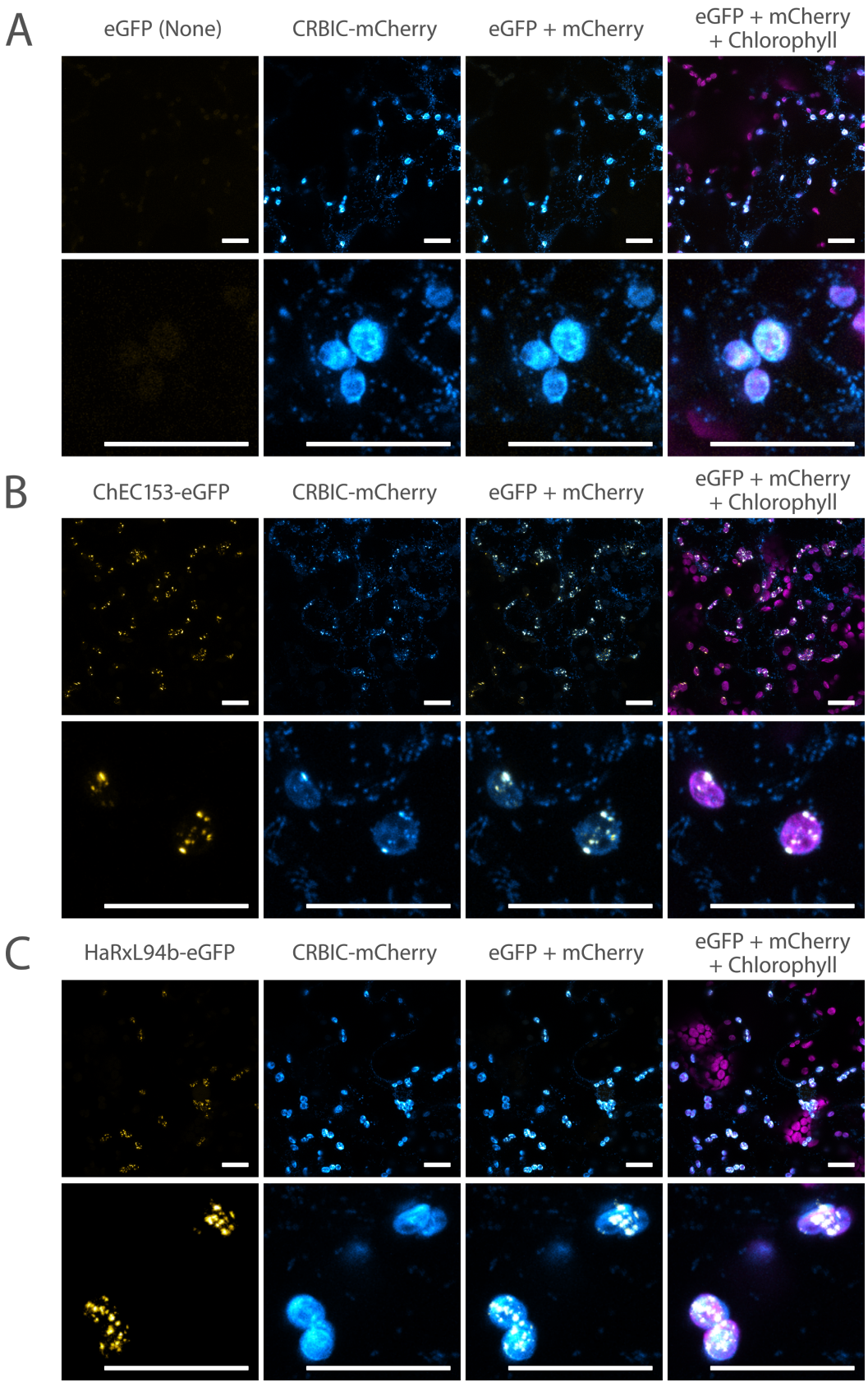


Figure 4.15: AtCRBIC localises to the chloroplast and is recruited to puncta by ChEC153. *Agrobacterium*-mediated transient expression in *N. benthamiana* of *35S::CRBIC-mCherry* **(A)** alone; **(B)** co-expressed with *35S::ChEC153-eGFP*; and **(C)** co-expressed with *35S::HaRxL94b-eGFP*. eGFP signal shown in yellow, with mCherry in cyan and chlorophyll autofluorescence in magenta. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

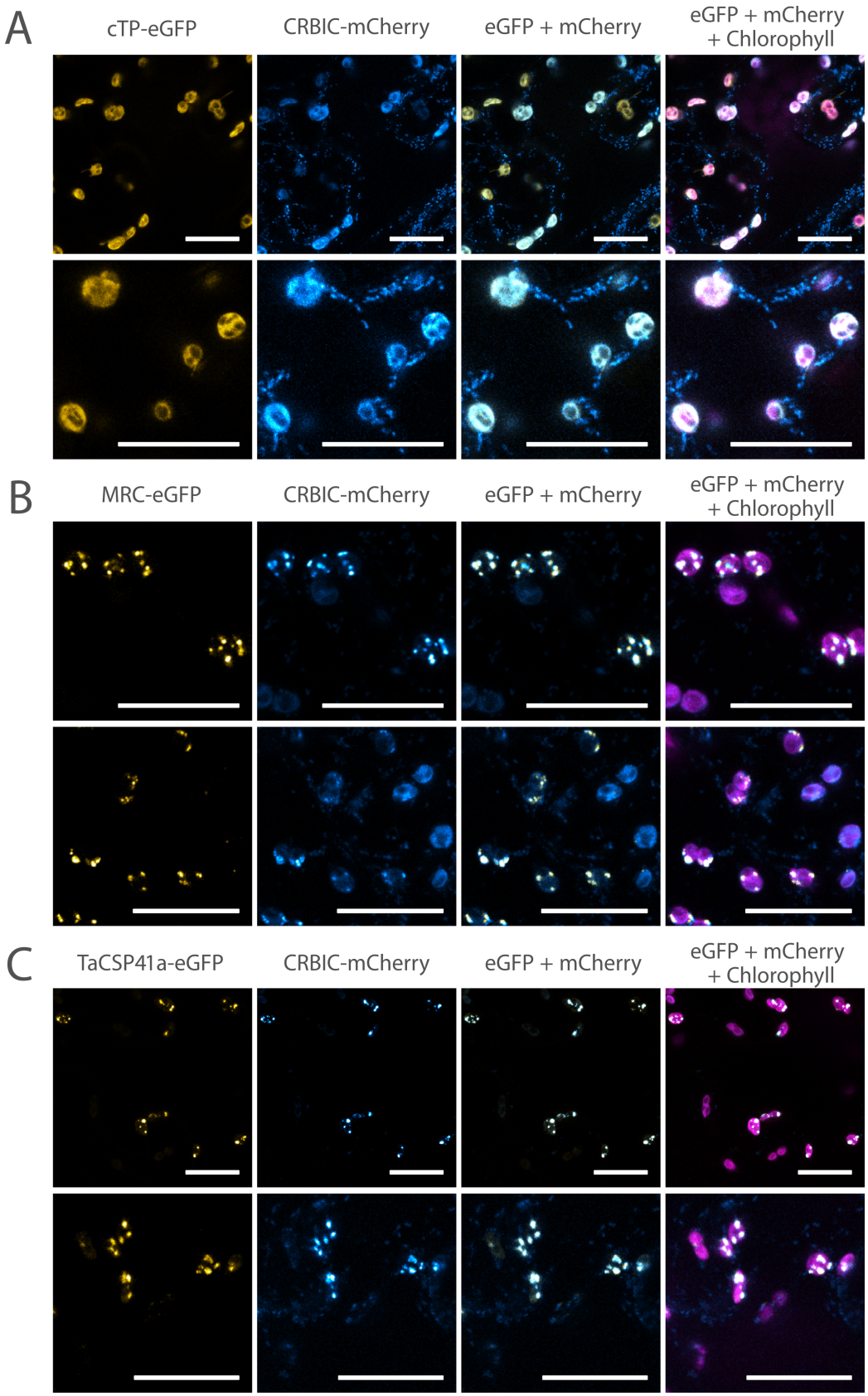


Figure 4.16: **CRBIC is recruited to chloroplastic puncta during *CRBIC* co-expression with *MRC* or *TaCSP41a*.** Confocal images showing *Agrobacterium*-mediated transient expression of 35S::CRBIC-mCherry in N. *benthamiana* when co-expressed with **(A)** *35S::cTP-eGFP*, **(B)** *35S::MRC-eGFP*, or **(C)** *35S::TaCSP41a-eGFP*. eGFP signal shown in yellow, with mCherry in cyan and chlorophyll autofluorescence in magenta. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

4.3.7 CRBIC localises to mitochondria, as well as chloroplasts

When expressing *CRBIC-mCherry* in *N. benthamiana*, in addition to the chloroplastic localisation, I also observed mCherry signal at small points in the cytoplasm that resembled mitochondria (**Figure 4.15**). Though the fluorescent signal in these points was comparatively faint, this dual localisation would be in keeping with the *in silico* predictions of the localisation of CRBIC. An illustrative confocal image of this localisation is presented in **Figure 4.17 A**. In order to confirm whether these regions of mCherry signal represent mitochondria, I co-expressed *CRBIC-mCherry* with a mitochondrial GFP marker (CoxIV sequence targeting GFP to mitochondria: TSL SynBio; pICSL11247; 35S::Mit localisation signal:GFP:Nos) in *N. benthamiana* using *Agrobacterium*-mediated transient expression. As shown in **Figure 4.17 B**, CRBIC-mCherry was seen to co-localise with the Mit-GFP mitochondrial marker, confirming that CRBIC localises to mitochondria as well as chloroplasts.

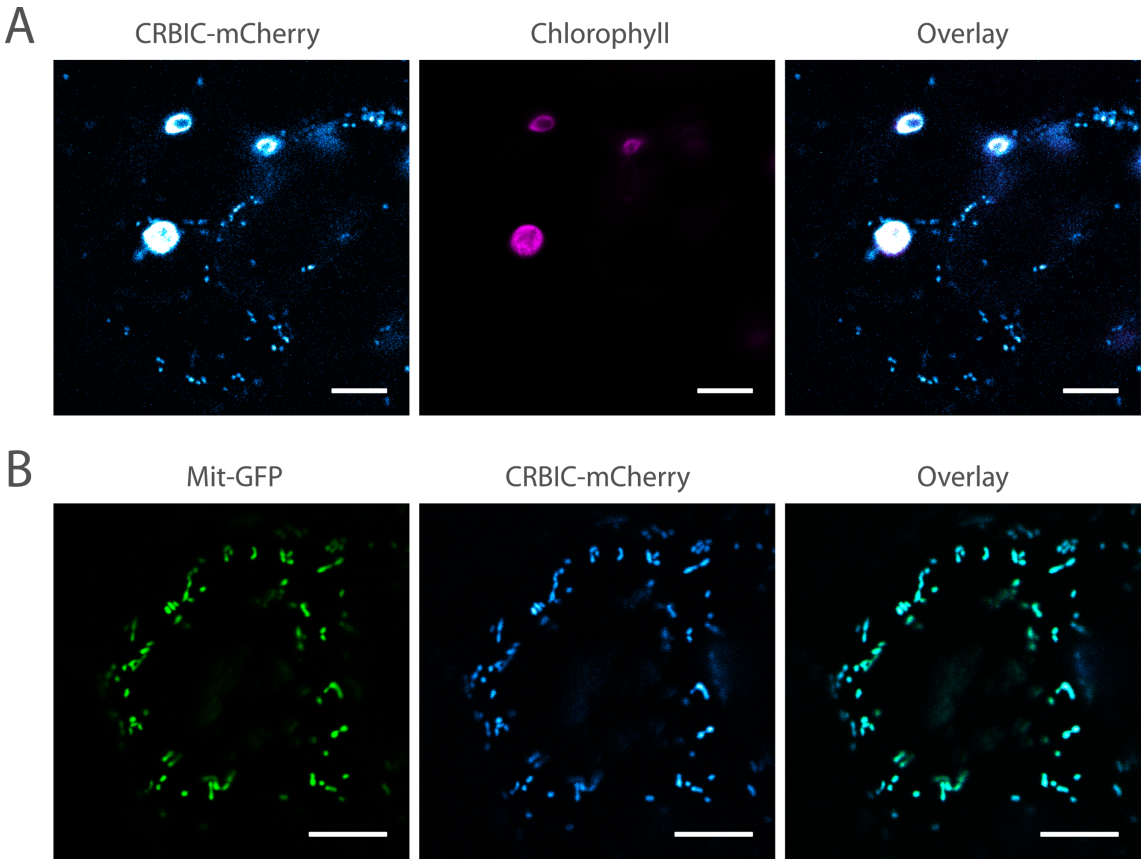


Figure 4.17: **CRBIC localises to mitochondria, in addition to chloroplasts.** Confocal images showing *Agrobacterium*-mediated transient expression in *N. benthamiana* of **(A)** *35S::CRBIC-mCherry* expressed alone; and **(B)** co-expressed with GFP mitochondrial marker *35S::Mit-GFP*. mCherry signal is shown in blue, with chlorophyll autofluorescence in magenta and GFP in green. Scale bars, 10 µm.

The localisation of CRBIC at mitochondria as well as the chloroplast prompted me to re-examine the localisation of ChEC153 to see whether any secondary targeting of mitochondria was observable. The punctate chloroplastic signal from the putative effectors generally appeared very bright in this system, and so I thought it possible that faint mitochondrial signal may have been missed with my imaging settings. Indeed, adjusting the brightness of images of *ChEC153-eGFP* expressed in *N. benthamiana* revealed faint signal at cytoplasmic puncta which may represent mitochondria. To validate this, I co-expressed *ChEC153-mCherry* and the mitochondrial GFP marker in *N. benthamiana*. I observed clear overlap between mCherry and GFP signals, confirming that ChEC153 localises at mitochondria, albeit at a significantly lower level than the accumulation seen in chloroplasts (**Figure 4.18**).

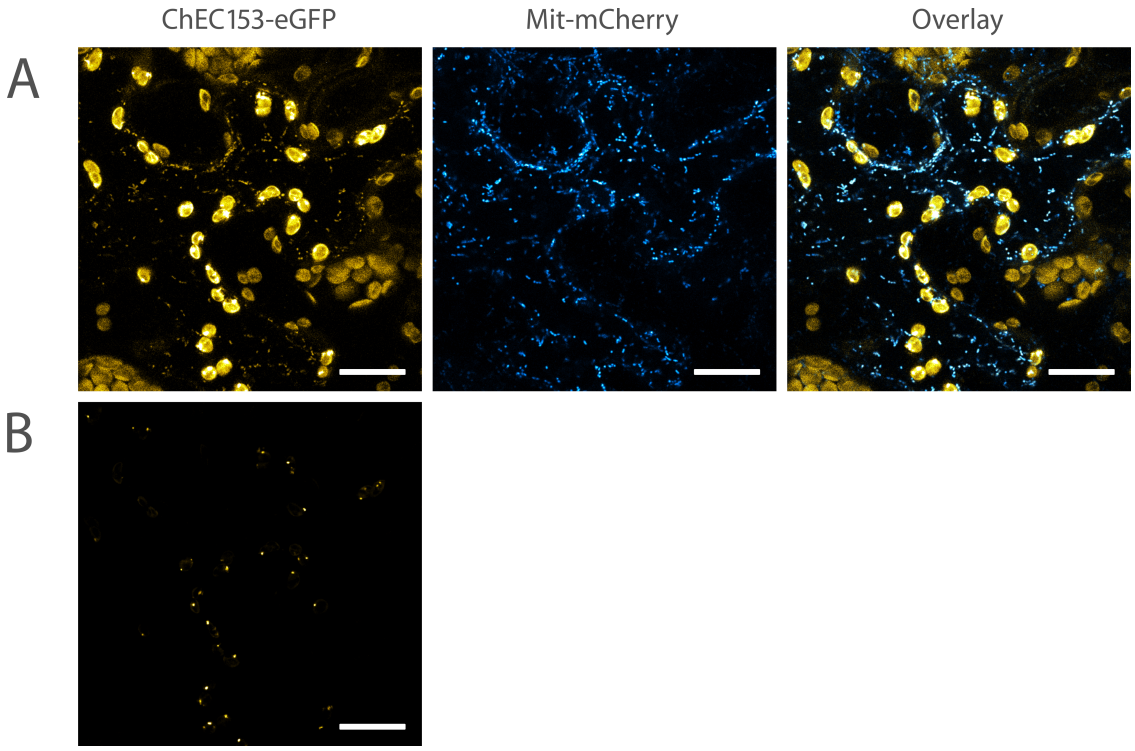


Figure 4.18: **ChEC153 localises to mitochondria at a low level. (A)** Confocal images taken during *Agrobacterium*-mediated transient expression of *ChEC153-eGFP* in *N. benthamiana,* showing *35S::ChEC153-eGFP* co-expression with *35S::mit-mCherry*. **(B)** The same image of ChEC153-eGFP as in (A), before adjusting the brightness in ImageJ. eGFP signal shown in yellow, with mCherry in cyan. Images are maximum intensity projections of z-stacks. Scale bars, 20 µm.

4.3.8 *crbic* Arabidopsis mutants do not have altered susceptibility to *C. higginsianum*

To inform on the role of CRBIC in Arabidopsis, I acquired *AT1G53120* (*crbic*) mutant Arabidopsis lines SALK_041981C and SALK_099429C from NASC (Alonso et al., 2003). The T-DNA insert in line SALK_041981C is expected to be within the 5' UTR, while the insert in SALK_099429C is in the second exon of the coding sequence (Figure 4.19 A). I extracted gDNA from these lines alongside a Col-0 control, and carried out PCR genotyping to confirm whether they are true mutants, and whether they are homozygous or heterozygous for the mutation of *CRBIC*.

To identify homozygous mutants in the SALK_041981C line, I used SALK_041981C gDNA as a template for gene-specific reactions (using SALK_041981C primers 071 and 072), and T-DNA insert reactions (using primers SALK_LB and 072). For SALK_099429C lines, I used SALK_099429C gDNA as a template, with SALK_099429C gene-specific primers 069 and 070 to amplify the wild-type *CRBIC* gene, and primers SALK_LB and 070 to amplify from the T-DNA insert (Figure 4.19 B; primers detailed in Appendix 1 Tab. E). While I found all seedlings of the SALK_041981C line that I tested lacked the expected T-DNA insertion, SALK_099429C seedlings were homozygous for the *CRBIC* mutation as shown by the loss of the control, wild-type gene-specific amplicon (Figure 4.19 C). Amplification of gDNA from a Col-0 control plant using test primers SALK_LB and 070 gave a faint, non-specific band of a smaller size than the T-DNA PCR band. Line SALK_099429C is henceforth referred to as *crbic* within this work.

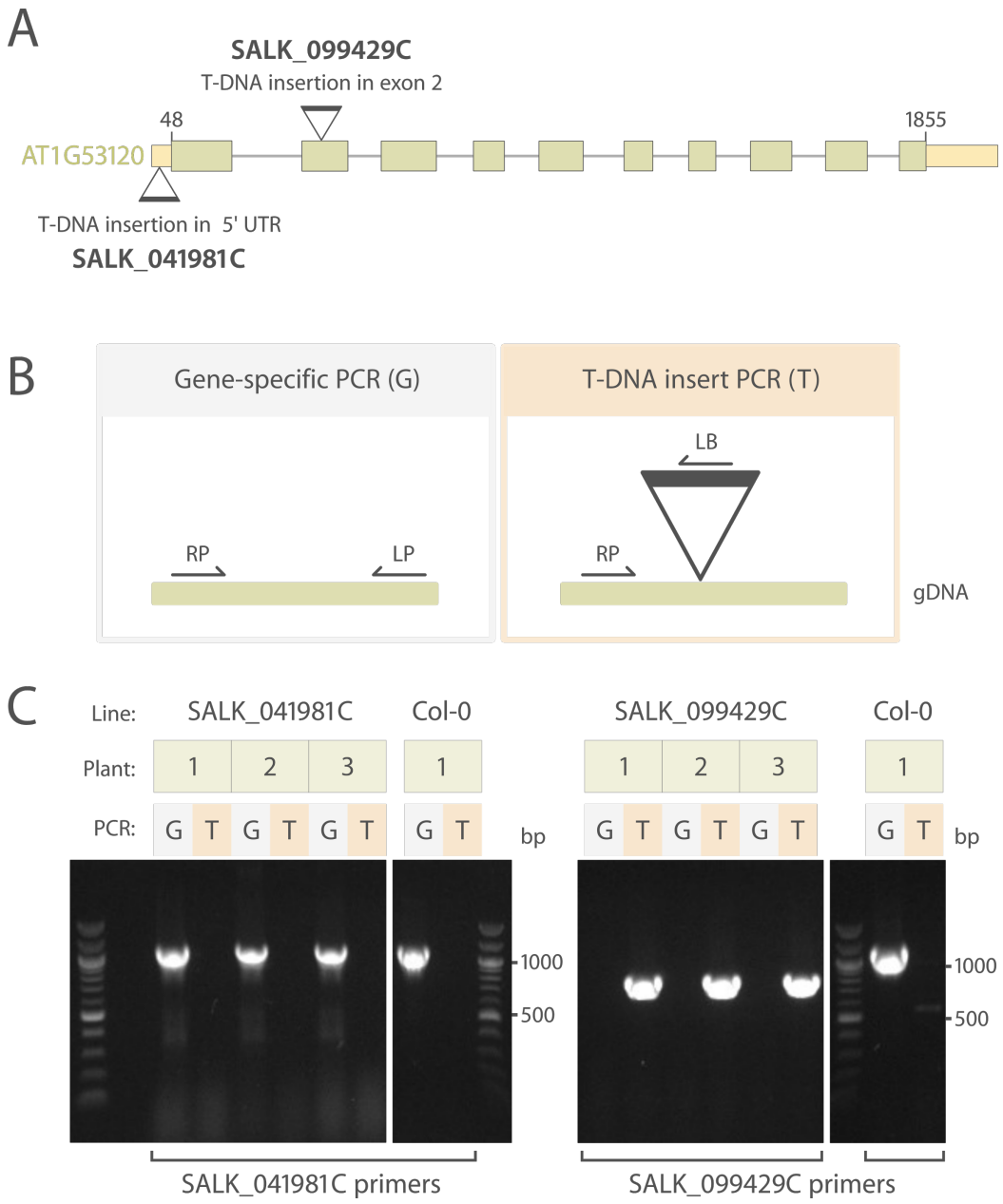


Figure 4.19: **Confirming T-DNA insertion in *CRBIC (AT1G53120)* in SALK_099429C plants. (A)** Gene model for *CRBIC* showing the site of T-DNA insertions for SALK_041981C and SALK_099429C Arabidopsis lines. 5′ and 3′ UTRs shown in yellow, with exons highlighted in green. **(B)** Genotyping PCR reaction schematic showing reactions testing for presence of the wild-type gene (G) and T-DNA insertion (T) in extracted genomic DNA (gDNA). **(C)** Agarose gel showing products of genotyping PCRs for gDNA extracted from three potential mutant plants and one Col-0 plant as a control with either primer set (indicated below the gels). Expected amplicon size for SALK_041981C primer reactions: gene-specific reaction (G, wild-type): 1114 bp; T-DNA insert reaction (T, mutant): 512-812 bp. Expected amplicon size for SALK_041981C primer reactions: gene-specific reaction (G, wild-type): 1163 bp; T-DNA insert reaction (T, mutant): 533-833 bp. PCR products were separated on a 1% agarose gel stained with EtBr alongside a 100 bp ladder (NEB).

I carried out infection assays with *C. higginsianum* to determine whether the loss of *CRBIC* in Arabidopsis affects the susceptibility of the host. I hypothesised that *crbic* lines may have an altered susceptibility to infection should the putative effector ChEC153 be unable to interact with candidate host target CRBIC.

I analysed infection assay data using a Satterthwaite's method t-test on a linear mixed effects model, with the individual replicate of the experiment included as a random factor (plotted as differently coloured points). I found no difference in necrotic lesion sizes between Col-0 and *crbic* genotypes at either four- or five-days post inoculation, with data presented in Figure 4.20 (4 dpi: $p = 0.056$; 5 dpi: $p = 0.524$). Further, I similarly observed no difference in chlorotic lesion size (4 dpi: $p = 0.306$; 5 dpi: $p = 0.729$) or lesion length (4 dpi: $p = 0.287$; 5 dpi: $p = 0.898$) between genotypes at either time-point (data presented in Appendix 2 Fig. D, full statistical test results are presented in Appendix 2 Tab. E). Thus, I conclude that loss of *CRBIC* does not have a clear impact on the susceptibility of Arabidopsis to *C. higginsianum*, despite interaction of CRBIC with *C. higginsianum* putative effector ChEC153.

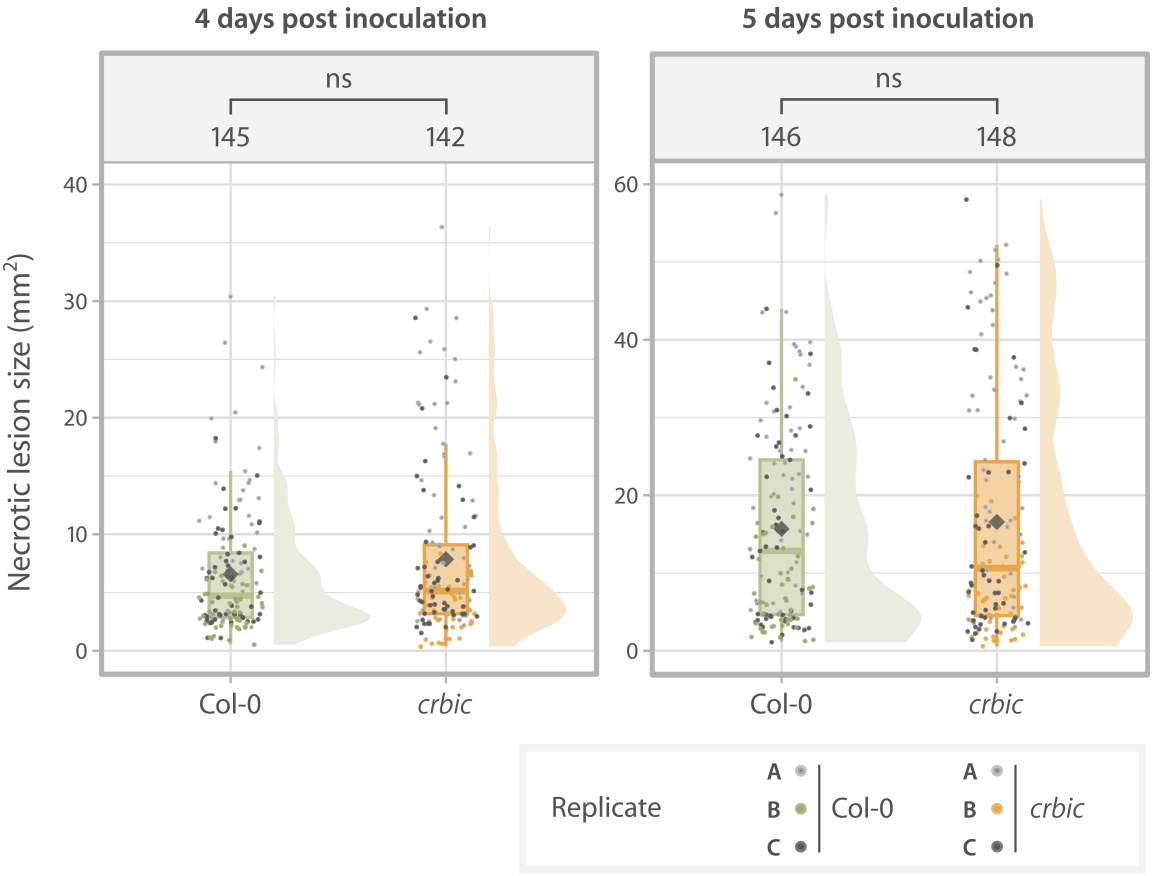


Figure 4.20: ***crbic*** **Arabidopsis mutants do not have altered susceptibility to infection by *C. higginsianum* at four- or five-days post inoculation.** Data shown are necrotic lesion areas in mm^2, with each point representing a single lesion. Infection assay data for individual replicates of the experiment are plotted as differently coloured points, with no significant difference (ns) seen between genotypes as determined by a t-test using Satterthwaite's method on a linear mixed effects model, with individual replicate included as a random factor. Values of n are shown above each plot, and mean lesion sizes are represented by grey diamonds. Full statistical test results shown in **Appendix 2 Tab. E**.

4.3.9 Extraction of ChEC153-eGFP from *N. benthamiana*

In order to carry out co-immunoprecipitation assays to test for direct interactions between ChEC153 and CRBIC, as well as other potential interactors such as MRC, I first sought to purify ChEC153 from *N. benthamiana* tissue transiently expressing *35S::ChEC153-eGFP*. Prior to harvesting samples for protein extraction and Western blotting, I used confocal microscopy to confirm the clear expression of *ChEC153-eGFP*. I tested several volumes of extraction buffer for protein extraction, with eight 8 cm diameter leaf discs being processed in 400, 250, or 125 µL of extraction buffer. The only clear band observed for ChEC153-eGFP samples corresponded to cleaved, free eGFP at around 27 kDa (**Figure 4.21 A**). I then tested whether immunoprecipitation of ChEC153-eGFP using anti-GFP beads would allow enrichment of ChEC153-eGFP such that the full-length protein would be detected by Western blotting. In the resulting anti-GFP Western blot

(Figure 4.21 B), for both the IP sample and the total soluble fraction (IN), free GFP was detected. Additionally, a band became apparent at around 70 kDa. The expected size of ChEC153-eGFP is 84.5 kDa, so whether this band truly represented the full-length putative effector was unclear, despite the band appearing slightly enriched in the IP sample. Following this observation, I made further attempts to detect ChEC153-eGFP, following transient expression in *N. benthamiana*. As shown in Figure 4.21 C, while this approximately 70 kDa band was present in samples containing ChEC153-eGFP, it was also seen in a sample for expression of only *MRC-RFP*, included as a negative control. Therefore, I conclude that this band is a non-specific band that does not represent the full-length ChEC153-eGFP protein.

While protein extracts from Arabidopsis samples stably expressing *35S::ChEC153-eGFP* also indicated some cleavage of the eGFP tag, detection of the full-length ChEC153-eGFP protein was achieved easily with similar extraction conditions, as was shown in Chapter 3 Figure 3.14. However, in *N. benthamiana*, the eGFP tag appears to be entirely cleaved from ChEC153, preventing its detection in protein extracts including following immunoprecipitation (Figure 4.21).

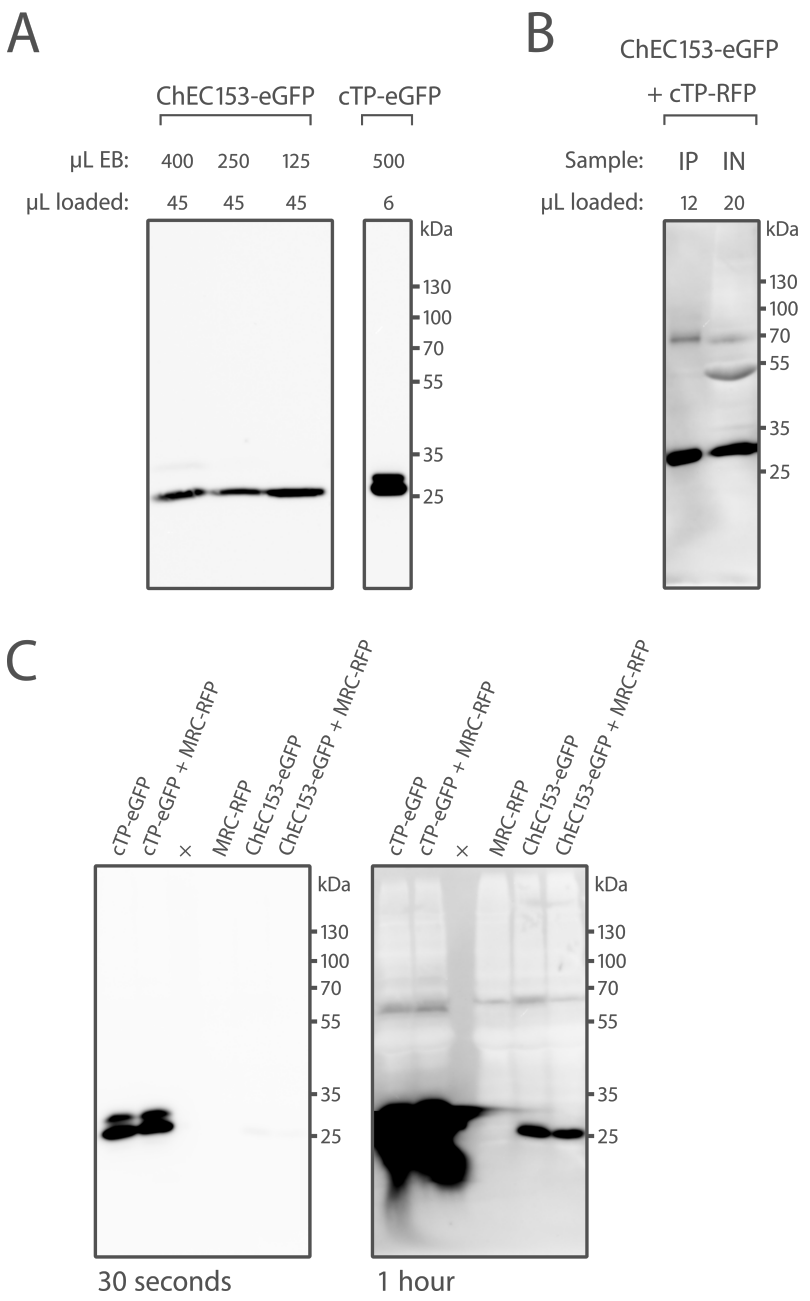


Figure 4.21: **Western blot analyses of *N. benthamiana* protein extracts do not show detection of full-length ChEC153-eGFP.** All blots were probed with an anti-GFP antibody. **(A)** Proteins were extracted from *N. benthamiana* tissue transiently expressing *35S::ChEC153-eGFP* or *35S::cTP-eGFP* in various volumes of extraction buffer (EB) as shown above the blot. The volumes of the soluble fraction of protein extract loaded for SDS-PAGE are also shown (µL loaded). **(B)** Proteins were extracted from *N. benthamiana* tissue transiently co-expressing *35S::ChEC153-eGFP* and *35S::cTP-RFP*. Samples loaded are proteins immunoprecipitated with anti-GFP beads (IP) and a sample of the soluble fraction used as the input (IN) of the immunoprecipitation. **(C)** Proteins were extracted from *N. benthamiana* tissue transiently expressing various constructs, indicated above each lane. Samples were loaded equally for SDS-PAGE and probed with an anti-GFP antibody. The same blot is shown twice, after imaging with 30 seconds or 1 hour of exposure. "×" denotes a lane intentionally left empty. Sizes in kDa are indicated to the right of each blot.

4.4 DISCUSSION

4.4.1 ChEC153 contributes to host susceptibility

Overexpression of *ChEC153* in the Arabidopsis host confers enhanced susceptibility to *C. higginsianum* infection (Figure 4.3), especially early in the infection process, supporting the notion that ChEC153 may function as an effector protein to enhance pathogen virulence or host susceptibility. Based on the results of my infection assays using the two *ChEC153*-knockout strains generated, removal of *ChEC153* from *C. higginsianum* is not sufficient to reduce pathogen virulence (Figure 4.9). *C. higginsianum* has around 365 putative effector proteins (O'Connell et al., 2012), so it may not be surprising that removal of *ChEC153* alone gave no perceptible virulence phenotype on a macroscopic level – it may have only a minor contribution to virulence or else display functional redundancy with other effector proteins.

While I sought to generate two *ChEC153*-knockout lines to assay for changes in virulence in the absence of ChEC153, one of these strains, *ΔChEC153-1*, showed a reduced rate of growth *in vitro* (Figure 4.8). The growth defect seen in *ΔChEC153-1* may reflect off-target integration of the *HygR* cassette perturbing the expression of other genes. Alternatively, there may be considerable differences in the level of *HygR* expression between the two vectors fv *ChEC153* KO 1 and fv *ChEC153* KO 2 due to the difference in promoter - should the expression driven by the cos1gpdA promoter be too high, this may present a metabolic burden and thus result in the observed slower growth of strain *ΔChEC153-1*. Based on the reduced growth rate of this strain *in vitro* (Figure 4.8), the reduced size of lesions it produced *in planta* could not be attributed to a *ChEC153*-specific effect, as the *ΔChEC153-2* strain (with normal *in vitro* growth) produced lesions *in planta* no different in size to the wild-type lesions (Figure 4.9).

No clear starch phenotype was apparent during infection based on Lugol staining of *ChEC153*-expressing Arabidopsis plants infected with wild-type *C. higginsianum* (Figure 4.10), or Col-0 Arabidopsis plants infected with *ChEC153*-knockout *C. higginsianum* strains (Figure 4.11). Due to the variability seen and lack of a clear phenotype, infection-specific starch phenotypes were not probed further. However, there may be differences in total starch distribution visible at a microscopic level, or by starch quantification at different proximities to the lesion.

As introduced in Section 4.1.1, several methods for assaying pathogen infection exist. While here lesion area was used as the predominant metric of infection, other methodologies evaluate lesion diameter, pathogen cell-entry rates and scoring of infection stage at a microscopic level, or pathogen DNA or mRNA levels via quantitative real-time or reverse-transcription PCR as a proxy for relative pathogen biomass (Narusaka et al., 2010). Analysing lesion area provides an affordable

and relatively high-throughput method for evaluating infection, and measurements can be automated through the application of ImageJ thresholds (Tsushima et al., 2019b). However, it should be considered that the sensitivity of this method will not match that of a PCR-based approach and subtle differences in susceptibility may be overlooked, particularly in the early stages of infection and prior to the appearance of visible, macroscopic disease symptoms. Given that *ChEC153* is expressed early in the infection process, at the stages of host penetration and biotrophy (O'Connell et al., 2012, Dallery et al., 2017), at which disease lesions will not yet be apparent, it may be appropriate to employ PCR-based approaches to establish whether ChEC153 effects pathogenicity at these early stages.

It is often observed that loss of a single effector gene from a pathogen is insufficient to confer a clear loss-of-virulence phenotype. This phenomenon is widely accepted for a number of phytopathogens, and high rates of effector functional redundancy may represent a consequence of pathogen evolution; multiple effectors targeting the same host protein/process may allow greater pathogen genomic plasticity in terms of loss of effectors which the host evolves to recognise. Recently, Yan and colleagues demonstrated a competition-based virulence assay for *M. oryzae* using differentially fluorescently tagged wild-type and gene-knockout strains (Yan et al., 2023). In this assay, leaves are inoculated with two strains at equal concentrations, and the ratio of control and knockout strains after inoculation monitored over numerous rounds of infection. In this way, differences in the proportion of the two strains can be monitored to determine relative fitness (Yan et al., 2023). This approach requires the generation of fluorescently tagged strains which can be observed as spores/conidia using confocal microscopy.

Here, I propose an alternate approach to evaluate the competitiveness of strains leveraging the KASP (competitive allele-specific PCR) genotyping assay (He et al., 2014). Using this approach, it may be possible to assay for the relative proportion of background and knockout strain present following infection with a 1:1 mixed population of conidia without the need to generate fluorescently-tagged strains. To do this, alternate primer pairs would be designed to amplify selectively from the background or knockout strain genomic DNA, as illustrated in Figure 4.22 A, with alternate primer tail sequences allowing binding of HEX and FAM oligos following amplification, and a shared non-genotype-specific primer. PCR reactions using the shared primer and both genotype-specific primers would be used to determine the relative amounts of wild-type or knockout template: based on the relative proportion of the strains in the KASP template input, different levels of HEX and FAM fluorescent signals would be apparent. Following infection, infected leaves would be used to harvest fungal DNA to serve as the KASP template – possibly using conidial suspensions obtained from "washing" the leaves in water directly as the DNA template. Background-only and knockout-only inoculum controls would be included, and a known 1:1 strain mixture used as a baseline from which to establish any shifts in strain proportions following *in*

planta growth. An illustration of a theoretical result in which a knockout strain is less competitive than the wild-type is presented in **Figure 4.22 B**.

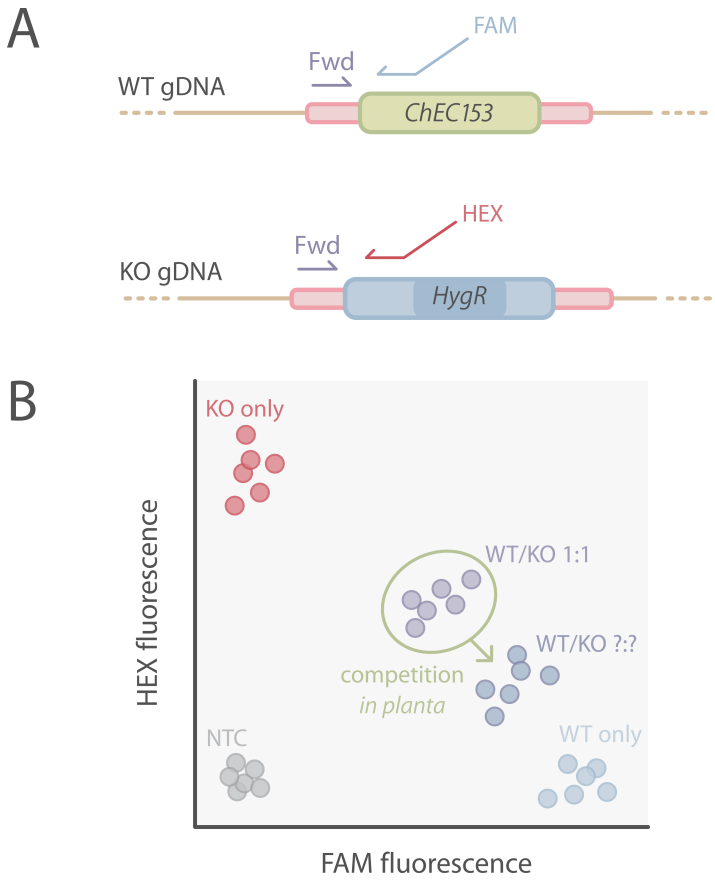


Figure 4.22: **Schematic for theoretical KASP-based competition assay. (A)** Example primer pairs for wild-type (WT) and knockout (KO) strains, with a common forward primer and genotype-specific reverse primers with extensions for FAM or HEX. **(B)** Following KASP reactions, HEX and FAM fluorescence would be recorded and plotted. KO only template would here be expected to only give HEX fluorescence, while WT only template would give only FAM fluorescence, with datapoints plotted in red and blue respectively. A known 1:1 ratio of WT and KO template would give an intermediate result (purple datapoints). Following in planta growth, the WT/KO ratio can be determined as being more WT or more KO based on a shift in the relative fluorescence signal seen. In this example, the template following in planta growth (WT/KO ?:?) is skewed toward the WT genotype, relative to the 1:1 control (>1:1). This would suggest that the KO strain is less virulent than the wild-type. No template controls (NTC) are also illustrated.

Further, while infection assays carried out in this work, with the exception of assays for starch staining (**Figure 4.10** and **4.11**), used detached leaves on water agar plates, it may also be informative to validate these results using attached leaf assays.

4.4.2 CRBIC interacts with ChEC153 at chloroplastic puncta, but loss of *CRBIC* from the host does not alter susceptibility to infection by *C. higginsianum*

I identified CRBIC (*AT1G53120*) as a putative interactor of ChEC153 via IP-MS (**Table 4.1**), and while this interaction remains to be confirmed with direct protein-protein interaction assays, it is supported by the altered localisation of CRBIC seen in the presence of ChEC153 (**Figure 4.15**), shifting from diffuse to punctate. Alteration of subcellular localisation presents compelling evidence for an interaction between proteins. Additionally, I observed the recruitment of CRBIC to puncta in the presence of either MRC or TaCSP41a (**Figure 4.16**), suggesting that these proteins may form a protein complex. Despite this targeting of CRBIC by the putative fungal effector, loss of *CRBIC* does not appear to alter the level of Arabidopsis susceptibility to infection by *C. higginsianum* (**Figure 4.20**). Given the lack of a clear virulence phenotype I observed for *ChEC153*-knockout *C. higginsianum* strains (**Figure 4.9**), the finding that the *crbic* mutant line and Col-0 show equal susceptibility to infection is not wholly surprising. Given more time for this project, it may also be informative to use microprojectile bombardment to observe the localisation of ChEC153-eGFP in the *crbic* mutant Arabidopsis background to test whether the putative effector is able to form puncta in the absence of this host interactor.

CRBIC is annotated in TAIR simply as an RNA-binding S4 domain containing protein, and has some homology to the bacterial cell division protein YlmH (Miyagishima et al., 2005). In the cyanobacterium *Synechocystis*, a homolog of this protein, Sll1252, has been implicated in photosynthetic electron transport (Inoue-Kashino et al., 2011). Little has been done to evaluate the role of this protein in Arabidopsis, but it is possible that it is involved in chloroplast division or electron transport, and its predicted RNA-binding capacity remains to be proven experimentally.

4.4.3 ChEC153 may target CPNB3, while EIF2-A2 represents an unlikely interactor

In addition to CRBIC, I evaluated the subcellular localisations of candidate interactors of ChEC153 CPNB3 and EIF2-A2 in *N. benthamiana*. That ChEC153 and CPNB3 are present within the same subcellular compartment (**Figure 4.13**) suggests potential for an interaction between these proteins, though any direct interaction remains to be confirmed experimentally using an assay such as co-immunoprecipitation to validate the IP-MS result. Chaperonins are a type of molecular chaperones that aid in the correct folding of proteins. Chloroplast chaperonins are hetero-oligomeric, with CPN60 (chloroplast homolog of GroEL) comprising Cpn60α and Cpn60β subunits (Nishio et al., 1999). CPNB3 is a minor subunit constituent of the latter in Arabidopsis (Hill and

Hemmingsen, 2001), and functions in association with CPNA2 (Ke et al., 2017). In addition to determining correct folding of chloroplastic proteins whether translocated or newly synthesised, CPN60 may be required for the correct targeting of chloroplast encoded thylakoid membrane proteins (Klasek et al., 2020) and normal plastid division (Suzuki et al., 2009). It is possible that pathogens may target chaperonins in order to disrupt normal chloroplast function. However, as a component of a CPN60, it is possible that CPNB3 interacts with many chloroplastic proteins, and so may have been pulled down my IP-MS experiment due to a potential role in aiding the import/folding of ChEC153-eGFP in the chloroplast, rather than due to a specific, functional targeting of CPNB3 by the putative effector. CPN60 has been implicated in putative ribonucleoprotein complexes with mTERF9, which localises to chloroplastic puncta (Méteignier et al., 2021). CPN60 interacts with mTERF9, alongside ISE2 and a number of PAP proteins (Méteignier et al., 2021), and so may be important in the folding of some of these proteins. In *Pseudomonas syringae*, effector HRP OUTER PROTEIN I1 (HopI1) has been found to localise to the host chloroplast and interact with chaperone HEAT-SHOCK PROTEIN, resulting in altered thylakoid structure and SA accumulation (Jelenska et al., 2007, Jelenska et al., 2010). This exemplifies to the potential role of host chaperone proteins as effector targets, though the mechanisms by which such chaperone-targeting results in a virulence effect is not yet clear.

EIF2-A2 was also identified as a potential interactor of ChEC153 by IP-MS (**Table 4.1**). However, the spatial separation of ChEC153-eGFP and EIF2-A2-mCherry in different subcellular compartments as shown in **Figure 4.14** suggests that a considerable interaction between ChEC153 and EIF2-A2 is unlikely. Cytosolic translation-related proteins may represent a common source of contamination in chloroplast proteomic analyses (Bouchnak et al., 2019). It is possible that EIF2-A2 was pulled down in this IP-MS experiment as a contaminant, as an artefact of the *in planta* translation of *ChEC153* in the Arabidopsis stable line (which would not be present in the infection context).

4.4.4 Candidate interactors of ChEC153 include PEP complex proteins

While not further examined in this work, I identified AtFSD3 (*AT5G23310*) and AtMurE (*AT1G63680*) as potential host interactors of ChEC153 (**Table 4.1**). These proteins stand out as being components of the PEP complex, and are also known as PAP4 and PAP11 respectively. FSD3 has been suggested to regulate chloroplast development through its role in the PEP complex (Lee et al., 2019). The *Chlamydomonas reinhardtii* MurE homolog (*Cre12.g519900*) has been seen to localise to chloroplastic puncta in *Chlamydomonas*, wherein it is also identified as CPP2 (Wang et al., 2023c). However, localisation of AtMurE with a GFP tag in *Nicotiana tabacum* and *Physcomitrella patens* by Garcia and colleagues showed a diffuse localisation within the chloroplast (Garcia et al., 2008). Further, though falling below my threshold for significance in my IP-MS data,

the tenth most enriched protein in ChEC153-eGFP samples is FRUCTOKINASE-LIKE PROTEIN 2 (FLN2; *AT1G69200*), which is also implicated in the PEP complex despite not being defined as a PAP (Pfalz et al., 2006). FLN2 interacts with bona fide PAPs PAP10 (Arsova et al., 2010) and PAP6 (also known as FLN1; Huang et al., 2013). This may point to targeting of chloroplast gene transcription during infection by the action of putative effector ChEC153.

4.4.5 The *mrc* starch granule initiation mutant has enhanced resistance to infection

Here, I have shown that loss of MRC from Arabidopsis leaves confers an increased resistance to infection by *C. higginsianum* (Figure 4.12). To support the data presented here, it would be beneficial to assay for susceptibility comparable to the wild-type in pMRC::MRC complementation lines in the *mrc* background. It is possible that the increased resistance to infection is due to the altered starch granule morphology seen in *mrc* lines: where fewer, larger starch granules are seen. These granules may be less accessible for degradation and therefore reduce the accessibility of host carbon to the pathogen. In order to investigate this hypothesis, it would be informative to assay for resistance phenotypes of other starch granule initiation mutants. For example, overexpression and knockout of PTST2 in Arabidopsis results in the formation of smaller and larger starch granules, respectively (Seung et al., 2017). Carrying out infection assays with these lines may allow potential correlations between granule sizes and resistance phenotypes to be elucidated.

While none of the punctate starch granule initiation proteins mentioned in Section 1.7.2 were present in my IP-MS data (Table 4.1) as putative interactors of ChEC153 in Arabidopsis, it would be informative to directly test for interactions between ChEC153 and MRC due to the co-localisation between these two proteins (Figure 3.8). However, due to difficulties in detecting the full-length ChEC153-eGFP in *N. benthamiana* protein extracts, I was not able to complete co-immunoprecipitation experiments within this project. It appeared that the eGFP tag was being cleaved from the ChEC153 protein in this system (Figure 4.21). While I therefore cloned constructs for the expression of *ChEC153* with a C-terminal 6× HA tag to potentially avoid this problem, time limitations precluded the completion of this experiment. Given the interactions between MRC and SS4 and SS5 (Vandromme et al., 2019, Abt et al., 2020), these also represent candidates to probe for interactions with ChEC153. As the relationship between starch granule initiation, plastidial transcription and nucleoids is not clear, these data may suggest some overlap between the location/proteins involved.

4.4.6 Chapter conclusions

In summary, while the virulence function of ChEC153 is not clear, it appears to alter Arabidopsis susceptibility to infection and to interact with a number of host proteins. Candidate host targets include several PEP complex associated proteins. Notably, ChEC153 candidate interactor CRBIC, which may also be linked to plastidial transcription, can be recruited to puncta within the chloroplast by ChEC153, MRC, and TaCSP41a. The potential link between starch granule initiation sites and sites of plastid transcription and RNA-processing remains to be clarified.

Chapter 5

Elucidating the *ChEC153* gene model and exploratory transcriptome analysis

5.1 INTRODUCTION

5.1.1 *C. higginsianum* genetic resources

In 2012, the first genome of *C. higginsianum* strain IMI 349063 was assembled using short-read sequencing (O'Connell et al., 2012). Building on this, Zampounis *et al.* published an improved genome assembly of IMI 349063 using long-read sequencing and optical mapping (Zampounis et al., 2016). This more complete genome has elucidated a link between transposable elements and predicted effector genes as well as secondary metabolite gene clusters (Dallery et al., 2017). More recently, Tsushima and colleagues published a genome assembly of *C. higginsianum* strain MAFF 305635-RFP (Tsushima et al., 2019a). The nuclear genome of *C. higginsianum* comprises approximately 50 Mb across 12 chromosomes, two of which are minichromosomes of a considerably smaller size than the other ten (O'Connell et al., 2012). Strains IMI 349063 and MAFF 305635-RFP are predicted to have 14,651 and 12,915 protein-coding genes, respectively (Zampounis et al., 2016, Tsushima et al., 2019a). Alongside other filamentous phytopathogenic genomes (Raffaele et al., 2010), the genome of *C. higginsianum* has been described as a bipartite, "two-speed" genome. The genome is compartmentalised into gene-sparse and gene-dense regions, with effector genes and transposable elements enriched in the gene-sparse regions. These regions may aid in genomic plasticity, by facilitating genetic flux via transposable elements, an especially valuable mechanism in a species lacking sexual reproduction (Tsushima et al., 2019a).

In addition to these genomic resources, RNA-Seq data for transcriptomic analyses of *C. higginsianum* were published by O'Connell and colleagues (O'Connell et al., 2012). These data comprise samples of *C. higginsianum* cultured *in vitro*, and at three stages of infection by *C. higginsianum* of Arabidopsis including: formation of *in planta* appressoria, the biotrophic phase of infection, and the necrotrophic phase of infection (O'Connell et al., 2012).

5.1.1.1 Genome annotations can present bottlenecks for the study of non-model organisms

Over the last decades, the genome sequencing revolution – where sequencing costs have drastically dropped thanks to technological advances – has allowed the sequencing and assembly of exponential number of genomes. However, the usefulness of genome sequences of any species can be limited by the accuracy of its annotations. Particularly in eukaryotes, where genes are often sparsely arranged and coding sequences interrupted by the presence of intronic regions, the accuracy of *in silico* gene annotations can be suboptimal. When genes are well-characterised, and annotated in other species, they may be recognised with higher accuracy by genome annotation

tools than novel or uncharacterised genes. However, reliance on previous annotations of the same or other species comes with its own caveats, with limitations and errors from draft genome assemblies/annotations being carried forward between genomes, leading to the inadvertent propagation of errors (Salzberg, 2019). The correct annotation of genes is also reliant on the quality of genome assemblies – should portions of a gene be separated in the assembly, annotation tools may be unable to recognise them as a single gene. Application of long-read sequencing may help to produce a less fragmented genome assembly and thereby aid gene annotation (Amarasinghe et al., 2020).

5.1.1.2 *In silico* gene annotation pipelines and the annotation of *C. higginsianum* genome assemblies

Computational approaches can be applied to predict genes, and these are generally divided into those approaches using sequence similarity (to other genomes, expressed sequence tags (ESTs), or proteins), and those based on identifying genes from gene structure and signals (e.g. start and stop codons and other motifs, as well as exon codon usage). This latter approach is known as *ab initio* gene annotation.

Annotation pipelines such as MAKER (Cantarel et al., 2008) seek to avoid some of the pitfalls associated with gene annotation predictions by incorporating RNA-Seq data to improve annotations. These tools can also signpost annotations that require manual review. In *C. higginsianum*, the IMI 349063 genome assembly ASM167251v1 published by Zampounis *et al.* was generated using PacBio long-read genome sequencing and improved upon the more fragmented IMI 349063 genome assembly ASM31379v2 previously published by O'Connell and colleagues using short-read sequencing (Zampounis et al., 2016, O'Connell et al., 2012). This resulted in a reduction of the number of annotated protein-coding genes from 16,172 to 14,651, possibly due to the short-read O'Connell *et al.* assembly splitting genes between contigs and resulting in multiple annotations for a single gene. For the original assembly, ASM31379v2, O'Connell and colleagues applied a combination of gene annotation pipelines trained with ESTs (a type of cDNA fragment) from multiple *C. higginsianum* libraries and secretome analysis (O'Connell et al., 2012). Zampounis and colleagues applied the MAKER2 pipeline (Holt and Yandell, 2011) for gene annotation in their *C. higginsianum* genome assembly ASM167251v1 (Zampounis et al., 2016), while Tsushima *et al.* used an analogous pipeline, BRAKER1 (Hoff et al., 2016), in their genome annotation efforts (Tsushima et al., 2019a). The authors report using the O'Connell *et al.* RNA-Seq data to support the *ab initio* gene annotations from the gene prediction pipeline (O'Connell et al., 2012, Tsushima et al., 2019a). While the abovementioned annotation pipelines use RNA-Seq data to aid predictions, other pipelines, such as BRAKER2 (Bruna et al., 2021), use protein databases to

inform gene model annotations. Further pipelines, such as BRAKER3 and FINDER, seek to integrate both approaches (Gabriel et al., 2023, Banerjee et al., 2021). However, despite improving pipelines, it is still likely that a portion of the genes in a genome will be annotated incorrectly, and gene-by-gene manual validation of annotations is not viable.

Pacific Biosciences and Oxford Nanopore Technologies offer long-read RNA-Sequencing using cDNA synthesised from the extracted RNA. While these approaches are currently lower throughput and more error-prone than the Illumina platform short-read RNA sequencing applied in the majority of RNA-Seq experiments, long-read RNA-Sequencing can aid genome assembly and gene annotation (Stark et al., 2019). Using this approach, the entire RNA transcript for a gene can be sequenced in a single read, thus removing any ambiguity regarding intron/exon junctions and splice variants (Hu et al., 2021). Further, direct long-read sequencing of native RNAs is now possible with the application of nanopore technologies (Workman et al., 2019, Wongsurawat et al., 2019, Wongsurawat et al., 2022). As well as being able to inform on the structure of genes for which RNA is sequenced, this approach also serves to confer information regarding post-transcriptional modification of RNA molecules, which may be lost where RNA is first converted to cDNA for sequencing.

5.1.1.3 Gene model validation and 5' rapid amplification of cDNA ends

To complement gene model predictions, gene models can be validated experimentally. This is particularly important where candidate genes lack well-characterised homologs in related species/strains. If sufficient knowledge of the sequence of the region of interest is available, it may be possible to amplify the coding sequence (CDS) from cDNA for cloning and sequencing. This will confirm any internal intron/exon boundaries. However, this approach relies on sufficient knowledge of the 5' and 3' regions of the gene to design cloning primers.

Rapid amplification of cDNA ends (RACE) is a technique that can be used to determine the 3' or 5' terminus of a cDNA molecule, thereby conferring information about the gene coding sequence. The schematic presented in **Figure 5.1** depicts the process of 5' RACE, in which knowledge of the 3' region of a CDS of interest allows the design of gene-specific primers to inform on the 5' region. First, cDNA is synthesised by reverse transcription of isolated RNA using a gene-specific primer (GSP1) that binds at the known 3' of the gene. The resulting cDNA library is then treated with RNase to remove the mRNA template, and purified to separate it from the primer, unincorporated dNTPs, and proteins from the cDNA synthesis reaction. The purified cDNA is then treated with a terminal transferase to add a 3' homopolymeric tail. In the example presented in **Figure 5.1**, the cDNA is polyadenylated. This provides a known 3' region to which a complementary homopolymeric PCR primer will bind. The primer used for this initial PCR contains the homopolymeric stretch of bases

for annealing to the tailed cDNA as well as an anchor sequence – for poly(A) cDNA, this will be an oligo(dT)-anchor primer as shown in Figure 5.1. This oligo(dT) anchor is used in combination with a second gene-specific primer (GSP2), which binds in a known region of the cDNA, to generate a PCR product spanning the length of the poly(A) cDNA. A second round of PCR may be desired (PCR2) during which an anchor PCR primer binds the anchor region incorporated by the oligo(dT)-anchor primer during PCR1, and a nested, third gene-specific primer (GSP3) binds the known 3' region of the cDNA conferring additional specificity. Products of PCR2 can then be analysed by agarose gel electrophoresis, cloned, and sequencing reactions performed to determine the sequence of the unknown 5' region of the cDNA.

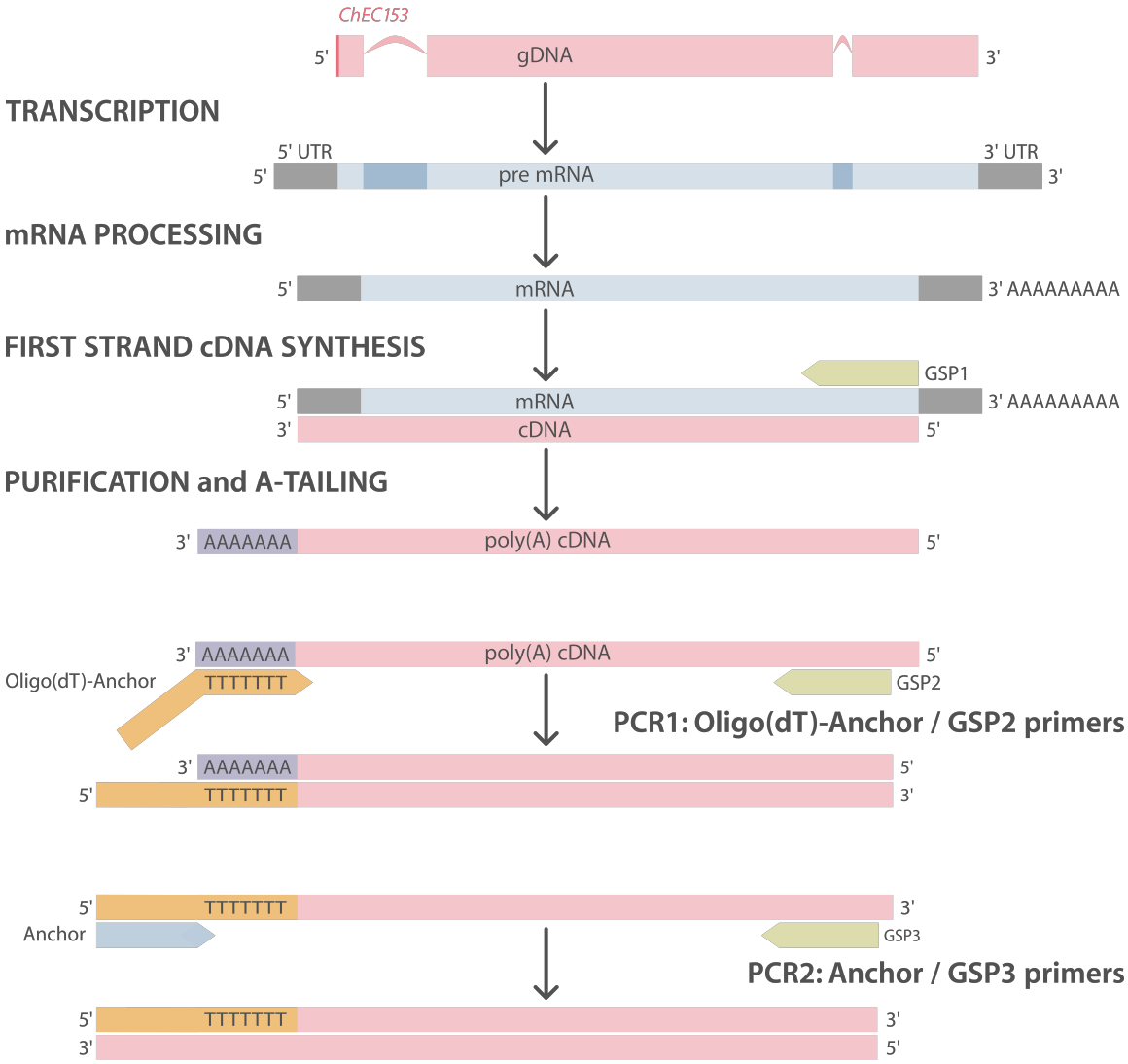


Figure 5.1: Schematic to illustrate the process of 5' RACE. For illustrative purposes, the *CH63R_12252* gene model of *ChEC153* is shown as the gene (gDNA) for which the 5' region is to be determined. Transcription and mRNA processing produce an mRNA molecule devoid of introns. This RNA is then used as a template for first strand cDNA synthesis using a gene-specific primer (GSP1), or oligo(dT) primer, with synthesis in the 5' to 3' direction. The RNA is then degraded, and cDNA purified. Homopolymeric tailing – in this case polyadenylation – is carried out to produce purified, poly(A)-tailed cDNA. The first PCR reaction (PCR1) uses a forward oligo(dT)-anchor primer and a second gene-specific reverse primer (GSP2). Using the product of PCR1 as a template, PCR2 uses a forward anchor-specific primer and a third, nested gene-specific primer (GSP3) to amplify the cDNA further. This product can then be cloned and sequenced to elucidate the 5' region of the gene of interest.

5.1.2 Dual transcriptome analysis of pathosystems

In order to study the process of infection, many approaches use transcriptomics to analyse both host and pathogen responses during various stages of infection. Due to the dynamic nature of infection in both host and pathogen, collection and analysis of combined RNA can provide a powerful and cost-effective tool in understanding the transcriptional response of a complete pathosystem. For this purpose, dual RNA-Seq (Westermann et al., 2012) can be applied. While analysis of such datasets can be complex, where reference genomes are available for both organisms, analysis approaches have been described (O'Keeffe and Jones, 2019).

Teixeira and colleagues applied dual RNA-Seq to the *Moniliophthora perniciosa-Theobroma cacao* pathosystem (Teixeira et al., 2014). *M. perniciosa* is a fungal hemibiotrophic pathogen that demonstrates a remarkably long biotrophic phase – lasting two to three months, considerably longer than the short biotrophic phases seen in the infection cycles of most hemibiotrophs, such as *M. oryzae* and *Colletotrichum* spp. including *C. higginsianum*. During this prolonged biotrophic phase, the authors observed host metabolic reprogramming in response to infection, including the downregulation of a number of photosynthesis-related genes (as is frequently seen during infection, see Section 1.6.1). Among the host genes with altered expression during infection, a number of lipases are strongly upregulated during infection as well as several amino acid metabolic and catabolic genes (Teixeira et al., 2014). Further, genes encoding starch-degrading α-amylases and a number of hexose transporters were upregulated, and enzymes related to starch biosynthesis (starch synthase and AGPase genes) were downregulated. Strikingly, in this pathosystem, during infection the number of starch granules in shoots of plants are visibly depleted, supporting the transcriptional reprogramming observed (Teixeira et al., 2014).

In the *C. higginsianum*-Arabidopsis pathosystem, transcriptional analyses have previously been largely focussed on analysis of pathogen transcripts (Takahara et al., 2009, O'Connell et al., 2012, Kleemann et al., 2012) or host responses (Gebauer et al., 2017, Tang et al., 2023) separately. Recently, Zhu and colleagues published a dual-transcriptome analysis of Arabidopsis infected by *C. higginsianum* strains with and without the presence of a pathogenicity gene *ChATG8* (*AUTOPHAGY RELATED GENE 8*; Zhu et al., 2023) using a dual RNA-Seq approach. The authors identify host genes which may be important in the host defence response, as well as fungal genes implicating ChATG8 in autophagy and melanin biosynthesis (Zhu et al., 2023).

5.1.3 Chapter aims

In this chapter, I aimed to validate the gene model for *ChEC153* in light of differences in gene annotations in the different genome assemblies available for *C. higginsianum*. Further, I aimed to carry out dual-transcriptome analyses of Arabidopsis infected with *C. higginsianum* strains *ΔChKu80* and *ΔChEC153-2* to evaluate any differences in gene expression during infection in either the host or pathogen in the presence/absence of ChEC153 that might highlight the function of the putative effector.

5.2 MATERIALS & METHODS

5.2.1 Reverse transcription PCR (RT-PCR) for gene model validation

5.2.1.1 Initial RT-PCR

cDNA samples of wild-type *C. higginsianum* strain IMI 349061 during infection were generated from drop-inoculated Arabidopsis leaves. Mature leaves of five-week-old Col-0 Arabidopsis plants were detached, their petioles inserted into water agar plates (2% agar), and the adaxial leaf surfaces drop-inoculated with approximately twenty 3 µL droplets of a *C. higginsianum* conidial suspension containing 2 ×10^6 conidia/mL. Plates were incubated at 25 °C (10 h:14 h, light: dark) and leaves harvested by flash-freezing in liquid nitrogen at 19-, 24-, 40-, and 46-hours post inoculation. Samples were ground to a fine powder with 5 mm metal bearings using a Geno/Grinder® (SPEX Sample Prep).

RNA was extracted from approximately 100 mg of ground infected leaf tissue using the RNeasy mini kit (QIAGEN, 74104) according to the manufacturer's instructions. DNase treatment was carried out using a TURBO DNA-*free*™ kit (Invitrogen, AM1907) according to the manufacturer's "rigorous" protocol. cDNA was synthesised using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4374966) with random primers for standard RT-PCR experiments, and 0.5 µL of cDNA in Phusion® High-Fidelity DNA polymerase amplification reactions (see Section 2.5.1). Independently, cDNA was also synthesised from RNA (40-hpi sample) using a targeted, gene-specific primer (GSP1) using the same protocols, for use as a template for 5' RACE.

5.2.1.2 Additional RT-PCR

For additional RT-PCRs, RNA was extracted from additional samples harvested from the same round of inoculations, by Dr Andrew Breakspear, as described above in Section 5.2.1.1. First-strand cDNA synthesis was performed using SuperScript™ IV Reverse Transcriptase (Invitrogen, 18090010) according to the manufacturer's instructions. One microgram of template RNA was added to each reaction and primed using oligo$(dT)_{12\text{-}18}$ primers (Invitrogen, 18418012). 2 µL volumes were then used in PCR with GoTaq® G2 Green Master Mix (Promega, M7822) using the manufacturer's recommended protocol, with 38 cycles and allowing the following extension times based on the expected amplicon sizes: Primers 100/101: 30 seconds; 097/101: 1 minute; 097/099: 2 minutes. PCR products were run on a 1% agarose gel and bands visualised with EtBr and a UV transilluminator.

5.2.2 5' RACE

5' RACE was carried out using a 5'/3' RACE kit, 2nd generation (Roche, 3353621001), according to the manufacturer's instructions. Briefly, cDNA synthesised using a gene-specific primer (GSP1) was purified using the QIAquick PCR Purification kit (QIAGEN, 28104) with adjustments to the QIAGEN protocol in accordance with the RACE kit protocol (100 µL of binding buffer instead of 500 µL, and increased centrifugation speeds). Purified cDNA was polyadenylated by incubation with dATP and a terminal transferase for 25 minutes at 37 °C. Initial PCR reactions (PCR1) were carried out using Phusion polymerase to amplify from 5 µL of the poly(A)-tailed cDNA in a final reaction volume of 50 µL, with primers GSP2 and the provided oligo(dT)-anchor primer. Secondary PCR reactions (PCR2) were then carried out using the product of PCR1 as a template, amplifying with primers GSP3 and a PCR anchor primer. To enable Golden Gate cloning of PCR products, a custom anchor primer was used with the same sequence as that provided by the kit but with the addition of a 5' extension for cloning to a Level 0 vector with BpiI. GSP3 contains a similarly designed cloning extension. Following amplification, PCR products were run on a 1.5% agarose gel for separation alongside 100 bp and 1 kb DNA ladders (NEB, N3231 and N3232). Control reactions were carried out as recommended by the kit manufacturers, using the provided control RNA in combination with the RNA extracted from the *C. higginsianum*-infected Arabidopsis sample (Section 5.2.1), and the provided control primers.

5.2.3 RNA-Sequencing

Samples for RNA-Sequencing (RNA-Seq) and analysis were generated, validated, and submitted for sequencing by Dr Andrew Breakspear, pre-processing of RNA-Seq data was carried out by Dr

Hannah Rae Thomas, and I analysed the pre-processed RNA-Seq data with the guidance of Dr Hannah Rae Thomas.

5.2.3.1 RNA-Seq sample preparation

The following methods for RNA-Seq sample preparation and quality control checks were both provided by and carried out by Dr Andrew Breakspear. To generate samples for RNA extraction, mature detached, five-week-old Arabidopsis Col-0 leaves were drop-inoculated with *C. higginsianum* strains *ΔChKu80* (CY6021, generated by Korn and colleagues from strain MAFF 305635 (Korn et al., 2015)) or *ΔChEC153-2* (generated in this work from *ΔChKu80*, see Chapter 4). Further, leaves were inoculated with water as a mock control. Three biological replicates were performed, with each sample consisting of six leaves (two per plant) on 2% water agar plates. Each leaf was inoculated with six 5 µL droplets of a 2×10^6 conidia/mL suspension. Plates were sealed with parafilm and incubated at 25 °C in short-day conditions. Samples were harvested at 24- and 36-hours post inoculation (hpi) and flash-frozen in liquid nitrogen. Samples were stored at -70 °C before being crushed in liquid nitrogen using a pre-cooled pestle and mortar. For each sample, 100 mg of ground tissue was subject to RNA extraction with a QIAGEN RNeasy Plant Mini kit (QIAGEN, 74904). Contaminating genomic DNA was removed using a TURBO DNA-*free*™ kit (Invitrogen, AM1907) according to manufacturer's "rigorous" protocol. Samples were sent to GENEWIZ for library construction and sequencing. RNA concentration and quality was assessed using Qubit and Bioanalyzer (RIN >6). rRNA was depleted by poly-A selection and libraries built using the NEB Next Ultra II RNA Library Prep Kit (NEB) and sequenced on an Illumina NovaSeq (Illumina) with condition of 2×150 bp at 20 million read depth.

5.2.3.2 RNA sample validation

5.2.3.2.1 Reverse transcription quantitative-PCR (RT-qPCR)

First-strand cDNA synthesis was performed using the High-Capacity cDNA Reverse Transcription kit with RNase inhibitor (Applied Biosystems, 4374966) using random primers according to the manufacturer's instructions, and subsequently diluted 1 in 10 with water. qPCR was carried out using a LightCycler® 480 (Roche) with LightCycler® 480 SYBR Green I Master Mix (Roche, 04707516001) according to the manufacturer's instructions, with each sample measured in technical triplicates.

5.2.3.2.2 Reverse-transcription PCR (RT-PCR)

First-strand cDNA synthesis was performed as described above for qPCR. Phusion® High-Fidelity DNA polymerase (NEB, M0530) was used to amplify a 293 bp region of *CH63R_13188* with primers 102 and 103. Amplification reactions were assembled in 20 µL volumes and according to the manufacturer's instructions, using 2 µL of cDNA per reaction. A standard thermocycling protocol was used: initial 30 second denaturation at 98 °C; 35 cycles of denaturation (98 °C, 10 seconds), annealing (65 °C, 20 seconds), and extension (72 °C, 15 seconds); and final extension at 72 °C for 10 minutes. PCR products were separated by electrophoresis on a 1.5 % agarose gel stained with EtBr, and bands visualised with a UV transilluminator.

5.2.3.3 RNA-Seq and data pre-processing

The following methodologies for RNA-Sequencing were provided and carried out by Dr Hannah Rae Thomas. For dual host-pathogen RNA-Seq, fasta files were quality checked using FastQC version 0.11.8 (Andrews, 2010). Low quality reads (phred < 20) and adaptors were trimmed using Trim Galore! version 0.4.2 and Cutadapt version 1.9.1 (Martin, 2011, Krueger, 2015). Sequences were first mapped to the TAIR10 (Lamesch et al., 2012) *Arabidopsis thaliana* genome using HISAT2 version 2.1.0 (Kim et al., 2019). Unmapped reads were then aligned to the *Colletotrichum higginsianum* genome (accession: GCA_004920355.1) using NextGenMap 0.5.5 (Sedlazeck et al., 2013). The resulting .sam files were sorted and used to generate .bai and .bam files using SAMtools version 1.4.1 (Danecek et al., 2021). Transcripts were assembled using StringTie version 1.3.3 (Pertea et al., 2015) and raw read counts generated with the prepDE.py script: (http://ccb.jhu.edu/software/stringtie/dl/prepDE.py). Transcripts were visualised with the Integrative Genomics Viewer (IGV; Thorvaldsdóttir et al., 2013). For comparative analysis of previously published RNA-Seq data, fasta files (SRA: SRP009324; O'Connell et al., 2012) were downloaded as pre-processed as previously described. Genome indexes were built for the Zampounis *et al.* genome (C. *higginsianum* strain IMI 349063, genome assembly ASM167251v1, assembly accession: GCA_001672515.1; Zampounis et al., 2016), the O'Connell *et al.* genome (C. *higginsianum* strain IMI 349063, genome assembly ASM31379v2, assembly accession: GCA_000313795.2; O'Connell et al., 2012), and the Tsushima *et al*. genome (C. *higginsianum* strain MAFF 305635-RFP, genome assembly ASM492035v1, assembly accession: GCA_004920355.1; Tsushima et al., 2019a) using HISAT2. Alignments were preformed against the specific genome of interest and post-processed as previously described.

5.2.3.4 Data analysis

Raw read counts were utilised for differential analysis in R (R Core Team, 2021) using DESeq2 with a TidyVerse pipeline (Love et al., 2014, Wickham et al., 2019). For fungal reads, genes with 0 counts were filtered and 1 pseudo-count was added to all remaining samples to overcome any statistical issues with low read counts. For Arabidopsis reads, any gene with no reads in 75% of the samples was filtered as well. The DESeq2 model incorporated the fungal genotype and time as variables. PCA plots were generated using ggplot2 (Wickham, 2009) and ggConvexHull (Martin, 2017). Normalised read counts and model dispersion were determined in DESeq2. Differential expression was determined based on Wald Test with an adjusted p-value cut-off of < 0.10 and $\log_2$ fold change > 0.5 or < -0.5.

5.3 RESULTS

5.3.1 The *ChEC153* gene model differs between published genome assemblies

The sequence of *ChEC153* used in **Chapter 3** and **Chapter 4** of my thesis is the mature, protein-coding sequence of *CH63R_12252*, as annotated in what I will henceforth refer to as the Zampounis *et al.* genome (C. *higginsianum* strain IMI 349063, genome assembly ASM167251v1v; Zampounis et al., 2016). This gene is equivalent and identical to *CH063_01906* in the earlier genome assembly for IMI 349063, ASM31379v2 (hereafter: O'Connell *et al.* genome; O'Connell et al., 2012). This coding sequence was synthesised as part of the *Colletotrichum higginsianum* putative effector screen carried out by Dr Mina Ohtsu and Dr Joanna Jennings in the Faulkner lab (Ohtsu et al., 2023, Jennings, 2021). For clarity, this version of *ChEC153*, described by the *CH63R_12252* gene model, is referred to as *ChEC153.1* (encoding protein ChEC153.1) for the remainder of this thesis.

During my project, I noticed that the gene model at the corresponding locus in the MAFF 305635-RFP genome assembly ASM492035v1 (hereafter referred to as the Tsushima *et al*. genome; Tsushima et al., 2019a), *CH35J_009900*, was inconsistent with the *CH63R_12252* Zampounis *et al*. genome annotated gene model at this location. The abovementioned gene IDs, genome assemblies and references are summarised in **Figure 5.2 A** for clarity. The published genomes are relatively similar in this region, with only 28 single nucleotide polymorphisms (SNPs) being seen between the Zampounis *et al.* and Tsushima *et al.* genomes when comparing the 1809 bp regions of the genomic sequences in which the Zampounis *et al*. gene annotation predicts *CH63R_12252* (chromosome 8: 4,300,557 - 4,298,749; associated region in the Tsushima *et al*. genome: contig 7:

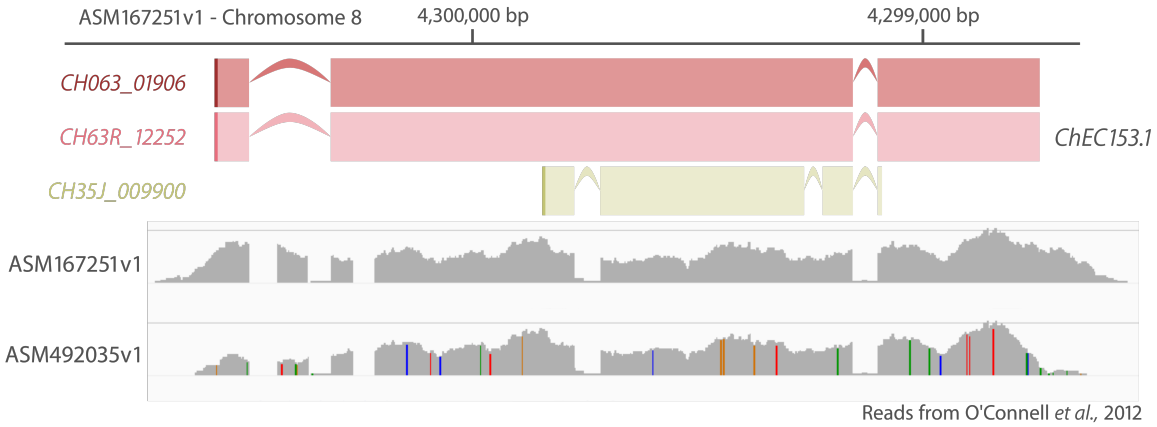
3,328,310 – 3,326,502, genes encoded on the reverse strand). Due to this sequence homology, it is unlikely that the gene encoded by this region differs between the two strains, and that discrepancies in the gene models are likely due to differences in the annotation prediction pipelines rather than evolutionary variation in the *ChEC153* gene.

Specifically, the start codon of the Tsushima *et al.* gene model begins 713 bp downstream of the Zampounis *et al.* gene model start codon. If true, this would result in a truncation at the 5' of *CH35J_009900* (Tsushima *et al.* gene model) relative to *CH63R_12252* (Zampounis *et al.* gene model). Further, the Tsushima *et al.* predicted gene model annotates two introns within the second exon of the Zampounis *et al.* gene model, the second of which would result in an early stop codon 341 bp upstream of the stop codon annotated in the Zampounis *et al.* gene model. The O'Connell *et al.*, Zampounis *et al.*, and Tsushima *et al.* gene models are depicted in Figure 5.2 B. Due to these gene model discrepancies, the proteins predicted by these two gene models differ significantly in their length, with the Tsushima *et al.* gene model encoding a protein consisting of 181 amino acids, much shorter than the 521 amino acid protein encoded by the Zampounis *et al.* gene (Figure 5.2 C). As the two proteins are mostly in the same reading-frame, the protein sequences encoded by the regions of the genomes predicted to be exons in both gene models are largely the same. However, the large N-terminal and smaller C-terminal truncations predicted by the Tsushima *et al.* gene model have the potential to alter the subcellular localisation of the protein, any postulated function in infection, and the prediction of a signal peptide for secretion. Due to the differences between these gene models, it therefore became critical for me to resolve the true gene model of *ChEC153*, establishing the true coding sequence for this gene.

A

Strain	Reference	Gene ID	Genome assembly	Assembly accession
IMI 349063	O'Connell *et al.*, 2012	*CH063_01906*	ASM31379v2	GCA_000313795.2
IMI 349063	Zampounis *et al.*, 2016	*CH63R_12252*	ASM167251v1	GCA_001672515.1
MAFF 305635-RFP	Tsushima *et al.*, 2019	*CH35J_009900*	ASM492035v1	GCA_004920355.1

B



C

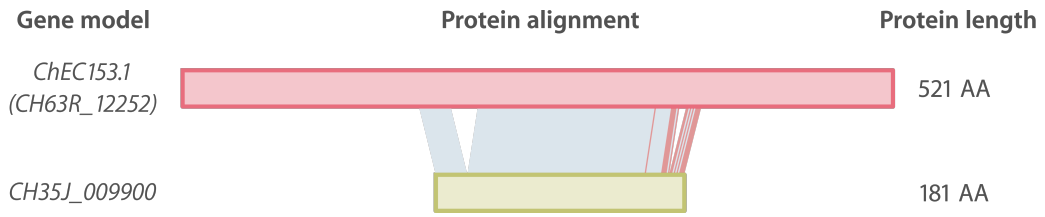


Figure 5.2: **Differences in gene models corresponding to *ChEC153* in different genome assemblies, and their encoded proteins.** **(A)** Summary of the three genome assemblies for *C. higginsianum*, with the gene IDs corresponding to *ChEC153* indicated. **(B)** Gene models at the locus for *ChEC153* in each genome assembly. The translation start site is indicated by a bold vertical line at the left-hand side, with predicted exons shown as rectangles and introns as arcs. The location in the Zampounis *et al.* genome assembly (ASM167251v1) is shown above the gene models. Below the gene models, coverage traces of RNA-Seq reads from a single biotrophic phase sample from the O'Connell *et al.* data are shown, mapped to the Zampounis *et al.* ASM167251v1 and Tsushima *et al.* ASM492035v1 genomes respectively. Coloured lines indicate SNPs relative to the reference. **(C)** Proteins encoded by the Zampounis *et al.* (pink) and Tsushima *et al.* (green) gene models. Regions in which the amino acid sequence is identical are shown in blue, with differences shown in red. The protein encoded by the O'Connell *et al.* gene model is identical to that of the Zampounis *et al.* gene model, and therefore for simplicity is not shown.

5.3.2 Elucidating the true gene model of the candidate effector gene *ChEC153*

5.3.2.1 RT-PCR analyses of *C. higginsianum* cDNA

To elucidate the true gene model of *ChEC153*, I infected Arabidopsis leaves with *C. higginsianum* and extracted RNA from samples harvested at several time-points early in the infection process. Based on the transcriptomic data from O'Connell *et al*., *ChEC153* is expressed only during *in planta* appressorial development and the biotrophic phase of infection (O'Connell et al., 2012). I therefore harvested samples at 19-, 24-, and 40-hours post inoculation to ensure the relevant infection stages would be sampled. From this RNA, I synthesised cDNA from which to amplify, and potentially clone and sequence, regions of the *ChEC153.1* gene. I designed primers 102 and 103 to amplify *ChEC6* (*CH63R_13188*), the most highly expressed putative effector gene expressed only in the same infection stages as *ChEC153* according to the O'Connell *et al.* transcriptomic data (O'Connell et al., 2012). This served as a positive control for the infection stage and fungal cDNA quality. As shown in **Figure 5.3**, I was able to amplify *ChEC6* at all time-points tested, indicating that I had isolated *C. higginsianum* RNA from the desired stage of infection. I also sought to amplify a small region of the putative *ChEC153* gene which is shared between Zampounis *et al.* and Tsushima *et al.* gene models to check for the expression of the putative effector gene irrespective of its gene model. I therefore used primers 100 and 101 to amplify a 426 bp region of the coding sequence, illustrated in **Figure 5.3 A**, from the cDNA samples. I observed amplicons of the expected size for each of the samples tested, indicating expression of the gene at each time-point, as shown in **Figure 5.3 B**.

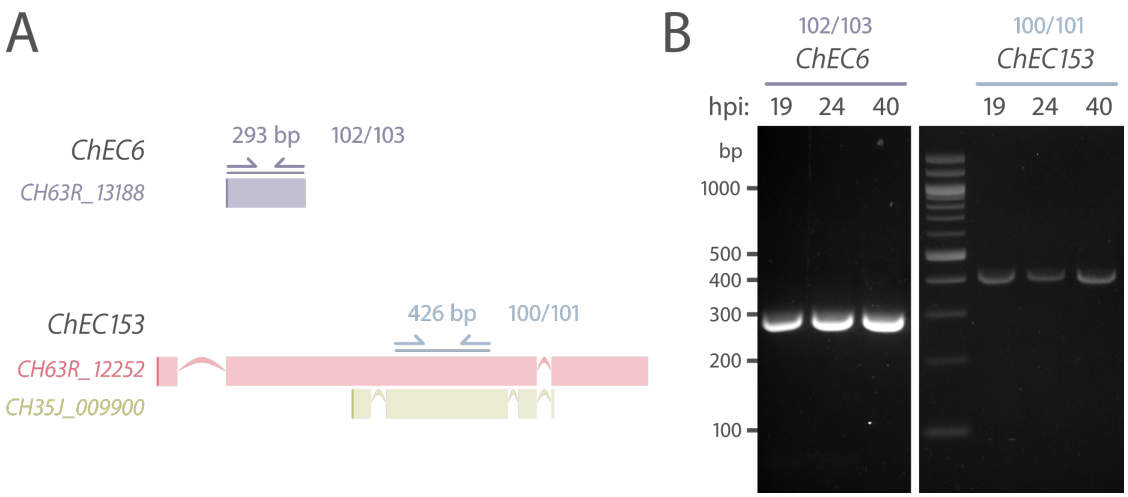


Figure 5.3: **RT-PCR amplification of *ChEC6* and a portion of *ChEC153*.** Amplification from cDNA generated from Arabidopsis leaf tissue infected with wild-type *C. higginsianum* strain IMI 349061, confirming expression of the putative effector ChEC153 at the timepoints sampled. **(A)** A schematic depicting the PCR primer binding sites and expected amplicon sizes from amplification with primer pairs 102/103 and 100/101. **(B)** PCR products were separated on a 2% agarose gel stained with EtBr, alongside a 100 bp DNA ladder (NEB). The PCR template used for each reaction was 0.5 µL of cDNA from samples harvested at 19-, 24-, or 40-hours post inoculation (hpi), as indicated above each lane of the gel.

Next, I sought to amplify cDNA of other regions of the putative effector gene to determine whether they were present, using the sample from tissue collected at 40 hpi as the PCR template. Alongside this, I used gDNA isolated from the fungus when cultured *ex planta* in liquid media as a template for some reactions, to serve as a positive control for the amplification reactions and to allow me to evaluate any size differences of amplicons produced, indicating potential introns. As expected based on the previous result (**Figure 5.3**), the 426 bp portion of the *ChEC153* sequence shared between gene models could be amplified from both cDNA and gDNA using primers 100/101 (**Figure 5.4**). My attempts to amplify the entire coding sequence of *ChEC153.1* based on the Zampounis *et al.* gene model from cDNA were not successful, though this region could be amplified from fungal gDNA with the same primers (**Figure 5.4**, primers 097/099). Primers to amplify a large portion of the Zampounis *et al.* gene model at the 5' of the gene from cDNA also yielded no PCR product (**Figure 5.4**, 097/101), but primers to amplify the 3' region gave a band of approximately the expected size (**Figure 5.4**, 100/099). This confirmed that this 3' region (chromosome 8: 4,299,682 - 4,298,752, Zampounis *et al.* genome) is present within the cDNA. On the other hand, the presence of the 5' region (chromosome 8: 4,300,557 - 4,299,683, Zampounis *et al.* genome) remained to be evidenced.

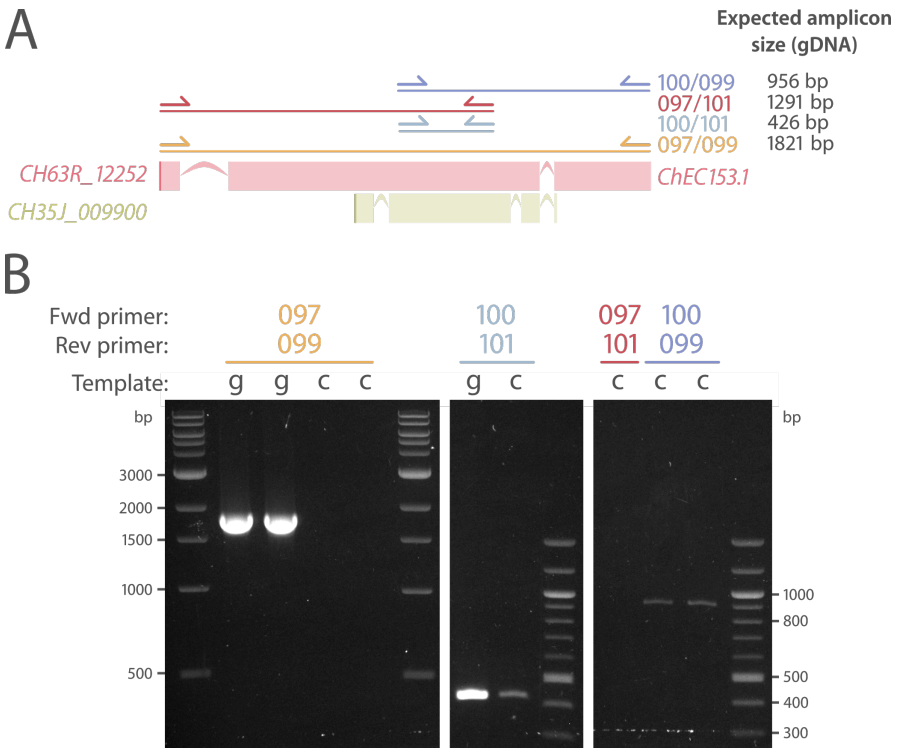


Figure 5.4: **RT-PCR amplification of different portions of *ChEC153.1*. (A)** A schematic depicting the PCR primer binding sites and expected amplicon sizes in base pairs (calculated based on amplification from gDNA). **(B)** PCR products following amplification with different primer pairs, as indicated above the gel, were separated on a 2% agarose gel stained with EtBr, alongside a 1 kb or 100 bp DNA ladder (NEB). The PCR template used for each reaction was 0.5 µL of: c = cDNA from samples harvested at 40-hours post inoculation (hpi), g = gDNA from *ex planta* fungal cultures, as indicated above each lane.

5.3.2.2 Sequencing of cloned regions of *ChEC153* from cDNA

The Zampounis *et al.* and Tsushima *et al.* gene models disagree on the presence of introns at 4,299,777 - 4,299,717 and 4,299,536 - 4,299,448 (chromosome 8, Zampounis *et al.* genome), as shown in **Figure 5.2**. To resolve the presence or absence of these introns, I cloned amplicons for these regions from both cDNA and gDNA into vectors for Sanger sequencing. Aligning the sequencing traces to the Zampounis *et al.* genomic sequence as a reference and comparing the sequencing reads obtained from cloned cDNA and gDNA, I was able to determine the presence of an intron at chromosome 8: 4,299,777 - 4,299,717 not predicted in the Zampounis *et al.* model. My sequencing data confirm the presence of this region in the sequence amplified from gDNA but not cDNA. This intron is 60 bp in length, and therefore would not result in a shift in reading-frame (**Figure 5.5 A**). On the other hand, I confirmed the absence of an intron at chromosome 8: 4,299,536 - 4,299,448 (Zampounis *et al.* genome), as this region was present in the region sequenced from both amplified gDNA and cDNA (**Figure 5.5 B**). Thus, I confirmed that neither of the published gene models correspond to the true coding sequence present in *C. higginsianum*

strain IMI 349061. Notably, with no 88 bp intron at chromosome 8: 4,299,536 - 4,299,448, the stop codon seen in the Tsushima *et al.* gene model (at chromosome 8: 4,299,095 - 4,299,093 of the Zampounis *et al.* genome) would not be in frame, and transcription would continue downstream. Without this frameshift, the next stop codon in frame is that employed in the Zampounis *et al.* gene model (at chromosome 8: 4,298,751 - 4,298,749). Thus, I conclude that the true *ChEC153* gene does not contain the 3' truncation relative to *ChEC153.1* that the Tsushima *et al.* gene model *CH35J_009900* predicts. Should the Tsushima *et al.* gene annotation for the *CH35J_009900* start codon be correct, the gene model would instead be as illustrated by *ChEC153.2a* in **Figure 5.5**.

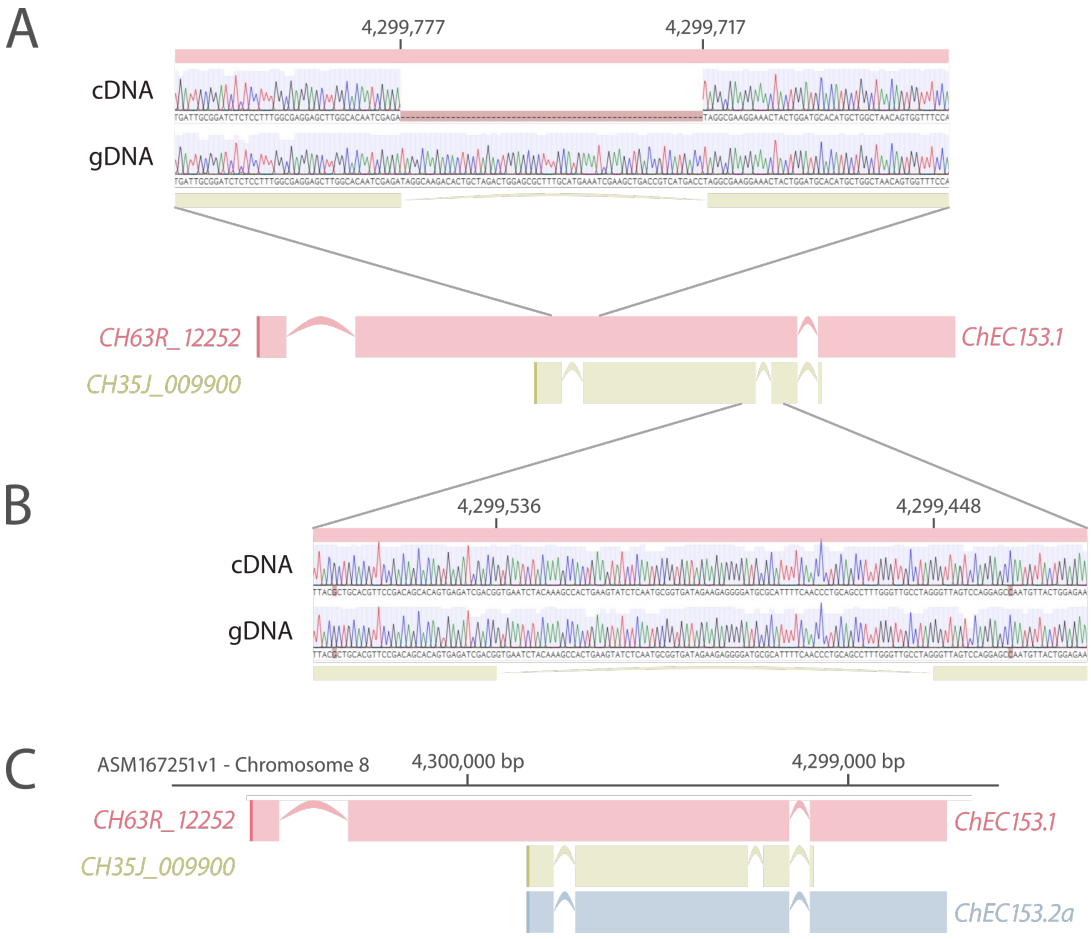


Figure 5.5: **Sanger sequencing of cloned regions of gDNA and cDNA for *ChEC153*. (A)** An intron is present at 4,299,777- 4,299,717, in agreement with the Tsushima *et al.* gene model for *CH35J_009900*. **(B)** An intron is not present at 4,299,536- 4,299,448, in agreement with the Zampounis *et al.* gene model for *CH63R_12252* (*ChEC153.1*). Bases highlighted in red indicate SNPs in the sequencing data (IMI 349061) relative to the Zampounis *et al.* gene sequence for IMI 349063. Genomic locations stated relate to chromosome 8 of the Zampounis *et al.* genome. **(C)** Potential new gene model, *ChEC153.2a* based on the Tsushima *et al*. *CH35J_009900* gene model start codon and introns confirmed by sequencing.

5.3.2.3 Localisation of ChEC153.2a in *N. benthamiana*

Based on the differences between *ChEC153.1* and potential gene model *ChEC153.2a*, I sought to localise ChEC153.2a in *N. benthamiana* to see whether this protein were capable of localising to the chloroplast. The predicted encoded sequence of *ChEC153.2a* was therefore synthesised, and cloned to produce vectors *35S::ChEC153.2a-eGFP* and *35S::ChEC153.2a-mCherry.* This was kindly carried out by Dr Andrew Breakspear. I transformed the assembled constructs that he provided into *Agrobacterium* and transiently expressed them in *N. benthamiana* by agroinfiltration. A lack of expression was observed for both constructs cloned with C-terminal eGFP and mCherry tags, and so it was not possible to determine the subcellular localisation of this hypothetical protein, thus suggesting that ChEC153.2a may not be a stable/functioning protein.

5.3.2.4 Determining the start site of *ChEC153*

5.3.2.4.1 Using 5' RACE as a strategy to determine the *ChEC153* transcription start site

Based on the above results, I sought to use 5' RACE to establish the full-length sequence of the *ChEC153* gene. There are two apparent potential start codons for the coding sequence. The codon predicted as the translation start site in the Tsushima *et al.* gene model (chromosome 8: 4,299,844 - 4,299,842, Zampounis *et al.* genome) is 713 bp downstream of that predicted as the start site in the Zampounis *et al.* gene model (chromosome 8: 4,300,557 - 4,300,555, Zampounis *et al.* genome). I aimed to use 5' RACE to inform on the transcriptional start site of the gene encoding the 3' region of interest, and determine the coding sequence present.

Difficulties in amplifying the 5' region of the gene from cDNA synthesised using random primers (**Figure 5.4**) may result from low yield cDNA synthesis, leading to inadequate template strands. To counter this, a gene-specific cDNA synthesis approach may help to increase the chances of targeted amplification producing the appropriate template. To carry out 5' RACE, I therefore used the *ChEC153* gene-specific primer GSP1 (**Appendix 1 Tab. E**) to synthesise gene-specific (GS) cDNA from the RNA I isolated from Arabidopsis leaves infected with *C. higginsianum*. Alongside processing the RNA I isolated, I carried out the control reactions recommended in the kit, mixing my sample RNA with the provided control RNA and amplifying at various points of the RACE protocol using the control primers. Control cDNA was synthesised using control primer C1. Amplification using control primers 2 and 3 (C2/C3) produced the expected 157 bp amplicon prior to and following the cDNA purification step (**Figure 5.6 A**). Further, amplification of the control sample with the oligo(dT) anchor primer and C2 produced an amplicon of 293 bp, as anticipated (**Figure 5.6 A**). As the control reactions produced the expected amplicons at each stage, I confirmed

that my test RNA sample was free of RNase and that no obvious technical issues occurred during the protocol.

Products of the test reaction PCR1, which sought to amplify using a gene-specific primer (GSP2) in combination with the oligo(dT)-anchor PCR primer, produced a smear on the agarose gel following electrophoresis, with no clear discrete bands visible. Similarly, the reaction PCR2, to amplify the product of PCR1 using a nested gene-specific primer (GSP3) and the anchor PCR primer resulted in a fainter smear, in which no bands could be distinguished (Figure 5.6 B).

I carried out further control reactions using the primer pair 100/101 which amplifies a 426 bp region of the *ChEC153* CDS reliably from other cDNA and gDNA samples (Figure 5.4), amplifying from GS cDNA samples taken before and after the polyadenylation reaction. As shown in Figure 5.6 C, I observed a faint amplicon of the expected size prior to the polyadenylation reaction, confirming the presence of at least part of *ChEC153* in the GS cDNA. However, this amplicon was not observed from the polyadenylated sample, suggesting that the cDNA may have been degraded at this stage. From these results, I was not able to infer any information about the 5' region of *ChEC153* to inform on the transcription or translation start site for this gene.

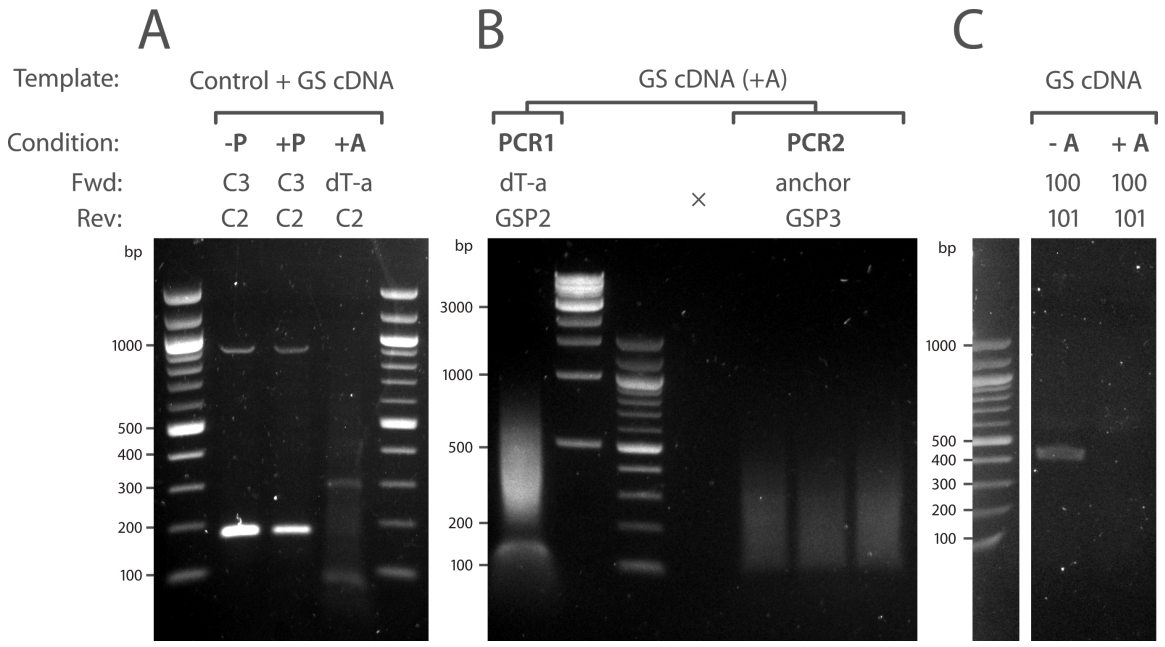


Figure 5.6: 5' RACE did not produce clear amplicons. (A) Control cDNA was mixed with cDNA synthesised from extracted RNA using a *ChEC153* gene-specific primer (GSP1) as a template for control reactions, either with or without purification (indicated by +P (with) or-P (without)). Control reactions: using control primers C2 and C3 to amplify 157 bp, and using the oligo(dT)-anchor (dT-a) primer with C2 to amplify 293 bp from the same template following purification and polyadenylation (+A). **(B)** Purified polyadenylated (+A) gene-specific (GS) cDNA was used as a template for PCR1 and PCR2 reactions using the primers indicated. PCR2 lanes are products when 0.5, 1, or 1.5 µL of PCR1 was used as template, from left to right. **(C)** Primers 100 and 101 used to amplify from GS cDNA, either with (+A) or without (-A) polyadenylation. Expected amplicon size: 426 bp. PCR products were run on 1.5% agarose gels stained with EtBr alongside 100 bp and 1 kb DNA ladders (NEB), with sizes in bp indicated to the left-hand side of each panel. "×" indicates an intentionally empty lane.

5.3.2.4.2 Further RT-PCR-based analysis of *C. higginsianum* cDNA

Based on inconclusive results from the 5' RACE experiment (**Figure 5.6**), I sought to confirm whether *ChEC153* transcription is likely to start at the Zampounis *et al.* or Tsushima *et al.* annotated start sites using RT-PCR-based approaches. As attempts to amplify the entire *ChEC153.1* gene were initially unsuccessful (**Figure 5.4**), I tried to amplify a shorter region at the 5' of the gene model, from both cDNA and gDNA. I hypothesised that as I used random primers for cDNA synthesis, amplification of a smaller fragment of the gene would be less technically challenging than amplifying the entire gene, as the likelihood of cDNA synthesis primers binding downstream of the amplification primers would be increased. I sought to amplify a short sequence at chromosome 8: 4,300,557 - 4,300,266 (Zampounis *et al.* genome) using primers 097 and 109, as shown in **Figure 5.7 A**. Amplicons were obtained from reactions with gDNA and cDNA templates, indicating that this region is likely to be transcribed (**Figure 5.7 B**). The predicted size of amplicon from the gDNA

is 318 bp (including primer cloning extensions), while the Zampounis *et al.* gene model predicts an amplicon from cDNA of 133 bp using these primers. The amplicon I obtained from the cDNA appears to be larger than 133 bp. The RNA-Seq reads mapped to this region of the genome indicate the presence of a possible additional exon between the two primers, indicated with a purple asterisk in Figure 5.7 A (chromosome 8: 4,300,435 - 4,300,369, Zampounis *et al.* genome). If this is a true exon present in the cDNA, the expected amplicon obtained with primers 097/109 would be 200 bp, which is much closer to the size of amplicon that I observed. Therefore, I suggest that this putative exon is part of the coding sequence in this region. The inclusion of this 67 bp exon in the coding sequence results in a frameshift and premature stop codon according to the Zampounis *et al.* gene model (chromosome 8: 4,300,268 – 4,300,266), such that the putative protein encoded in this region is only 57 amino acids in length. Therefore, it is possible that the genome region covered by *ChEC153.1* contains two separate coding sequences: one annotated in Figure 5.7 A as *ChEC153.2b*, and a second CDS starting downstream, at the Tsushima *et al.* gene model predicted start codon, annotated as *ChEC153.2a.*

The RNA-Seq reads mapping to this region from the O'Connell *et al.* data indicate a possible intron at chromosome 8: 4,300,269 – 4,300,221 (Zampounis *et al*. genome, Figure 5.7 A, orange asterisk) which, if present, would mean that the stop codon at chromosome 8: 4,300,268 – 4,300,266 would be absent from the coding sequence. This would result in the reading-frame subsequently returning to that seen in the Zampounis *et al.* gene model, upstream of the Tsushima *et al.* gene model start codon. This additional potential gene model is annotated in Figure 5.7 A as *ChEC153.3*, with the coding sequence having a predicted size of 1524 bp without introns, spanning the entire length of the *ChEC153.1* gene model (chromosome 8: 4,300,557 - 4,298,749, Zampounis *et al.* genome) but with different intron/exon boundaries. *ChEC153.3* would thereby encode a 507 amino acid protein.

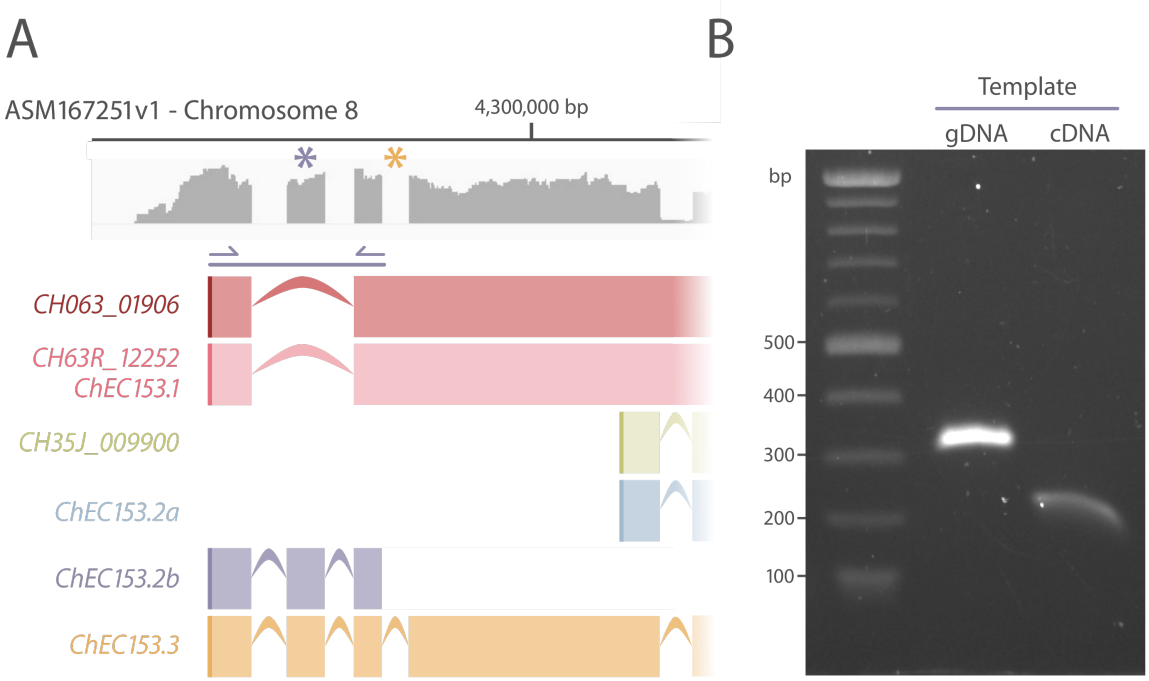


Figure 5.7: **RT-PCR amplification of an upstream region of *ChEC153*. (A)** A schematic depicting the PCR primer binding sites. Only the 5' of the region of interest is shown for simplicity, the gene models fading to white continue downstream. The purple asterisk over the RNA-Seq coverage plot (O'Connell *et al.* data mapped to Zampounis *et al.* assembly) annotates a potential exon resulting in the *ChEC153.2b* gene model, while the orange asterisk annotates a potential intron which, in combination with the potential exon, would result in the continuation of the CDS (*ChEC153.3*). **(B)** PCR products following amplification with primer pair 097/109 were separated on a 2% agarose gel stained with EtBr, alongside a 100 bp DNA ladder (NEB). The PCR template used for each reaction was 0.5 µL of: cDNA from samples harvested at 40-hours post inoculation (hpi), or gDNA from *ex planta* fungal cultures, as indicated above each lane.

Based on this result, and to provide further evidence for the *ChEC153.3* gene model, Dr Andrew Breakspear kindly carried out some further gene model validation experiments, as described in **Section 5.2.1.2**. He extracted RNA from tissue I had harvested at 24-, 40-, and 46-hours post inoculation of Arabidopsis with either *ΔChKu80* (background strain: MAFF 305635) or wild-type IMI 349061 *C. higginsianum* conidia. From this, he synthesised cDNA using oligo(dT) primers in order to enrich for full-length transcripts. As shown in **Figure 5.8**, he was able to acquire amplicons of the expected sizes from all samples, with primers 100/101 (positive control: 426 bp region previously successfully amplified (**Figure 5.3** and **Figure 5.4**)), and with primer pairs 097/101 and 097/099 to amplify the 1049 bp 5' region (1078 bp with primer cloning extensions) and the full length gene *ChEC153.3*, respectively, both of which had previously failed to produce amplicons in my hands from the original infection random primer cDNA (**Figure 5.4**). As the amplicon that Dr Andrew Breakspear obtained with primers 097/099 is approximately the size expected for the *ChEC153.3* gene model (1524 bp, 1553 bp with primer cloning extensions), he purified, cloned and sequenced this amplicon. The sequencing results confirmed the gene model *ChEC153.3* to be

correct, with all expected intron/exon boundaries as predicted. Based on a comprehensive approach using RT-PCR, cloning and sequencing, 5' RACE, and computational alignments with published RNA-Seq data, I conclude that *ChEC153.3* is the true model for the gene (*ChEC153*) encoded in this region.

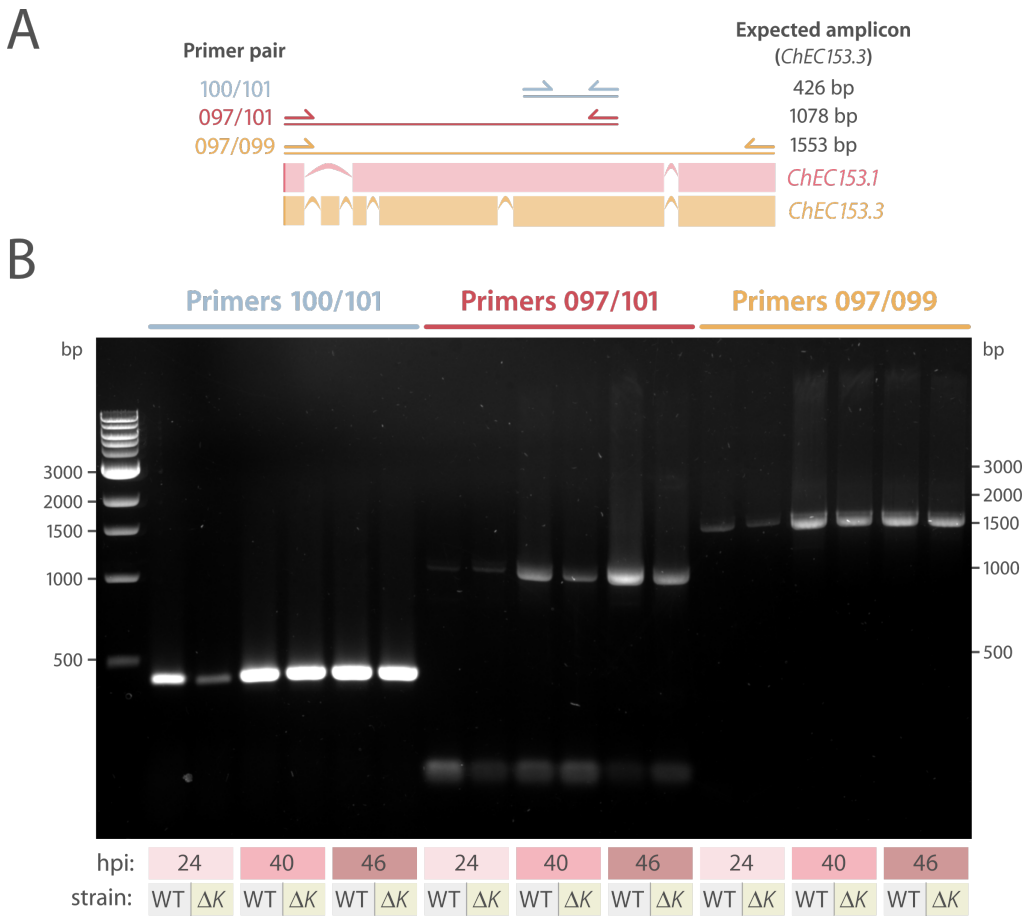


Figure 5.8: RT-PCR of *ChEC153.3*. **(A)** Schematic depicting the relevant primer binding sites with respect to the *ChEC153.1* and *ChEC153.3* gene models, and expected amplicon sizes in bp (including primer cloning extensions) assuming the *ChEC153.3* gene model to be correct. **(B)** Products of RT-PCR with the primer pairs depicted in (A) amplifying from RNA extracted from Col-0 infected with either wild-type (WT) or *ΔChKu80* (*ΔK*) *C. higginsianum* conidia at 24-, 40-, and 46-hours post inoculation, as indicated below the gel. PCR products were separated on a 1% agarose gel stained with EtBr alongside a 1 kb DNA ladder (NEB).

5.3.2.5 Summary of the gene model for *ChEC153: ChEC153.3*

Through the results of this section, I progressively identified several potential gene models for *ChEC153* (**Figure 5.9**). The work presented here validated that the true gene model is *ChEC153.3*. This model *ChEC153.3* is similar to the previously assumed gene model *ChEC153.1* (*CH63R_12252*; Zampounis et al., 2016) which informed the work I conducted throughout **Chapter 3** and **Chapter 4** of this project. With respect to *ChEC153.1*, *ChEC153.3* contains an additional exon at 4,300,435 - 4,300,369 and additional introns at 4,300,269 - 4,300,221 and 4,299,777 - 4,299,717 (Zampounis et al., 2016). Notably, neither of the published gene model annotations were completely correct according to my findings, with the *CH35J_009900* gene model of the Tsushima *et al*. genome assembly (Tsushima et al., 2019a) predicting the most divergent coding sequence.

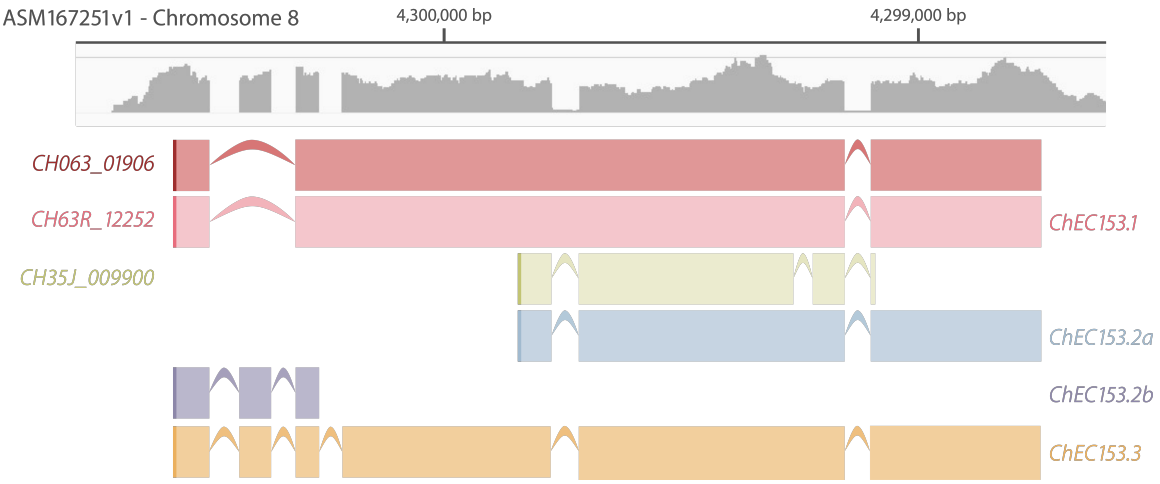


Figure 5.9: Summary of *ChEC153* gene models. Schematic depicting the three published *ChEC153* gene models, *CH063_01906*, *CH63R_12252*, and *CH35J_009900* from the O'Connell *et al.*, Zampounis *et al.*, and Tsushima *et al.* genome assemblies, respectively, and the new gene models elucidated here. *ChEC153.3* (shown in orange) has been confirmed by sequencing of cloned cDNA and is assumed to be the predominant gene isoform. The alternative two coding sequences in this region (*ChEC153.2a* and *ChEC153.2b*) are annotated in blue and purple, and could represent splice variants of *ChEC153.3*. The genome location is shown relative to the Zampounis *et al.* genome above the gene models, with read coverage from the O'Connell *et al.* RNA-Seq data shown in grey. Exons are shown as coloured rectangles, with introns shown as arcs. The start codon is highlighted in a darker colour at the start of the first exon for each gene model.

5.3.2.6 Protein ChEC153.3 is encoded by the new gene model *ChEC153.3*

The *ChEC153.3* gene model would encode a 507 amino acid protein (ChEC153.3), without the large N-terminal and C-terminal truncations suggested by the Tsushima *et al.* gene model *CH35J_009900* (Tsushima et al., 2019a), but still with some differences relative to ChEC153.1 (**Figure 5.10 A**). ChEC153.3 and ChEC153.1 agree on the first 19 amino acids, but due to the presence of an additional exon at chromosome 8: 4,300,435 - 4,300,369, and subsequent frameshift, followed by an intron at chromosome 8: 4,300,269 - 4,300,221 in *ChEC153.3*, the protein sequences diverge. This region consists of 38 amino acids in ChEC153.3 and 32 amino acids in ChEC153.1, differing until sequence identity is restored at ChEC153.3 residue Ile58 (from chromosome 8: 4,300,219). The subsequent 148 amino acids are identical to those seen in the ChEC153.1, until a 20 amino acid deletion in ChEC153.3 relative to ChEC153.1 (where an intron was identified in *ChEC153.3*, chromosome 8: 4,299,777 - 4,299,717). The C-terminal 303 amino acids are identical between the two proteins. Both proteins are expected to be cytochrome P450 monooxygenases according to domain prediction analysis.

I used AlphaFold2 via the ColabFold platform to predict the structures of both ChEC153.1 and ChEC153.3 (**Figure 5.10 B**; Jumper et al., 2021, Mirdita et al., 2022). The two predicted structures are very similar, with one of the major differences being the extended alpha helix at the N-terminus of ChEC153.3 relative to ChEC153.1, which correlates with the prediction of an N-terminal transmembrane region rather than signal peptide for this protein. As highlighted by an asterisk in **Figure 5.10 B**, the region near the centre of the ChEC153.1 sequence now known to correspond to an intron and therefore absent in ChEC153.3 was predicted with a lower confidence than the surrounding regions of the protein.

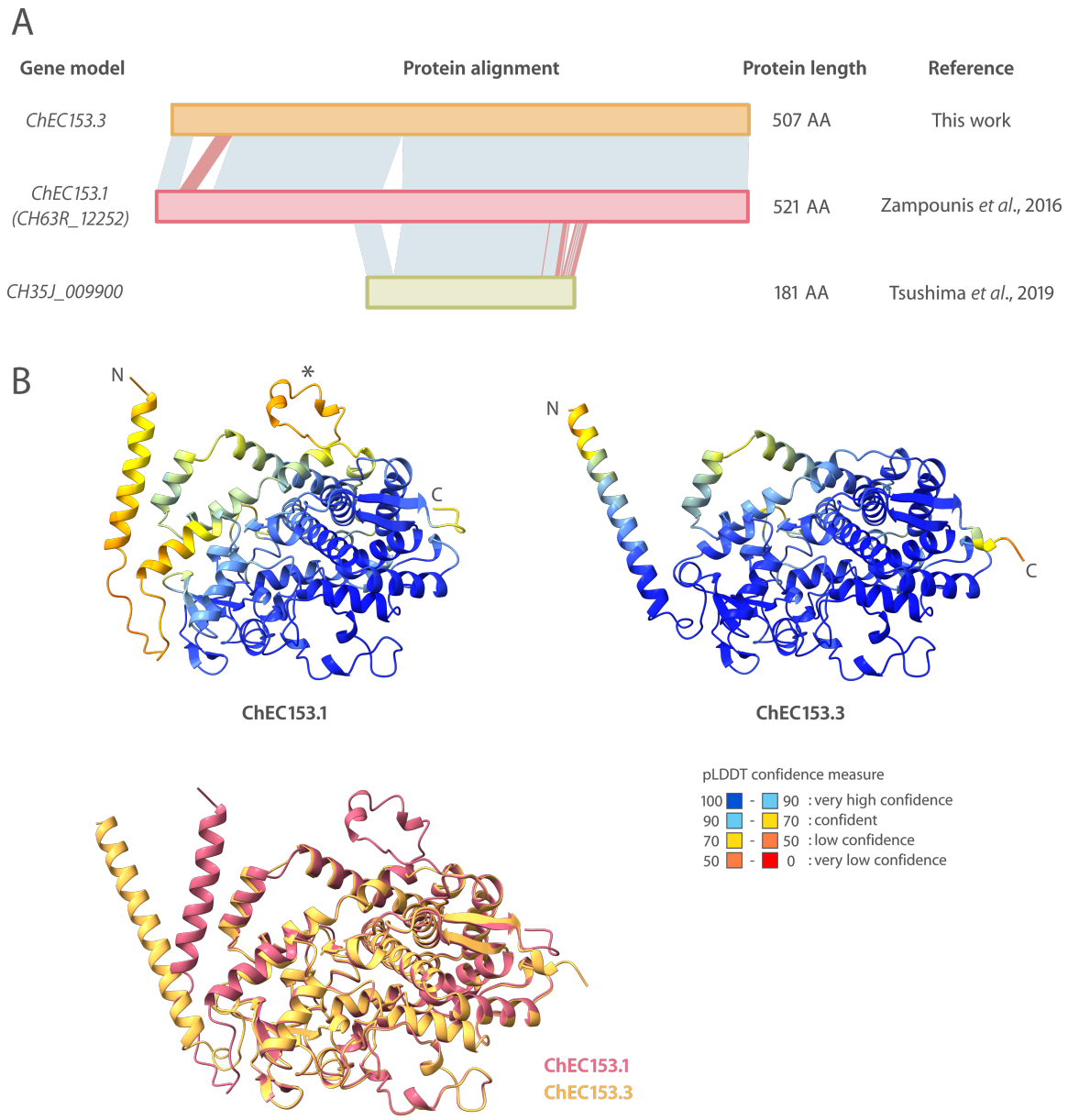


Figure 5.10: **Proteins encoded by *ChEC153* gene models: alignments of sequences and predicted structures. (A)** Schematic of the proteins predicted to be encoded by *ChEC153.3* (this work) and *CH35J_009900* (Tsushima et al., 2019a) are aligned relative to the ChEC153.1 sequence (*CH63R_12252*, Zampounis et al., 2016). Regions where amino acid sequences are identical are mapped in blue, with regions where amino acids differ shown in red. **(B)** Structure predictions of ChEC153.1 and ChEC153.3, generated using AlphaFold2. Protein structures are shown separately, coloured according to the pLDDT values, with N and C termini labelled. An asterisk highlights the region of ChEC153.1 now know to be an intron and therefore absent in ChEC153.3. The predicted structures are also shown aligned relative to one another using the Matchmaker function in ChimeraX.

The mature coding sequence of putative effector ChEC153.1 used in **Chapters 3** and **4** of this thesis was synthesised based on the Zampounis *et al.* gene model *CH63R_12252* (*ChEC153.1*) with the signal peptide (residues 1-20) absent. Given the differences between ChEC153.3 and ChEC153.1, I sought to evaluate whether ChEC153.3 also possessed a predicted signal peptide. I used *in silico*

signal peptide prediction tools to evaluate whether the various predicted proteins encoded for in the region of *ChEC153* (**Figure 5.10**) are predicted to have signal peptides for secretion from the fungus. The results of these various tools are presented in **Table 5.1**. The various signal peptide prediction tools used gave varying predictions for each gene model-encoded protein. While the original ChEC153.1 protein sequence was predicted by SignalP 3.0, SignalP 4.1 and TargetP 2.0 to contain a signal peptide, it was not predicted as an effector by EffectorP 3.0, possibly due to its large size relative to the majority of conventional effectors. Conversely, the protein encoded by the *CH35J_009900* gene model was identified as a putative effector by EffectorP 3.0, but the SignalP and TargetP tools did not predict a signal peptide to be present. The protein encoded by putative gene model *ChEC153.2a* was consistently predicted not to possess a signal peptide, or to be an effector. The protein encoded by *ChEC153.3* was predicted to have a signal peptide according to SignalP 3.0 and TargetP 2.0, but not SignalP 4.1 – neither was it predicted to be an effector by EffectorP 3.0 (Bendtsen et al., 2004, Petersen et al., 2011, Armenteros et al., 2019a, Sperschneider and Dodds, 2022). The diverse predictions assigned by various tools highlight that while prediction tools remain a valuable asset in molecular biology, experimental validation is still required to determine conclusively whether ChEC153.3 is a secreted effector protein.

Table 5.1: **Prediction of signal peptides and effector prediction using *in silico* tools for the proteins encoded by the various *ChEC153* potential gene models.** The presence (SP) or absence (No SP) of a predicted signal peptide is indicated. For EffectorP 3.0 prediction of whether proteins are effectors, N indicates not an effector, while Y (cyt) indicates that the sequence was predicted as a cytoplasmic effector. The signal peptide cleavage points are indicated by the amino acid numbers in brackets. The *ChEC153.3* gene model found through this work to be the true gene model present at this region of the *C. higginsianum* genome is highlighted in yellow.

Gene model	Protein length (AA)	SignalP 3.0	SignalP 4.1	TargetP 2.0	EffectorP 3.0
ChEC153.1	521	SP (20/21)	SP (20/21)	SP (30/31)	N
CH35J_009900	181	No SP	No SP	Other	Y (cyt)
ChEC153.2a	325	No SP	No SP	Other	N
ChEC153.3	507	SP (23/24)	No SP	SP (27/28)	N

5.3.3 Dual transcriptome analysis of Arabidopsis infected with *C. higginsianum* strains *ΔChEC153-2* and *ΔChKu80*

Irrespective of the *ChEC153* gene model and the secretion of ChEC153.1/ChEC153.3 as an effector during infection, transcriptomics data show upregulation of this gene during infection (O'Connell et al., 2012, Dallery et al., 2017), suggesting that this gene has a virulence-associated function. To explore this, and simultaneously assay for *ChEC153.3*-dependent processes in both infection and/or host responses, I carried out dual transcriptome analyses with the assistance of Dr Andrew Breakspear and Dr Hannah Rae Thomas. With this experiment, I sought to identify transcriptional differences during infection of Arabidopsis leaves with *ΔChEC153-2* relative to the background strain *ΔChKu80* in both the plant and pathogen, which could suggest potential roles of ChEC153.3. As the *ΔChEC153-2 C. higginsianum* strain I generated in **Section 4.3.2.1** was designed to remove the entire *ChEC153.1* gene, it functions as a knockout of *ChEC153.3* and any putative gene isoforms, thereby providing an appropriate strain for analysis.

5.3.3.1 RNA-Seq sample preparation and validation

The experimental preparation of samples was carried out by Dr Andrew Breakspear according to the methods presented in **Section 5.2.3.1**. He inoculated detached, mature Col-0 Arabidopsis leaves with multiple droplets of conidial suspensions from *C. higginsianum* strains *ΔChEC153-2* and *ΔChKu80*, alongside a mock-inoculated control (inoculated with water), collecting tissue at 24- and 36-hours post inoculation and extracting RNA.

Dr Andrew Breakspear carried out quality control checks to ensure the presence of both plant and fungal RNA in the samples. To assay for Arabidopsis RNA, he employed RT-qPCR to amplify the Arabidopsis *UBIQUITIN 10* (*UBQ10*) housekeeping gene (*AT4G05320*) with primers UBQ_Fwd and UBQ_Rev (detailed in **Appendix 1 Tab. E**). Amplification of *AtUBQ10* was seen for all samples, as shown by the cycling threshold values presented in **Figure 5.11 A**, and relative expression values plotted in **Figure 5.11 B**. These results confirm that *AtUBQ10* is expressed in all samples, and to a relatively similar extent.

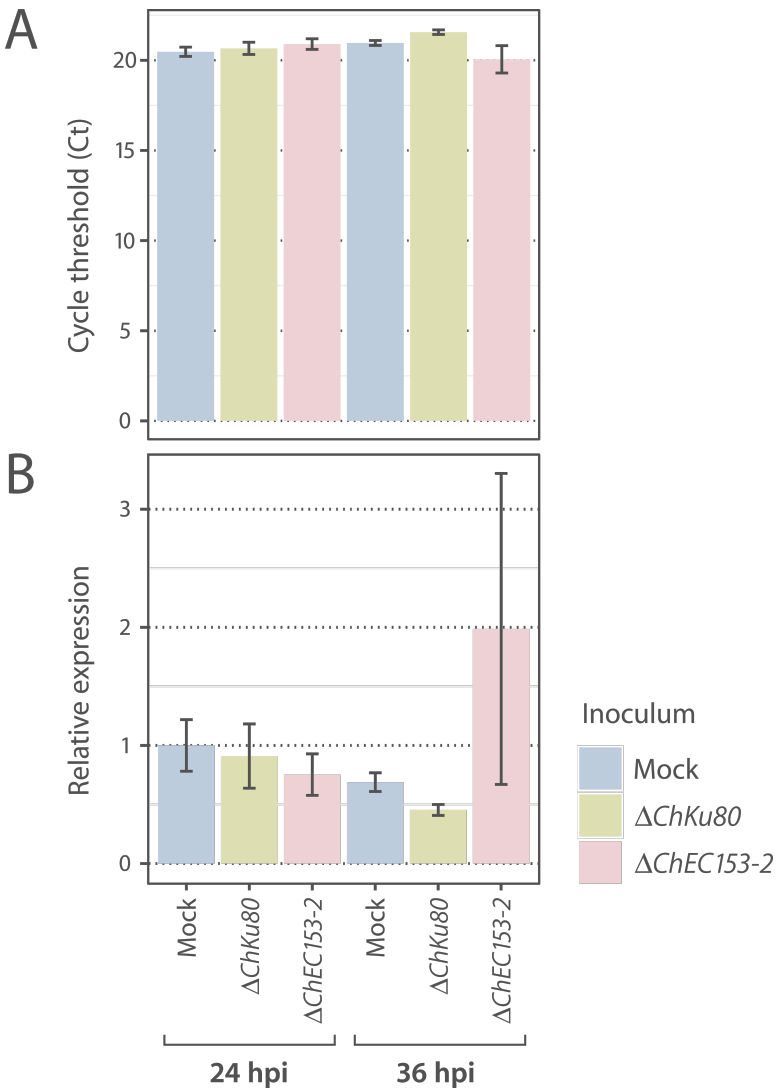


Figure 5.11: **Validation of RNA: reverse transcription quantitative PCR amplification of *AtUBQ10* from cDNA samples, confirming Arabidopsis gene expression. (A)** Mean raw cycle threshold (Ct) values for each sample. **(B)** Relative expression of *AtUBQ10* in each sample, calculated as the average relative quantity ($E^{-\Delta Ct}$) relative to the expression of *AtUBQ10* in the 24-hours post inoculation (24 hpi) mock sample. Data represent mean values for the three biological replicates for each time-point/inoculum combination (each being the average of three technical replicates), with error bars representing the standard error.

Further, to assay for *C. higginsianum* gene expression, one sample for each time-point/inoculum combination was subject to RT-PCR. Dr Andrew Breakspear carried out first strand synthesis from RNA samples using random primers to generate cDNA template for validation via PCR. He used primers 102 and 103 in amplification reactions to confirm the presence of cDNA for the highly expressed putative effector *ChEC6* (*CH63R_13188*, expressed at the same infection stages as *ChEC153*; Dallery et al., 2017) in infected samples, with the resulting agarose gel of PCR products shown in **Figure 5.12**. From this gel, the presence of fungal RNA was confirmed at both time-points in both infected samples by the presence of the expected size (293 bp) cDNA amplicon band. Resulting bands were of comparable intensities, suggesting that there are no obvious drastic

differences in the quantity of fungal RNA between *ΔChKu80* and *ΔChEC153-2* inoculations. The mock-inoculated control samples showed only faint, non-specific amplification, confirming the absence of *C. higginsianum* RNA in these samples, as expected.

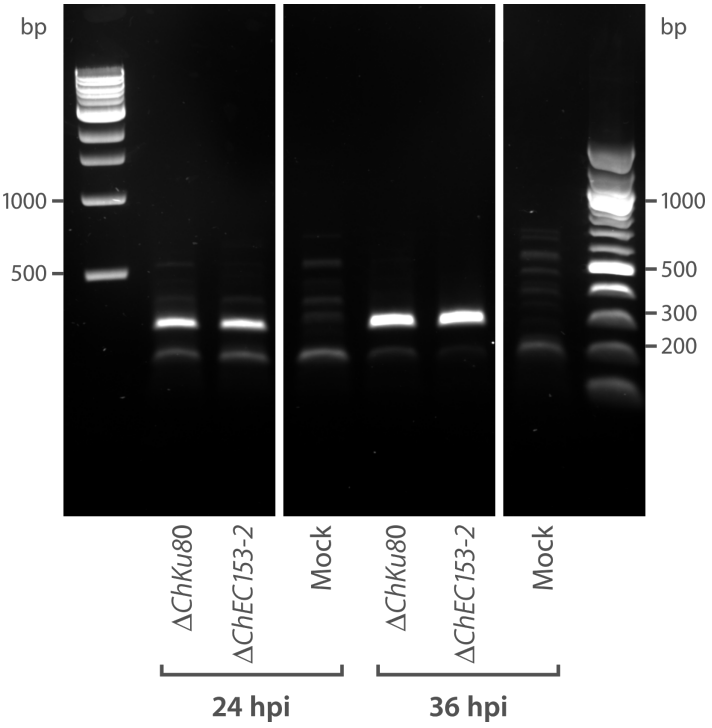


Figure 5.12: Validation of RNA: RT-PCR to amplify *ChEC6* from RNA samples, confirming *C. higginsianum* gene expression. Samples are derived from Arabidopsis leaves inoculated with *C. higginsianum* strains *ΔChKu80* or *ΔChEC153-2*, or mock (water) harvested at 24- or 36-hours post inoculation (hpi). One biological replicate of each RNA sample was used for cDNA synthesis and amplification to confirm the detected expression of *ChEC6* (*CH63R_13188,* expected amplicon size: 293 bp), as an indicator of *C. higginsianum* RNA in samples. Mock samples amplified only faint, non-specific bands. Products were separated on a 1.5% agarose gel stained with EtBr alongside 1 kb and 100 bp ladders.

5.3.3.2 RNA-Seq data preliminary analyses

Following the abovementioned validation experiments to confirm the presence of RNA from both the host and pathogen, the RNA samples were submitted to GENEWIZ for RNA-Sequencing. Dr Hannah Rae Thomas carried out pre-processing of the resulting RNA-Seq data as described in Section 5.2.3.4, using a cross-mapping approach whereby reads were first aligned to the Arabidopsis genome (TAIR10), and unmapped reads subsequently aligned to the *C. higginsianum* genome, thus preventing reads from potentially being mapped to both genomes. This resulted in a small read pool capable of mapping uniquely to the fungal genome. The Tsushima *et al.* genome assembly ASM492035v1 (Tsushima et al., 2019a) was chosen as the mapping reference for fungal

reads. The strain sequenced by Tsushima and colleagues was MAFF 305635-RFP (Hiruma et al., 2010, Tsushima et al., 2019a), generated from the MAFF 305635 strain which is also the background strain for the *ΔChKu80* strain used as the control in this RNA-Seq experiment and which I used to generate the *ΔChEC153-2* knockout strain.

I then analysed the pre-processed RNA-Seq data in R with guidance from Dr Hannah Rae Thomas. I observed that the number of normalised read counts acquired for each sample are consistent within each species, with the total number of normalised read counts for reads mapping to the Arabidopsis and *C. higginsianum* genomes shown in **Figure 5.13**. Notably, the vast majority of reads mapped to the Arabidopsis genome (> 96% of reads for each sample), with comparatively few reads mapping to the *C. higginsianum* genome. Some reads appear to map to the *C. higginsianum* genome in the mock-inoculated samples. In **Figure 5.13**, this is exacerbated by the normalisation process. Non-normalised read count data are presented in **Appendix 3 Fig. A**. These likely represent background reads, possibly derived from other species present in the system due to the non-sterile conditions used for plant growth, rather than *C. higginsianum* contamination. The number of reads mapping to the *C. higginsianum* genome is considerably lower in the mock-inoculated samples than in the infected samples.

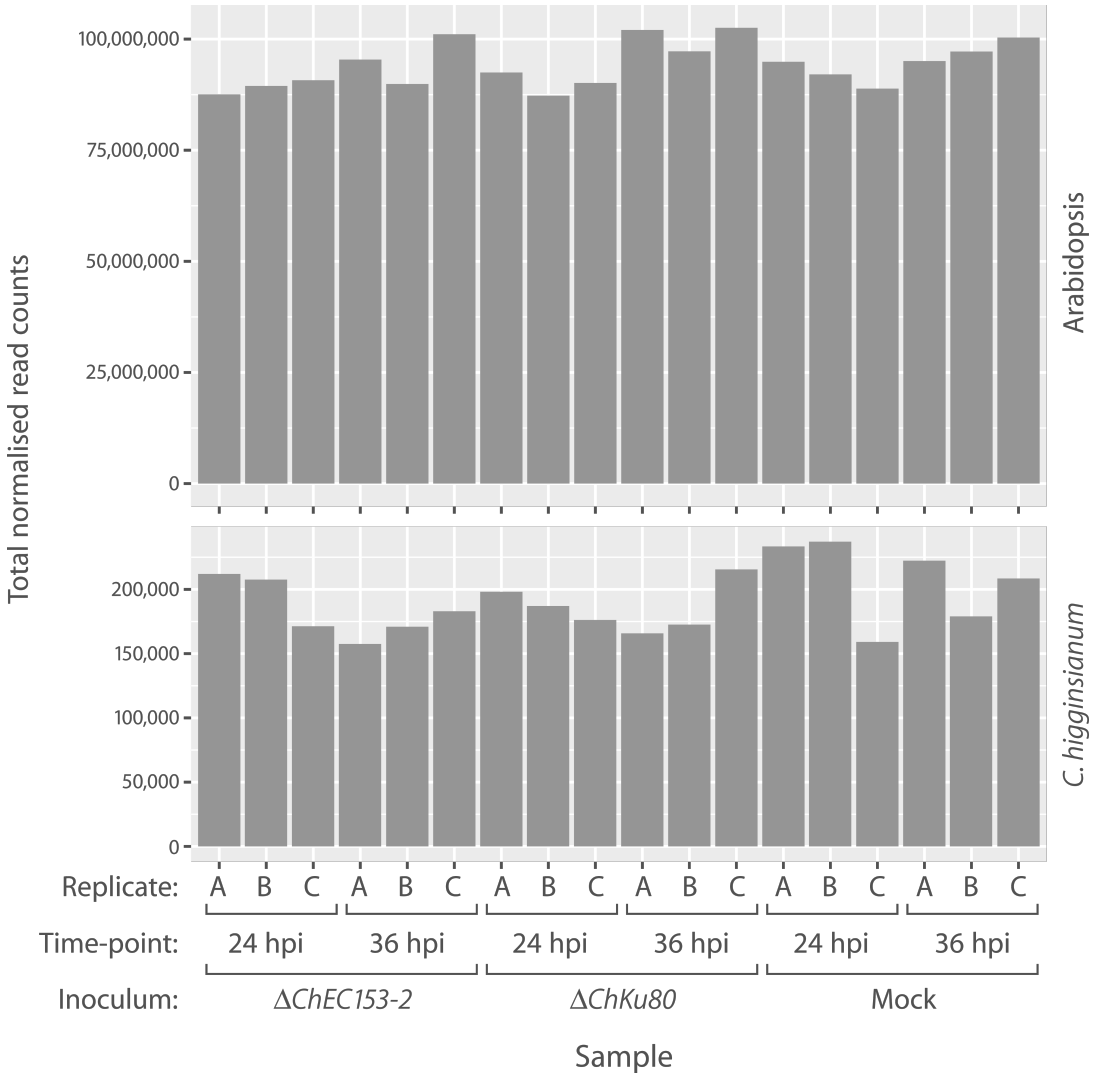


Figure 5.13: **Total normalised read counts for three biological replicates of Arabidopsis treated with *ΔChEC153-2*, *ΔChKu80*, or mock (water) at 24- or 36-hours post inoculation.** Total normalised read counts mapping to the Arabidopsis genome (top), and the *C. higginsianum* genome (bottom, Tsushima *et al. C. higginsianum* genome, ASM492035v1). Biological replicates are indicated by letters A, B, and C.

To evaluate the suitability of using the DESeq2 negative binomial distribution model for analysing these data, I generated and evaluated dispersion plots (**Figure 5.14 A** and **B**). The dispersion plots show that the model fits the data well for both species, indicating that it is appropriate for use in examining differences between samples.

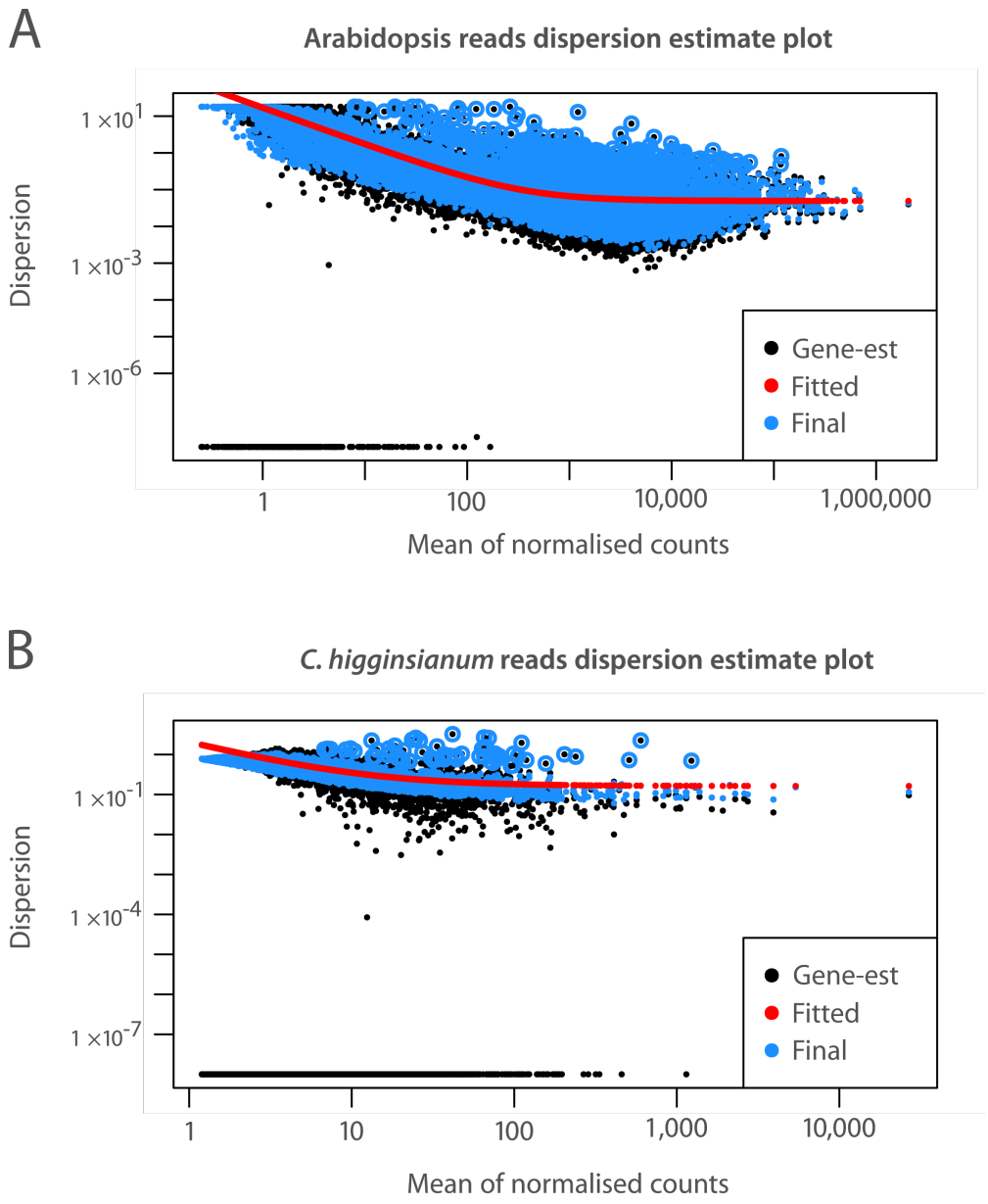


Figure 5.14: Dispersion plots of RNA-Seq reads for the *C. higginsianum*-Arabidopsis pathosystem. **(A)** Dispersion plot of reads mapping to the Arabidopsis genome. **(B)** Dispersion plot of reads mapping to the *C. higginsianum* genome. Plots show the dispersion and mean expression level (mean of normalised counts) for each gene. Black points ("Gene-est") show the maximum likelihood dispersion estimates for each gene. Red points ("fitted") show the curve fitted to the dispersion estimates. The blue points ("final") show the final dispersion estimates for each gene following shrinkage toward the fitted curve using a Bayes approach. Black points within blue circles represent dispersion outliers, to which shrinkage was not applied.

5.3.3.3 Principal component analysis reveals little variation between samples based on Arabidopsis reads

I first used principal component analysis (PCA) to evaluate the degree of variation between samples within and between inocula. I found that PCA of the Arabidopsis reads following inoculation with *ΔChEC153-2*, *ΔChKu80*, and mock (water) shows the main component of variance (PC1) to be time (**Figure 5.15 A**). Therefore, I plotted each time-point separately to look for variation between inocula (**Figure 5.15 B**). Within each inoculum, there is little variance between samples. However, there also appears to be little variance between inocula. Aside from time, there is no other component clearly driving variance. Notably, I observed little separation between Arabidopsis reads at either time-point between mock- and *C. higginsianum*-inoculated samples.

At 24 hpi, 18 Arabidopsis genes were identified as differentially expressed between *ΔKu80*- and mock-inoculated samples using a loose threshold for differential expression of $\log_2$ fold change between samples of > 0.5 or < -0.5, and an adjusted Wald test p-value of < 0.1. At 36 hpi, 42 Arabidopsis genes met these criteria. The Arabidopsis genes identified as differentially expressed between *ΔKu80*-infected and mock-inoculated samples are listed in **Appendix 3 Tab. A**. While the scale of the transcriptional response to infection generally appears to be small, several of the differentially expressed genes point to the host responding to biotic stress during infection (e.g. *CYSTEINE-RICH RLK 31* (*CRK31*), *AT2G43590* (*PR-3* like gene), *PEROXIDASE 33* (*PRX33*)).

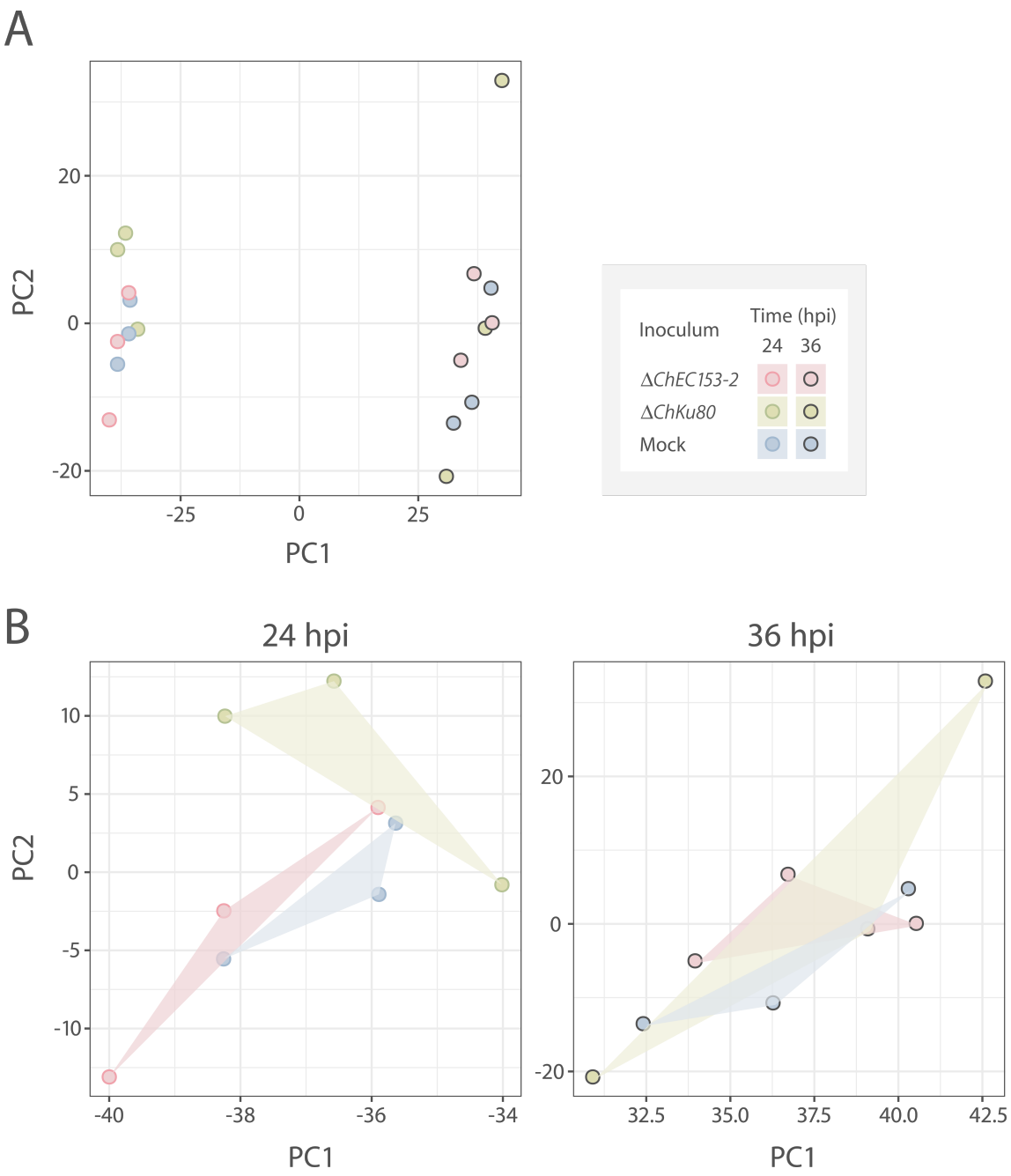


Figure 5.15: PCA of RNA-Seq reads mapping to the Arabidopsis genome. (A) PCA of all data, including both time-points. **(B)** PCA of data separated by time-point. Samples are coloured based on the inoculum with which they were treated as indicated in the key: *ΔChEC153-2*, pink; *ΔChKu80*, green; mock, blue. Samples are outlined based on the time-point at which they were harvested: colours = 24 hpi, grey = 36 hpi.

5.3.3.4 Principal component analysis reveals little variation between *ΔChKu80* and *ΔChEC153-2* samples in terms of fungal reads

Next, I carried out similar PCA of the reads mapping to the *C. higginsianum* genome in order to gauge the degree of variation between samples within and between inocula. Similarly to PCA of the reads mapping to the Arabidopsis genome for each sample (**Figure 5.15**), I found that one of the main components of variance for the fungal reads is time (**Figure 5.16 A**).

The mock-inoculated samples for both time-points cluster closely together, with the low variance between these control samples suggesting there is little background noise in these data. To better separate additional components of variance, I then plotted the two time-points separately (**Figure 5.16 B**). Within each inoculum there is little variance between samples. The separation of the mock-inoculated samples from the infected samples at either time-point remains clear, as expected. The infection status (infected versus non-infected) appears to be a major component of variance between the samples in terms of fungal reads. There is little separation between the samples following infection with *ΔChEC153-2* compared to *ΔChKu80*. This lack of variance suggests there is no additional component driving variance between these inocula.

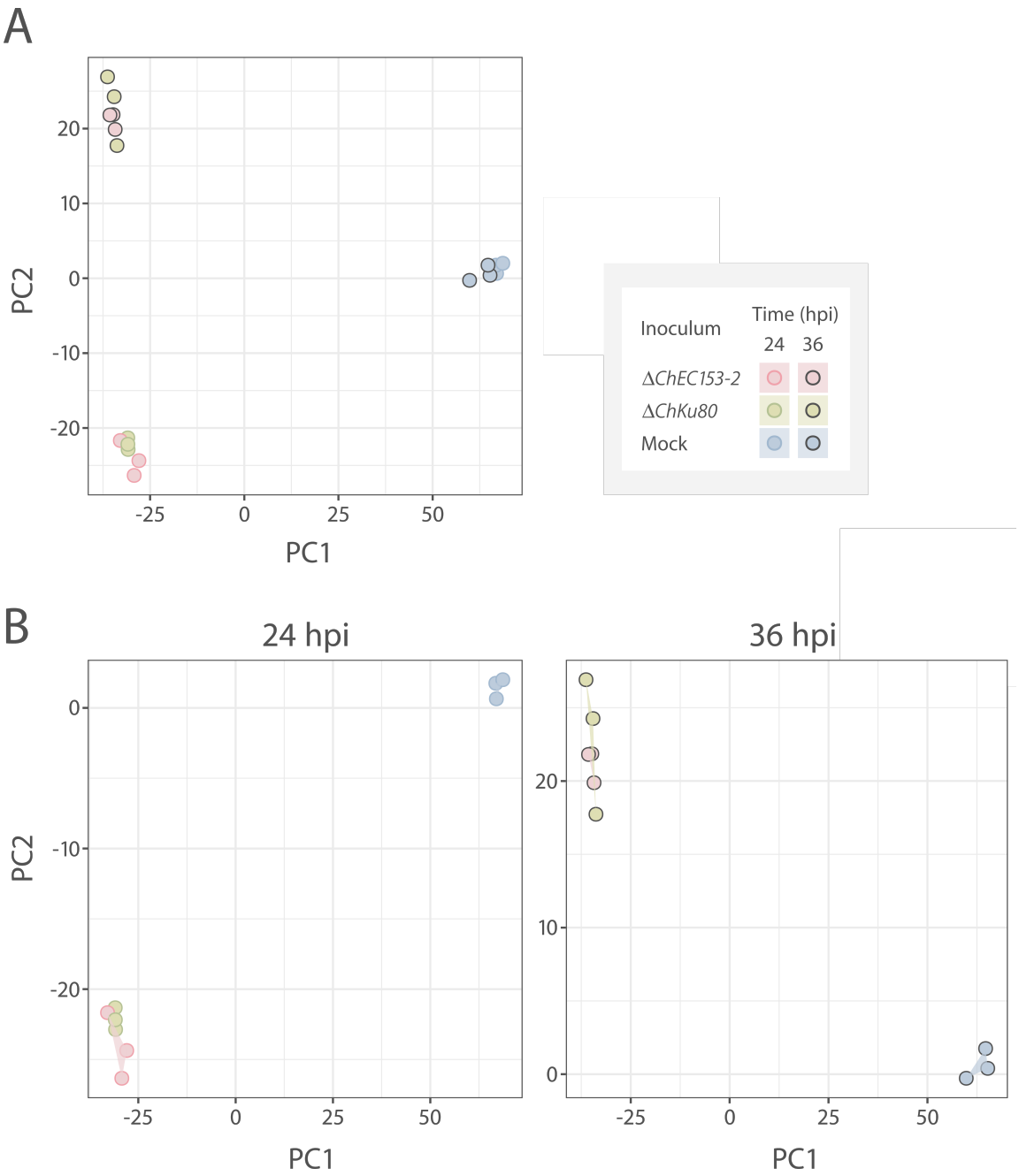


Figure 5.16: **PCA of RNA-Seq reads mapping to the *C. higginsianum* genome. (A)** PCA of all data, including both time-points. **(B)** PCA of data separated by time-point. Samples are coloured based on the inoculum with which they were treated as indicated in the key: *ΔChEC153-2*, pink; *ΔChKu80*, green; mock, blue. Samples are outlined based on the time-point at which they were harvested: colours = 24 hpi, grey = 36 hpi.

5.3.3.5 A number of Arabidopsis and *C. higginsianum* genes are differentially expressed between *ΔChKu80* and *ΔChEC153-2* infections

I hoped to probe the function of *ChEC153.3* as a putative effector in the infection context, and to identify any mechanisms being affected by the deletion of the putative effector gene. In order to do this, I sought to identify genes expressed at different levels during infection with *ΔChEC153-2* relative to the background strain *ΔChKu80*, in either the Arabidopsis host or the *C. higginsianum* pathogen.

5.3.3.5.1 Arabidopsis genes differentially expressed during infection with *ΔChEC153-2* versus *ΔChKu80* strains

I started by looking for host genes differentially expressed between *ΔChEC153-2*- and *ΔChKu80*-infected leaf samples to elucidate any differential host transcriptional responses that might point to the role of the putative effector. Due to the relatively small variance seen between samples inoculated with different *C. higginsianum* strains (**Figure 5.15**), I defined differential expression as a $\log_2$ fold change between samples of > 0.5 or < -0.5, and an adjusted Wald test p-value of < 0.1. **Table 5.2** lists the Arabidopsis genes differentially expressed during infection with *C. higginsianum* strain *ΔChEC153-2* relative to infection with *ΔChKu80*. At 24 hpi, six Arabidopsis genes are upregulated and two are downregulated in *ΔChEC153-2*-infected samples relative to *ΔChKu80*-infected samples, while at 36 hpi, ten Arabidopsis genes are upregulated and seven downregulated.

Table 5.2: **Differentially expressed Arabidopsis genes at 24- or 36-hours post inoculation with *ΔChEC153-2* relative to *ΔChKu80*.** AGI codes and names/descriptions (TAIR) of genes differentially expressed in *ΔChEC153-2*-infected samples are listed, alongside their log_2 fold change (log_2 FC) relative to samples infected with the background strain *ΔChKu80*, and respective adjusted Wald test p-values (p-adj). Differential expression defined as log_2 FC > 0.5 or <-0.5, p-adj < 0.1. Genes more highly expressed in samples inoculated with *ΔChEC153-2* relative to *ΔChKu80* are highlighted in pink, while those that are less highly expressed in *ΔChEC153-2*-inoculated samples are shown in green. Genes for which p-adj < 0.05 are highlighted in yellow.

Arabidopsis genes differentially expressed at 24 hpi with *ΔChEC153-2* vs. *ΔChKu80*			
Gene ID	Gene name / description	Log_2 FC	p-adj
AT4G39838	Antisense long non-coding RNA	19.00	0.00000915
AT5G01712	*CPuORF48*	-11.96	0.00177
AT5G39190	*GERMIN-LIKE PROTEIN 5A* (*GLP5A*)	2.29	0.0495
AT5G66640	*DA1-RELATED PROTEIN 3* (*DAR3*)	1.77	0.0524
AT2G43880	Pectin lyase-like superfamily protein	6.63	0.0641
AT5G35550	*TRANSPARENT TESTA 2* (*TT2, MYB123*)	-3.42	0.0684
AT2G41690	*HEAT SHOCK TRANSCRIPTION FACTOR B3* (*HSFB3*)	6.43	0.0684
AT5G39160	*GERMIN-LIKE PROTEIN 2A* (*GLP2A*)	2.07	0.0830
Arabidopsis genes differentially expressed at 36 hpi with *ΔChEC153-2* vs. *ΔChKu80*			
Gene ID	Gene name / description	Log_2 FC	p-adj
AT3G49620	*DARK INDUCIBLE 11* (*DIN11*)	20.85	0.000636
AT5G01712	*CPuORF48*	-12.57	0.000636
AT4G34588	*CPuORF2*	25.43	0.000880
AT3G01960	Hypothetical protein	-1.06	0.0350
AT5G62370	Tetratricopeptide repeat-like superfamily protein	1.22	0.0429
AT5G61330	rRNA processing protein-like protein	-0.57	0.0435
AT1G48598	*CPuORF31*	-5.04	0.0455
AT1G09240	*NICOTIANAMINE SYNTHASE 3* (*NAS3*)	3.68	0.0471
AT1G29680	Histone acetyltransferase	-21.74	0.0574
AT3G08040	*FERRIC REDUCTASE DEFECTIVE 3* (*FRD3*)	1.71	0.0724
AT4G13505	Antisense long non-coding RNA	4.04	0.0726
AT5G66440	tRNA-methyltransferase non-catalytic subunit	1.11	0.0726
AT5G22500	*FATTY ACID REDUCTASE 1* (*FAR1*)	-3.02	0.0966
AT1G65481	Transmembrane protein	-5.79	0.0966
AT5G20830	*SUCROSE SYNTHASE 1* (*SUS1*)	1.15	0.0966
AT3G61490	*POLYGALACTURONASE CLADE F9* (*PGF9*)	0.62	0.0993
AT3G19550	Glutamate racemase	1.59	0.0993

5.3.3.5.2 *C. higginsianum* genes differentially expressed between *ΔChEC153-2* and *ΔChKu80* strains during infection

To further probe the role of ChEC153.3 in infection, I also sought to look for transcriptional differences between *ΔChEC153-2* and the *ΔChKu80* background strain during their infection of Arabidopsis. I considered this to be particularly pertinent in light of the ambiguity regarding the secretion status of ChEC153.3 (**Table 5.1**). If ChEC153.3 is not a secreted protein, it may play a role during infection in the pathogen itself. Transcriptional analysis of the fungal *ChEC153.3*-knockout strain may thereby inform on the role of ChEC153.3 irrespective of its secretion. Due to the small variance seen between strains (**Figure 5.16**), and relatively low read counts seen for *C. higginsianum* genes, I defined differential expression as a $\log_2$ fold change between strains of > 0.5 or < -0.5, and an adjusted Wald test p-value of < 0.1. I first validated that fungal reads were present in the infected tissue; between *ΔChKu80*- and mock-inoculated samples, 3380 and 3386 *C. higginsianum* genes were differentially expressed at 24- and 36-hpi, respectively. Next, I compared *C. higginsianum* genes differentially expressed in *ΔChEC153-2* relative to *ΔChKu80* during infection of Arabidopsis (**Table 5.3**). At 24 hpi, four *C. higginsianum* genes are upregulated and eight are downregulated in *ΔChEC153-2*, while at 36 hpi, one *C. higginsianum* gene is upregulated and six are downregulated. As no reads in *ΔChKu80* mapped to *CH35J_009900* due to the low coverage, *ChEC153.3* is not listed as differentially expressed between strains.

Table 5.3: ***C. higginsianum* genes differentially expressed between strains *ΔChEC153-2* and *ΔChKu80* at 24- or 36-hours post Arabidopsis inoculation.** Gene IDs and names/descriptions (Ensembl fungi) of genes differentially expressed in *ΔChEC153-2* are listed, alongside their $\log_2$ fold change ($\log_2$ FC) relative to background strain *ΔChKu80*, and respective adjusted Wald test p-values (p-adj). Differential expression defined as $\log_2$ FC > 0.5 or <-0.5, p-adj < 0.1. Genes with higher expression in samples inoculated with *ΔChEC153-2* relative to *ΔChKu80* are highlighted in pink, while those with lower expression in *ΔChEC153-2*-inoculated samples are shown in green. Genes for which p-adj < 0.05 are highlighted in yellow.

C. higginsianum* genes differentially expressed at 24 hpi with *ΔChEC153-2* vs. *ΔChKu80			
Gene ID	**Gene name / description**	**$\log_2$ FC**	**p-adj**
CH35J_006543	*FUSARIN EFFLUX PUMP 6* (*FUS6*)	-5.04	0.0543
CH35J_006545	Putative zinc-type alcohol dehydrogenase-like protein AdhB	-4.54	0.0579
CH35J_000842	Dihydroxyacetone kinase 1	-4.36	0.0579
CH35J_002539	Kelch repeat-containing protein 3	-4.69	0.0579
CH35J_004760	Lipase 1	-3.16	0.0650
CH35J_010323	*SUCROSE NON-FERMENTABLE 7* (*SNF7*)	3.10	0.0678
CH35J_010175	60S acidic ribosomal protein P1	4.99	0.0678

CH35J_012265	Fasciclin-like arabinogalactan protein	-3.94		0.0825
CH35J_003984	Putative amino-acid permease P7G5.06	4.58		0.0825
CH35J_006544	Hypothetical protein	-3.30		0.0825
CH35J_006578	Hypothetical protein	-4.26		0.0825
CH35J_011758	Putative peroxiredoxin pmp20	3.90		0.0988
C. higginsianum* genes differentially expressed at 36 hpi with *ΔChEC153-2* vs. *ΔChKu80				
Gene ID	**Gene name / description**	**Log_2 FC**		**p-adj**
CH35J_008101	Periplasmic β-glucosidase	-3.80		0.00000937
CH35J_004709	Lactose permease	-3.46		0.000810
CH35J_011389	General α-glucoside permease	-2.71		0.00292
CH35J_010947	Hypothetical protein	4.87		0.00292
CH35J_006578	Hypothetical protein	-4.74		0.0344
CH35J_006545	Putative zinc-type alcohol dehydrogenase-like protein AdhB	-4.59		0.0486
CH35J_010346	Adenosine deaminase CECR1-A	-4.63		0.0880

5.3.3.6 RNA-Seq reads from strain *ΔChEC111* map to *ChEC153.3*

Finally, I sought to evaluate how well the reads acquired from the RNA-Seq data align to the genomic region corresponding to *ChEC153.3* due to the low coverage. No reads from the *ΔChKu80* samples were found to map to the *CH35J_009900* gene model. While the data presented here are analyses of only *ΔChEC153-2-*, *ΔChKu80-*, and mock-inoculated Arabidopsis samples, inoculation with another *C. higginsianum* putative effector knockout strain, *ΔChEC111*, was carried out and subject to RNA sequencing in the same experiment. The *ΔChEC111* strain was generated by Dr Andrew Breakspear from the *ΔChKu80* background strain using the same strategy as I used to generate *ΔChEC153-2* (Section 4.2.3).

Samples inoculated with strain *ΔChEC111* contained some fungal reads that aligned to the *CH35J_009900* gene model. Due to the relatively low number of reads seen in these samples, all reads from the three biological replicates of Col-0 36 hpi with *ΔChEC111* were pooled for alignment to the reference fungal genome. Figure 5.17 shows these pooled reads mapped to the Tsushima *et al. C. higginsianum* genome assembly (Tsushima et al., 2019a), with one replicate of the O'Connell *et al.* RNA-Seq data (O'Connell et al., 2012) shown for reference, alongside the gene models referred to in this chapter. This alignment shows that reads for the entire region of the genome encoding *ChEC153.3* are present in the *ΔChEC111* samples, and the putative intron/exon boundaries suggested by these data match with those seen suggested by the publicly available RNA-Seq data (O'Connell et al., 2012).

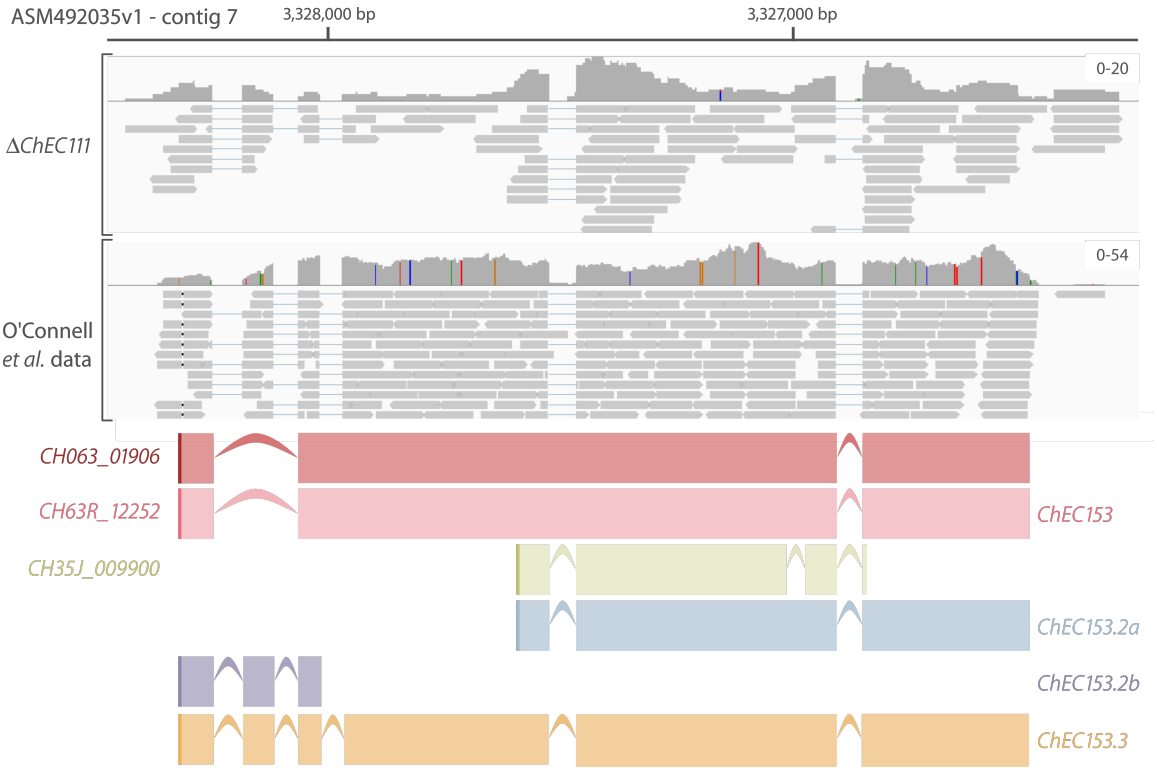


Figure 5.17: RNA-Seq reads from *C. higginsianum* strain *ΔChEC111* map to *ChEC153.3* in the Tsushima *et al.* genome (ASM492035v1). *ΔChEC111* data represent reads pooled from three biological replicates. A single biological replicate of the O'Connell *et al.* data is shown as a reference (O'Connell et al., 2012). Coverage and alignment tracks are shown, with individual reads coloured grey, and SNPs relative to the genome shown as coloured bars in the coverage track. The alignment track is cropped to show a read depth of 13 reads for each dataset, with scales for the coverage tracks to the right-hand side of each alignment. Data are visualised in IGV.

5.4 DISCUSSION

5.4.1 The true gene model for *ChEC153*, *ChEC153.3*, differs from both published gene models

Through the work in the Section 5.3.2, I revealed that neither published gene model for gene *ChEC153, CH63R_12252* (*ChEC153.1*) nor *CH35J_009900,* is definitively correct. Instead, I present a new gene model, sequence-verified from cDNA, referred to here as *ChEC153.3*. It is possible that a number of splice variants of this gene exist (for example, gene isoforms *ChEC153.2a*, *ChEC153.2b,* and *ChEC153.3*). As no expression of *35S::ChEC153.2a-eGFP* nor *35S::ChEC153.2a-mCherry* was observable during transient expression in *N. benthamiana* (Section 5.3.2.3), it is possible that the protein was unstable and may not exist in nature. The publicly available *C. higginsianum* RNA-Seq data (O'Connell et al., 2012, Tsushima et al., 2019a), and the RNA-Seq data acquired in this project (Figure 5.17) agree strongly with my analysis of the coding sequence of *ChEC153* and the new gene model, *ChEC153.3*, that I establish here. RNA-Seq data can indicate splice variants where clear changes in coverage across the putative coding region are seen. Given that neither the O'Connell *et al.* RNA-Seq data, nor the RNA-Seq data generated in this project (Figure 5.17; O'Connell et al., 2012), suggest clear splice variants, I hypothesise the true and predominant (if not only) gene isoform present to be *ChEC153.3*. However, further analyses could be carried out to elucidate potential gene isoforms and their relative abundances (Cmero et al., 2021, Harvey and Cheng, 2016).

5.4.1.1 Is ChEC153.3 likely to be secreted from the fungus?

The prediction of a signal peptide for secretion from the pathogen for ChEC153.3 is ambiguous (Table 5.1). Both SignalP 3.0 and TargetP 2.0 predict ChEC153.3 to possess a signal peptide (Bendtsen et al., 2004, Armenteros et al., 2019a). Some *in silico* tools, such as DeepTMHMM (Hallgren et al., 2022) and Phobius (Madeira et al., 2022), predict ChEC153.3 to have a transmembrane domain near the N-terminus (residues 19 - 29, or 12 - 33 with either tool, respectively). As signal peptides and transmembrane domains are both predicted based on the presence of sequences enriched in hydrophobic amino acids, there can be difficulties in discriminating between the two. SignalP 4.1 seeks to distinguish between N-terminal transmembrane domains and signal peptides, and does not predict the sequence of ChEC153.3 to have a signal peptide unless it is specified that the sequence does not possess a transmembrane domain. Thus, SignalP 4.1 appears to predict ChEC153.3 to contain an N-terminal transmembrane domain, rather than signal peptide (Petersen et al., 2011).

Secretion of fungal apoplastic proteins, is predominantly believed to be via the conventional secretory pathway whereby proteins with signal peptides are targeted via the endoplasmic reticulum to the Golgi apparatus, at which they are packaged into secretory vesicles for trafficking to the plasma membrane (Delic et al., 2013). However, as mentioned in Section 1.2.2, non-conventional secretion methods, particularly for cytoplasmic effectors, are increasingly being identified (Balmer and Faso, 2021). Kim and colleagues reported around 48% of secreted *M. oryzae* proteins to lack canonical signal peptides for the conventional secretion pathway (Kim et al., 2013). Indeed, effector proteins without canonical signal peptides have been characterised (Chen et al., 2023 Liu et al., 2014). EVs have been suggested to play a role in effector secretion from pathogens such as *Magnaporthe* (Oliveira-Garcia et al., 2023), and the presence of EVs has also recently been observed in *C. higginsianum* (Rutter et al., 2022). This represents a growing area of effector biology research and supports increasing evidence that the requirement for effector proteins to have conventionally predicted signal peptides is not absolute. To complement the *in silico* annotation of signal peptides, it would therefore be informative to experimentally assay for secretion and translocation of ChEC153.3. Approaches for assaying the secretion and translocation of effectors are discussed in Section 6.2.1.

Further, given the successful establishment and cloning of the *ChEC153.3* coding sequence, it may be informative to transiently express this gene in *N. benthamiana* to determine its subcellular localisation. To do this, one could assume the signal peptide sequence to be that predicted by SignalP 3.0 or TargetP 2.0 and express the mature sequence as both of these N-terminal truncations to establish whether or not ChEC153.3 could localise to the chloroplast, if translocated into the host.

5.4.2 What are the effects of *ChEC153.3* in the host or fungus?

5.4.2.1 Host gene expression differences in *ΔChEC153-2* and *ΔChKu80* infections

The specific expression of *ChEC153.3* by *C. higginsianum* during the early stages of infection strongly suggests a role in pathogenicity (O'Connell et al., 2012, Dallery et al., 2017). From the dual RNA-Seq data presented here, I identified a number of genes in either the pathogen or host of the *C. higginsianum*-Arabidopsis pathosystem that are differentially expressed during infection of Arabidopsis with *ΔChEC153-2* relative to *ΔChKu80* (Table 5.2 and Table 5.3).

DARK INDUCIBLE 11 (DIN11)

At 36 hours post inoculation, the Arabidopsis gene showing the largest difference in expression between inoculations with either *C. higginsianum* strain is *DARK INDUCIBLE 11* (*DIN11*). *DIN11* is more highly expressed (with a $\log_2$ fold change of 20.85) at 36 hours post inoculation with *ΔChEC153-2* compared to infection with *ΔChKu80*. Expression of *DIN11* is significantly lower during infection with *ΔChKu80* compared to the mock-inoculated control at 36 hpi ($\log_2$ fold change -18.41; **Appendix 3 Tab. A**). There is no significant difference in the expression of *DIN11* when *ΔChEC153-2*- and mock-inoculated Arabidopsis samples are compared. Therefore, it is possible that *ChEC153.3* somehow downregulates *DIN11* during infection, whether directly or indirectly. Expression of dark inducible genes is known to be senescence-related and repressed by sugars (Fujiki et al., 2001), so the differential expression of *DIN11* in infection in a *ChEC153.3*-dependent manner could support the association of ChEC153.3 with a role in carbohydrate targeting. If ChEC153.3 causes enhanced accumulation of sugars in host tissues in an infection context, this would correlate with the observed lower expression of *DIN11* in the presence of ChEC153.3 compared to in its absence. Interestingly, silencing of *DIN11* in Arabidopsis has been associated with reduced *Tobacco rattle virus* (TRV) accumulation, with TRV usually resulting in an increase in *DIN11* expression, pointing to a potential role for *DIN11* expression in susceptibility to some viruses (Fernández-Calvino et al., 2016). Further, *DIN11* expression appears to be induced following infection of Arabidopsis with the fungal necrotroph *A. brassicae* (Johnson et al., 2014). While this suggests a link between *DIN11* and infection more broadly, the specific implications for the function of ChEC153.3 are unclear.

Conserved peptide upstream open reading frames (CPuORFs)

Other Arabidopsis genes observed to be differentially expressed during infection with *ΔChEC153-2* relative to *ΔChKu80* include several *CPuORF* genes. CPuORFs are generally associated with regulatory genes, such as transcription factors (Jorgensen and Dorantes-Acosta, 2012). CPuORF48 is less highly expressed in Arabidopsis at both 24- and 36-hours post inoculation with the *ΔChEC153-2* strain relative to *ΔChKu80.* CPuORF48 is upstream of a methyltransferase gene (*AT5G01710*), for which it may play a regulatory role. *CPuORF31* is less highly expressed in *ΔChEC153-2* than *ΔChKu80*-infected Arabidopsis at 36 hpi and is involved in the regulation of PHOSPHOETHANOLAMINE METHYLTRANSFERASE 2, PMT2 (*AT1G48600*). Finally, *CPuORF2* is more highly expressed in infection with *ΔChEC153-2* than *ΔChKu80* at 36 hpi. *CPuORF2* regulates *bZIP11* (*AT4G34590.1*), which is a transcription factor influencing amino acid metabolism (Hanson et al., 2008). *CPuORF2* is known to be involved in inhibition of *bZIP11* translation through ribosome stalling, with this translation inhibition promoted by sucrose (Yamashita et al., 2017). The lower expression of *CPuORF2* seen in the presence of *ChEC153.3* may correlate with an increase in

bZIP11, and points to a potential link with host carbohydrate metabolism, amino acid metabolism and potential misregulation of sucrose signalling.

GERMIN-LIKE PROTEIN 2A and 5A (GLP2A/GLP5A)

Two host germin-like protein genes, *GLP2A* and *GLP5A*, are more highly expressed 24-hours after inoculation with *ΔChEC153-2* than with *ΔChKu80*. Plant germins are glycoproteins known to play diverse roles in development as well as in responding to a number of stresses (Karlik, 2021). Indeed, *GERMIN-LIKE PROTEIN* genes clustered in the genomes of cereal species have been linked to fungal defence responses (Breen and Bellgard, 2010). This could suggest a role for *ChEC153.3* in indirectly reducing host defence responses.

DA1 RELATED PROTEIN 3 (DAR3)

DAR3 is more highly expressed in infection with *ΔChEC153-2* than with *ΔChKu80. DAR3* has been linked to host responses to the oomycete *A. candida* by Gu and colleagues (Gu et al., 2023). *A. candida* infection leads to a decrease in DAR3, which through the activation of DA1 leads to host cell endoreduplication and increased pathogen virulence. The authors show that overexpression of Arabidopsis *DAR3* confers enhanced resistance to *A. candida*, while mutation of *DAR3* increases susceptibility to *A. candida*, as well as *H. arabidopsidis* and *P. syringae* pv. *tomato* DC3000. They suggest that DAR3 is a putative effector target, guarded by the CHS3/DAR4 sensor NLR which contains an integrated DAR3 decoy (Gu et al., 2023). Thus, there are clear links between *DAR3* and host defence responses, and the reduction of *DAR3* expression in the presence of *ChEC153.3* could point to disruption of the host defence response to facilitate pathogen growth.

IRON HOMEOSTASIS GENES

Both *NICOTIANAMINE SYNTHASE 3* (*NAS3*) and *FERRIC REDUCTASE DEFECTIVE 3* (*FRD3*) are more highly expressed in Arabidopsis 36 hours post infection with *ΔChEC153-2* relative to *ΔChKu80*. These genes both have roles in iron homeostasis, and so could point to the pathogen manipulating the host in order to acquire iron, or else an aspect of the plant immune response. In maize, susceptibility to *C. graminicola* is conferred by host iron deficiency, potentially due to the requirement of the host to recruit iron to the infection site as part of the production of the oxidative burst (Ye et al., 2014). Therefore, it is possible that *ChEC153.3* perturbs normal iron homeostasis in order to facilitate infection.

SUCROSE SYNTHASE 1 (SUS1)

Finally, *SUCROSE SYNTHASE 1* (*SUS1*), which is involved in both the synthesis and cleavage of sucrose, is slightly more highly expressed in Arabidopsis during infection with the *ΔChEC153-2* strain relative to *ΔChKu80*, further supporting a potential role for ChEC153.3 in manipulating host carbohydrates in some way. There are six sucrose synthase genes in Arabidopsis, with some functional redundancy observed (Baroja-Fernández et al., 2012). *SUS1* appears to be expressed predominantly in the Arabidopsis leaf vasculature (Bieniawska et al., 2007), localising specifically in companion cells of the phloem with a potential role in phloem loading (Yao et al., 2020). Differential expression of *AtSUS1* is seen in response to certain abiotic stresses (Déjardin et al., 1999, Baud et al., 2004).

In summary, a number of host genes expressed at different levels during infection with *C. higginsianum* strains with and without *ChEC153.3* have roles in defence and/or carbohydrate metabolism. This supports the hypothesis that *ChEC153.3* has a carbohydrate-related role in infection, though direct mechanisms are yet to be elucidated. Furthermore, these data have identified genes/processes putatively targeted by ChEC153.3, whether directly or indirectly, to promote infection.

5.4.2.2 Pathogen gene expression differences in *ΔChEC153.3* versus *ΔChKu80*

I identified a number of *C. higginsianum* genes with different levels of expression during infection of Arabidopsis by *ΔChEC153-2* and *ΔChKu80*. While little if anything is directly published about these *C. higginsianum* genes, inferences on their roles can be made from their annotations. The *C. higginsianum* genes most strongly differentially expressed between the *ΔChEC153-2* and *ΔChKu80* strain infections at 36 hpi include three genes with carbohydrate-related functions. The most significantly differentially expressed of these, *CH35J_008101* (corresponding to *CH63R_08058* in the Zampounis *et al.* genome assembly), is annotated as encoding a periplasmic β-glucosidase, here less highly expressed in the *ΔChEC153-2* strain. The degradation of β-glucans in infection may reflect targeting of the host cell wall cellulose. The action of β-glucosidases on their cellulose-derived cellobiose substrates releases glucose monomers, which may be taken up by the pathogen as a source of carbon. Secondly, a lactose permease gene, *CH35J_004709* (*CH63R_03333*), is also downregulated significantly at the same time post inoculation with *ΔChEC153-2* relative to *ΔChKu80*. Lactose permeases are symporters for uptake of the disaccharide lactose, for example characterised in the saprophytic filamentous fungus *Aspergillus nidulans* (Fekete et al., 2012). Finally, *CH35J_011389* (*CH63R_06790*), a gene annotated as encoding a general α-glucoside permease, is also downregulated in *ΔChEC153-2* during infection. This permease may play a role

in carbon uptake by the fungus, with α-glucoside permeases being capable of transporting a number of α-glucosides such as maltose, trehalose and sucrose (Han et al., 1995). Notably, maltose is one of the major degradation products of starch. In IMI 349063, each of these genes shows expression throughout infection, with highest expression during the necrotrophic phase of infection (O'Connell et al., 2012, Dallery et al., 2017). The lower expression of these fungal genes in the absence of *ChEC153.3* could point to the importance of *ChEC153.3* for carbon acquisition by the pathogen.

5.4.2.3 Dual RNA-Seq enables pathosystem analysis, but challenges posed by variable infection rates may confer limitations to the data

A relatively small number of host genes were identified as being differentially expressed during Arabidopsis infection with *C. higginsianum* (*ΔChKu80*) compared to the uninfected control (18 and 42 genes at 24- and 36-hours, Appendix 3 Tab. A). Given the small scale of host transcriptional responses to infection, and the low number of reads mapping to the *C. higginsianum* genome in this dual RNA-Seq experiment (Figure 5.13), it appears that the infection stage sampled may not have been optimal to probe the biotrophic interaction between host and pathogen. It is possible that the time-points of 24- and 36-hpi selected for this experiment did not give sufficient time for *C. higginsianum* to progress into the biotrophic phase of infection in this instance. While the verified expression of *ChEC6* in infected samples at both sampling time-points (Figure 5.12) supported their use as appropriate samples for the biotrophic phase, it may have been informative to carry out a more in-depth histological analysis of the pathosystem under our conditions to ensure the infection had sufficiently progressed into biotrophy. Through the infection assays that I carried out in Chapter 4, I found rates of infection to be fairly variable within and between experiments set up using identical conditions. For the single-cell approach employed by Tang and colleagues to analyse Arabidopsis cell transcriptional responses to *C. higginsianum* infection, they selected 24- and 40-hpi, which they identified to be the late biotrophic stage based on microscopic analysis of the infection. Though these time-points are very similar to those chosen in this work, the authors state that they observed uneven infection, with a fraction of the host epidermal cells directly contacting fungal cells (Tang et al., 2023).

While the *C. higginsianum* inoculum for these infections was fairly high (approximately 60,000 conidia per leaf), as *C. higginsianum* biotrophy is restricted to the primary infected cell, and much of the leaf surface may not have contacted conidia directly due to the drop-inoculation method used, it is possible that the relative proportion of uninfected leaf tissue may have diluted transcriptional responses seen in infected cells. Some published approaches use epidermal peels of infected tissue to analyse the *C. higginsianum* transcriptome (O'Connell et al., 2012, Kleemann

et al., 2012), and while labour-intensive to generate sufficient material for RNA-Seq analysis, these approaches would likely result in a much larger fraction of the acquired reads mapping to the fungal genome. However, whether sufficient Arabidopsis material would acquired with this method to probe host transcriptional responses remains to be determined.

5.4.3 Chapter conclusions

In this chapter, I identified inconsistencies between the gene annotations corresponding to *ChEC153* in different published genome assemblies for *C. higginsianum*. I was able to elucidate the true gene model of *ChEC153*, referred to here as *ChEC153.3*, which differs from both published gene models. Host transcriptional responses to infection in the presence and absence of *ChEC153.3* point to the relevance of this putative effector gene in both defence and aspects of host carbohydrate metabolism. Meanwhile, analyses of fungal transcripts reveal differences in expression of genes likely relating to carbon acquisition in the presence and absence of *ChEC153.3*. These results support a putative role for *ChEC153.3* in carbohydrate-related processes during infection.

CHAPTER 6

General discussion

This chapter aims to discuss the main results of the thesis, their broader implications, and the future questions and proposed experiments that these results raise. In **Chapter 3** I investigated the localisation of putative effectors ChEC153.1 and HaRxL94b, and the impact of ChEC153.1 on host starch. I found *ChEC153.1* expression to induce the formation of small starch granules in *N. benthamiana*. In **Chapter 4**, I probed the impact of *ChEC153.1* on infection, revealing that expression of the putative effector contributes to host susceptibility. I also identified likely interactors of the putative *C. higginsianum* effector in the Arabidopsis host, including a chloroplast-localised protein with predicted RNA-binding capabilities, referred to here as CRBIC, which was specifically recruited to puncta in the presence of ChEC153.1. In **Chapter 5**, I first sought to probe the gene model of gene *ChEC153*, elucidating the true gene model, *ChEC153.3*, to be different to those annotated in published genome assemblies. For clarity, the putative effector sequence previously referred to as *ChEC153* in **Chapters 3** and **4** of this thesis, inferred from published gene model *CH63R_12252* (Zampounis et al., 2016), is renamed as *ChEC153.1* here. Finally, I aimed to evaluate differences in transcriptional responses during infection in both the host and pathogen in the presence and absence of *ChEC153.3*. I identified a number of genes that are expressed at different levels in the pathogen or host during infection of Arabidopsis with *C. higginsianum ChEC153.3*-knockout strain *ΔChEC153-2* or the background strain *ΔChKu80.* The differential expression of these genes suggests a role for ChEC153.3 in infection and/or carbohydrate targeting.

6.1 A role for starch granule initiation in pathogenicity

The co-localisation of putative effectors ChEC153.1 and HaRxL94b with key starch granule initiation protein MRC suggests a role for starch granule initiation in pathogenicity. ChEC153.1 is able to induce the formation of small starch granules in *N. benthamiana*, and hosts with defects in starch granule initiation have different susceptibility to infection. In addition to the putative effectors focussed on here, this work therefore also highlights other, broader, links between starch granule initiation and infection.

6.1.1 ChEC153.1 perturbs starch granule initiation and may make host carbon more accessible to the pathogen

I found that *ChEC153.1* expression induces the formation of small starch granules (Section 3.3.7, discussed in Section 3.4.3). The fate of these small starch granules in infection remains to be determined. It would be informative to see whether GBSS puncta are induced during infection of Arabidopsis with *C. higginsianum*. While transient expression of *ChEC153.1* in *N. benthamiana* was sufficient to form GBSS-RFP puncta, whether the levels of the putative effector present during infection would be sufficient to initiate these small starch granules could be determined by infecting Arabidopsis lines expressing fluorescently-tagged GBSS with *C. higginsianum*. The use of a fluorescent strain of *C. higginsianum* may facilitate the identification of infection structures to guide the search for induced GBSS puncta. Should GBSS puncta be observed in this infection context, *ChEC153*-knockout *C. higginsianum* strains could be used to help to deduce whether a theoretical induction of small starch granules in Arabidopsis by *C. higginsianum* is specifically and exclusively caused by ChEC153.1/ChEC153.3. Here, I postulate that smaller starch granules may be more amenable to degradation, and that this could provide a more accessible source of carbon to the pathogen – expression of *ChEC153.1* in the host correlating with increased susceptibility. In addition to size, other factors also affect the degradability of starch granules. For example, the aspect ratio/circularity of granules is also believed to play a role in their degradability and other physiochemical properties, as well as the internal granule structure, and the amylose/amylopectin ratio (Chen et al., 2021a), all of which could link to the accessibility of carbon to invading pathogens.

The majority of starch granules of *ChEC153.1*-expressing Arabidopsis lines purified in this work were morphologically similar to those of the wild-type plants, as discussed in Section 3.4.2. Particle size analysis suggested that the majority of starch granules from *ChEC153.1*-expressing Arabidopsis lines are slightly larger than those of the wild-type, and microscopic analysis of the starch revealed no clear differences in granule shape (Section 3.3.11). However, it may be interesting to probe for more subtle starch granule phenotypes such as those mentioned above. In addition to confirming the presence of small starch granules (see Section 3.4.3), the "normal" size starch granules of *ChEC153.1*-expressing Arabidopsis lines could be compared to those of the wild-type in terms of their degradability using starch degradation assays.

6.1.2 The involvement of starch granule initiation in disease susceptibility is highlighted by granule initiation mutants

While the hyper-susceptibility of some starch mutants to *C. higginsianum* has been described in the literature (Engelsdorf et al., 2013, Engelsdorf et al., 2017), the impact of starch granule

initiation on host resistance had not previously been explored. Here, I revealed an enhanced resistance to infection by *C. higginsianum* in the absence of key granule initiation protein MRC (Section 4.3.3). A number of experiments would complement this observation, including testing whether complementation lines revert the enhanced resistance phenotype, and whether overexpression of MRC in the host would confer a more susceptible phenotype. There are a number of other starch granule initiation genes mutant lines which could be tested for similar resistance phenotypes, such as *ss4* mutants (Roldan et al., 2007), though these may be confounded by more severe growth and granule phenotypes.

One hypothesis to explain the enhanced resistance of *mrc* lines is that the starch granules present in this line will have a lower surface area to volume ratio than the wild-type, and that this could impede starch degradation in an infection context. In order to test this hypothesis, a number of other Arabidopsis lines could be employed, for example PTST2 mutant and overexpression lines which have larger and smaller starch granules compared to the wild-type, respectively. These lines, similarly to *mrc*, have similar total starch content to the wild-type as a comparable amount of substrate is differentially distributed to larger but fewer, or smaller but more numerous, starch granules per chloroplast. Therefore, for example, *ptst2* mutants may be more resistant to infection, similarly to the *mrc* mutant tested here – this would support a correlation between granule size and resistance to infection. These lines are particularly appropriate for this experiment as the PTST2 protein in involved at the same specific stage of starch granule initiation as MRC, with which it interacts (Seung et al., 2018).

In addition to the absolute size of starch granules, other factors, such as subtle changes in the granule structure, are known to contribute to their degradability, as mentioned in Section 6.1.1. Probing for susceptibility/resistance phenotypes in lines with known alterations in starch granule structure or composition, rather than size, may therefore also present an interesting area for further investigation to inform on whether degradability of host starch can impact pathogen virulence.

Further, it may be informative to assay for enhanced resistance in starch granule initiation mutants challenged by other pathogens to see how conserved this phenotype is. For example, *H. arabidopsidis* may be a logical choice given the targeting of MRC by HaRxL94b, and the biotrophic nature of the pathogen. As *H. arabidopsidis* is a biotrophic pathogen, the host will be kept alive for an extended period relative to that seen in *C. higginsianum* infection where the pathogen transitions to necrotrophy. It is possible that in extended biotrophy, any benefit of targeting starch granules would be exacerbated. Further, it may be interesting to see whether the *mrc* resistance phenotype is lost in infection with necrotrophic pathogens. Whether a change in starch granule surface area would represent a significantly different carbon source for pathogens during

necrotrophy is not clear. In necrotrophy, where carbon sources become abundant to the pathogen with the degradation of entire plant cells, it is possible that the increased accessibility of carbon in the form of starch granules with high surface area to volume ratios would be too small to significantly alter virulence.

6.1.3 Starch is likely to act as a carbon source for *C. higginsianum*

I hypothesise that *C. higginsianum* is capable of using starch as a carbon source. While I have not found direct evidence in the literature of *C. higginsianum* degrading starch *ex planta*, this would represent a relatively simple experiment to carry out via growth assays on various media. The related *Colletotrichum* species *C. gloeosporioides* is able to degrade starch in *ex planta* cultures, mainly via the action of a glucoamylase (Krause et al., 1991), and *C. fructicola* (Velho et al., 2018), *C. fioriniae, C. noveboracense*, and *C. chrysophilum* also exhibit amylolytic activity (Khodadadi et al., 2020). Further, *C. higginsianum* appears to encode a glucoamylase protein, and several α-amylases, predicted to be secreted from the fungus. The expression of this putative *C. higginsianum* glucoamylase gene (*CH63R_07722*) increases throughout infection, being particularly upregulated in the biotrophic and necrotrophic phases (O'Connell et al., 2012, Dallery et al., 2017). This suggests that *C. higginsianum* can degrade starch derived from the host. Glucoamylases are specific exo-acting α-glucosidases that target non-reducing ends of α-glucans. While non-secreted fungal glucoamylases may function in glycogen degradation, the predicted secretion of a glucoamylase points to potential in targeting host starch. Therefore, it is likely that changes to starch granule initiation would alter the accessibility of carbon to the pathogen.

6.1.4 Granule size may have implications in pathogen targeting of storage starches

The targeting of starch storage organs by some pathogens represents an alternate avenue by which starch, and granule size/degradability, may play a role in pathogenicity. For example, where pathogens infect potato tubers and cereal grains, the major source of carbon available will be in the form of starch. Whether these pathogens would be able to manipulate the host starch processes prior to hydrolysing the starch prevalent in these storage organs is not clear. As storage starches accumulate over much longer timescales than the transitory starches of leaves, their theoretical perturbation may look different – for example, the relevance of impacting granule initiation is not obvious. Certainly, in the cases of post-harvest pathogens and necrotrophs such as *Fusarium graminearum*, it seems unlikely that the pathogen would manipulate host starch processes. However, pathogens such as the hemibiotrophic *M. oryzae* pathotype *Triticum* (wheat blast) may still benefit from pre-necrotrophy targeting of host starch in some way. While much

research in the starch field is focussed on probing granule structure for potential benefits in terms of nutrition or suitability for industrial applications (Chen et al., 2021a), the work presented here suggests that susceptibility of plants may also be impacted in these pursuits. It would be particularly apposite if less degradable starch, for example high-amylose "resistant" starch, which is known to have nutritional benefits when used in foods (Li et al., 2023), correlated with an increase in plant disease resistance. Additional aspects of granule structure may also be relevant, for example, starch granules of some species, such as those of the cereal endosperm, have relatively large/numerous surface pores. Granules with more or larger pores are more amenable to degradation, likely due to the increased surface area that can be accessed by degradation enzymes (Fannon et al., 1992, Sujka and Jamroz, 2010). Thus, there are potential implications for reduced granule degradability in resistance of crop species.

6.1.5 Potential impacts of starch targeting on host defence

In addition to the hypothesis that starch granule initiation impacts pathogen virulence, for example by altering the accessibility of organic carbon, granule initiation may also impact host defence. As mentioned in Section 1.8, sugars form important players in host defence responses as signalling molecules. It is likely that abnormal starch granule initiation would correlate with perturbed sugar status of the infected cell. It is possible that perturbations in the chloroplast caused by aberrant granule initiation, such as altered sugar signalling, could have downstream impacts on phytohormone production. Therefore, the initiation of small starch granules by ChEC153.1 could represent a mechanism by which to indirectly perturb host defence responses by altering carbohydrate distribution between sugars and starch. Targeting of granule initiation proteins by pathogen effectors could thereby have wide-ranging effects on the capacity of the host to mount an appropriate defence response. It is possible that ChEC153.1-induced formation of small starch granules would correlate with a more rapid degradation of starch and release of sugar shortly after the end of the photoperiod. My RNA-Seq analysis of the *C. higginsianum*-Arabidopsis pathosystem comparing transcriptional responses during infection with a *ChEC153*-knockout strain to the background strain may support this hypothesis. Differential expression of a number of host genes is seen, including sugar-responsive genes such as *DIN11* and *CPuORF2*, as well as other defence-related genes.

6.2 Verification of ChEC153.3 and HaRxL94b as effectors

6.2.1 Experimental confirmation of putative effector secretion and translocation

To verify that ChEC153.3 and HaRxL94b are bona fide effectors, further work remains to be carried out to confirm the secretion of these proteins from the pathogen and their translocation into the host. The prediction of signal peptides of these putative effectors has some ambiguity. As highlighted in Section 5.4.1.1, a number of cytoplasmic effectors are secreted by nonconventional secretion pathways, and may not require canonical signal peptide sequences (Liu et al., 2014, Chen et al., 2023 Boevink et al., 2016). Given the limitations and ambiguity of *in silico* signal peptide prediction, it would be pertinent to experimentally evidence secretion of ChEC153.3 and HaRxL94b from their respective hosts. Approaches to directly evidence the secretion of a putative effector from the pathogen may not be straightforward, particularly in the case of HaRxL94b given the genetic intractability of the *H. arabidopsidis* pathogen.

Secretion assays in yeast have been used to evidence secretion of a number of putative effectors. Jacobs and colleagues describe a yeast signal sequence trap system to screen for secreted proteins (Jacobs et al., 1997). When yeast is grown on sucrose or raffinose media, it must secrete invertase in order to hydrolyse the di- or tri-saccharides in the media. By cloning potential signal peptides within an yeast expression cassette as 5' fusions to the invertase gene lacking its native secretion signal, the growth of yeast can be used to indicate the functionality of the putative signal peptide (Jacobs et al., 1997). This assay has been employed to identify secreted effector proteins (Lee and Rose, 2012, Kuppireddy et al., 2017), and has the advantage of not requiring that the pathogen of interest be genetically transformed, and therefore may be applicable for confirming the secretion of HaRxL94b.

For some genetically tractable pathogens, such as *M. oryzae*, observation of fluorescently-tagged effector proteins moving from the pathogen to the host has been used to verify translocation (Khang et al., 2010). Direct translocation of *Colletotrichum* putative effectors in this way has proven technically challenging (Irieda et al., 2014, Kleemann et al., 2012). While Kleeman and colleagues have demonstrated secretion of *C. higginsianum* effectors in the host apoplast, they did not observe translocation into the cytoplasm for any of the effectors they tested (Kleemann et al., 2012). Fusion of a fluorophore to a putative effector protein may preclude its secretion based on its increased size. To counter this, split-GFP translocation assays have been applied to demonstrate the secretion of bacterial effectors (Park et al., 2017). The split-GFP approach enables a putative effector to be tagged with only the 11^{th} β-strand of GFP, with the remaining 1^{st}-10^{th} β-strands being expressed in the host. Successful translocation of the putative effector with this minimal size

alteration reconstitutes the GFP fluorescence *in planta* (Park et al., 2017). A similar approach could involve fusion of the putative signal peptide sequence to a fluorescent protein for microscopic detection. For example, generating a *C. higginsianum* strain expressing, a *ChEC153.3*$_{1\text{-}23(SP)}$*-eGFP* construct under a constitutive promoter for high expression may allow the visualisation of eGFP in the host following infection, supporting the prediction of residues 1-23 as the signal peptide. As a positive control for such an assay, the signal peptide of an effector known to be secreted could be fused to eGFP for analysis *in planta*. This control would be necessary to establish whether eGFP itself is capable of being secreted with a known signal peptide, or whether specific properties of the fluorophore would preclude secretion. Targeting of the eGFP to a given part of the host cell may also facilitate its visualisation, such as by inclusion of a nuclear localisation signal or chloroplast transit peptide.

In addition to observing translocation of fluorescently-tagged effectors to evidence their secretion, immunoelectron microscopy has also been used to evidence effector translocation, as in the case of the translocation of a *U. maydis* effector into host cytoplasm during infection (Djamei et al., 2011). There are a number of biochemical approaches which could be used to assay for translocation. For example, biotinylation of putative effectors tagged with a short "Avitag" peptide in the host has been used as an indication of translocation (Lo Presti et al., 2017). In summary, a number of approaches are available to assay for secretion of ChEC153.3 from *C. higginsianum*, and its translocation to the host.

6.2.2 Can ChEC153.1 characterisation be used to infer roles of ChEC153.3?

Chapters 3 and 4 of this thesis use the coding sequence for the *ChEC153* gene as predicted by Zampounis *et al.*, *ChEC153.1* (*CH63R_12252*, Zampounis et al., 2016), a gene model which I proved to be imperfect in Chapter 5 – instead identifying *ChEC153.3* as the true gene model for *ChEC153* (Section 5.3.2). Therefore, whether the observations from Chapters 3 and 4 regarding ChEC153.1 truly relate to the genuine putative effector function remain to be confirmed. Given the similarity between ChEC153.1 and ChEC153.3 (Section 5.3.2.6), I hypothesise that a mature sequence of ChEC153.3 would localise similarly to ChEC153.1 in *N. benthamiana*. The signal peptide presence and cleavage point would ideally be explored further before localisation experiments are attempted. This would ensure that chloroplast import is not obstructed by potential N-terminal signal peptide sequence, but as neither ChEC153.1 nor ChEC153.3 possesses a canonical chloroplast transit peptide, this may not be essential to begin probing the subcellular localisation. It is possible that the differences between ChEC153.1 and ChEC153.3 would result in a different localisation for ChEC153.3. This could render the observations seen here for ChEC153.1 reflective

of a new-to-nature targeting of starch granule initiation. Expression of novel genes *in planta* that perturb starch granule initiation may present valuable tools in probing the granule initiation process.

6.2.3 Dual transcriptomic analysis supports the involvement of *ChEC153.3* in infection

The relevance of *ChEC153.3* in infection is supported by the specific expression of the gene during development of *in planta* appressoria and in the biotrophic stage of infection (O'Connell et al., 2012). The RNA-Seq analyses presented in **Chapter 5** suggest that *ChEC153.3* impacts the infection transcriptome on both sides of the pathosystem. A number of Arabidopsis genes are differentially expressed during infection with a *C. higginsianum ChEC153.3*-knockout strain relative to the background strain (**Section 5.3.3.5.1**, discussed in **Section 5.4.2.1**), pointing to a role for ChEC153.3 in directly or indirectly impacting the host defence response. Also, differential expression of a number of *C. higginsianum* genes relating to carbohydrate targeting seen in the presence versus absence of *ChEC153.3* (**Section 5.3.3.5.2**, discussed in **Section 5.4.2.2**). The mechanisms by which ChEC153.3 impacts the expression of these genes have not been explored. Examples of effector proteins directly manipulating host gene expression are apparent in the literature, as mentioned in **Section 1.1.2.2**. Most notably, TALEs localise to the nucleus and manipulate host gene expression to suppress the defence response. In *C. higginsianum*, a number of putative effectors localise to the nucleus (Robin et al., 2018, Ohtsu et al., 2023, Tsushima et al., 2021). Transcriptional reprogramming of the host has been suggested to be a key mechanism of effector action in the *C. graminicola*-maize pathosystem, with several nuclear-localised putative effectors with DNA-binding predictions being identified (Vargas et al., 2016). Therefore, it is likely that *C. higginsianum* also possesses effector proteins which target host gene expression, though these have not been thoroughly characterised. Given the localisation of ChEC153.1 to the chloroplast, and the lack of any DNA-binding motifs in ChEC153.3, it is unlikely that ChEC153.3 directly regulates host nuclear gene expression. It is possible that plastid-to-nucleus retrograde signalling is altered by ChEC153.3, and that this may present an alternative mechanism by which ChEC153.3 could impact nuclear gene expression from the chloroplast. It is possible that if the punctate localisation of ChEC153.1 at the site of starch granule initiation is linked to the site of nucleoids, as may be suggested by the partial co-localisation of ChEC153.1 with the nucleoid marker PEND (**Section 3.3.6**), despite its lack of DNA-binding motifs, ChEC153.1 could impact chloroplast gene expression. However, no chloroplast-encoded genes were seen to be differentially expressed in infection with *ΔChEC153-2* relative to the background strain. Instead, I propose that the differential expression of host defence and carbohydrate-related genes is seen as an indirect impact of ChEC153.3. For example, if

ChEC153.3 perturbs starch granule initiation, this could have downstream impacts on sugar levels and signalling, which could cause expression differences for host genes such as *DIN11*.

6.2.4 Putative effector HaRxL94b: similarities and differences to ChEC153.1

6.2.4.1 How conserved is the targeting of punctate chloroplastic processes by effectors?

It is striking that the localisations of two putative effector proteins from pathogens as unrelated as *C. higginsianum* and *H. arabidopsidis*, a fungus and oomycete respectively, should be so similar, and that both putative effector proteins should co-localise with MRC. This may point to the targeting of granule initiation by pathogen effectors as a relatively conserved process. Further, as mentioned in Section 1.10.2, similar localisations have been observed for putative effectors of other species such as fungal pathogens *Pst* and *Melampsora lini* (Andac et al., 2020, Petre et al., 2016a). These effectors could be investigated in the context of starch granule initiation. Wheat mutants in starch granule initiation, including MRC (Chen et al., 2022), have been characterised (Hawkins et al., 2021, Guo et al., 2017, Chia et al., 2020, Kamble et al., 2023) and would present appropriate hosts in which to probe for starch-related infection phenotypes for putative effectors of pathogens that infect wheat.

Although seen for the abovementioned putative effectors, this punctate sub-chloroplastic localisation is not reported very widely for effectors. Some effector localisation screens have employed N-terminal fluorescent tags for visualisation of subcellular localisations – for example the *C. higginsianum* effector localisation screen carried out by Robin and colleagues (Robin et al., 2018). Given that canonical chloroplast transit peptides are present at the N-terminus of proteins, N-terminal tags may interfere with the localisation of such proteins (Carrie et al., 2009) and preclude their identification as chloroplast-targeted putative effectors. However, at least in the case of non-canonical chloroplast localised effectors, N-terminal fluorophores may not disrupt their localisation. For example, Andac and colleagues show that PstCTE1, which lacks a canonical chloroplast targeting sequence, is able to localise to the chloroplast irrespective of the terminus at which the protein is tagged (Andac et al., 2020). Further, the accurate prediction of signal peptides may be critical in establishing the true subcellular localisations of effector proteins for similar reasons – exemplified by the work presented in Section 3.3.3. Different predictions for the signal peptide cleavage point can result in markedly different localisations being seen when the expected mature putative effector sequence is expressed in *N. benthamiana*. Should signal peptides be falsely predicted to be cleaved further downstream than their bona fide cleavage points,

chloroplast transit peptides may be unknowingly lost during cloning of mature effector sequences. It is further possible that C-terminal fluorophore tagging could interfere with other localisation signals such as peroxisome targeting sequences (Gould et al., 1987). Thus, it may be informative to evaluate the sub-cellular localisation of proteins of interest using both C-terminal and N-terminal fluorophores independently, to allow inferences on the true localisation (Tanz et al., 2013). It is therefore possible that effector sub-chloroplastic localisations similar to those of ChEC153.1 and HaRxL94b are more common than current research suggest. Effector targeting of host processes taking place at chloroplastic puncta may represent a widely adopted strategy by pathogens.

6.2.4.2 ChEC153.1 and HaRxL94b localisations and biological functions may differ

Notably, the localisations of ChEC153.1 and HaRxL94b differ by the additional targeting of the nucleus by the *Hpa* putative effector (Section 1.9.2). Targeting of multiple subcellular compartments is fairly common for plant proteins (Carrie et al., 2009, Carrie and Whelan, 2013), and it therefore stands to reason that pathogen effectors targeting plant processes may also localise to multiple organelles. For example, Liu and colleagues identify a candidate RxLR effector of the oomycete *Plasmopara viticola* to localise to both chloroplasts and mitochondria in *N. benthamiana* (Liu et al., 2018). This may reflect ambiguity between chloroplastic and mitochondrial transit peptides (Kunze and Berger, 2015), or a functional requirement of the effector to target both organelles. Dual targeting to the chloroplast and nucleus is not believed to be as common as dual targeting of other organelles, such as the chloroplast and mitochondria. However, some host proteins display dual chloroplastic/nuclear targeting. For example, the Arabidopsis SIGMA FACTOR BINDING PROTEIN 1 is dually targeted to the chloroplast and nucleus, interacting with key defence transcription factor WRKY33 in the nucleus (Lai et al., 2011).

The dual localisation to chloroplastic puncta and the nucleus raises the possibility of HaRxL94b impacting host gene expression. The partial co-localisation of HaRxL94b with nucleoid marker PEND (Section 3.3.6) indicates some correlation between HaRxL94b and the sites of both chloroplastic and nuclear gene expression, despite its lack of any clearly annotated DNA-binding domains. Demonstrating the potential for proteins localised to both the chloroplast and nucleus to impact nuclear gene expression, the Arabidopsis transcription factor pTAC1 (also known as WHIRLY1 and WHY1), despite being nuclear-encoded, has been suggested to act as a retrograde signal able to translocate from the chloroplast to the nucleus (Isemer et al., 2012). pTAC1 has been suggested to be nucleoid-associated, but the extent to which it localises to nucleoids is not clear, with one paper describing it as rarely associated with nucleoids, and pointing to a potential more dominant role in RNA-processing compared to DNA-binding in the chloroplast (Melonek et al., 2010). In the nucleus, pTAC1 impacts host gene expression, including in defence-related responses

(Lai et al., 2022), and playing a role in SA-mediated defence (Desveaux et al., 2004). Additionally, two further PEP-complex associated proteins PAP5 (aka pTAC12/HEMERA) and PAP8 (pTAC6), with which PAP5 interacts, are dually localised to chloroplasts and nuclei (Nevarez et al., 2017, Liebers et al., 2020). It may therefore be informative to assay for host transcriptional responses in the context of *HaRxL94b*. Due to the genetic intractability of *H. arabidopsidis*, generation of *HaRxL94b*-knockout strains with which to probe this is not currently achievable, but host-induced gene silencing, or spray induced gene silencing (Bilir et al., 2019), could instead be leveraged to explore this question, and the role of HaRxL94b in infection more broadly. Host transcriptional responses could alternatively be probed during transient expression of the putative effector in Arabidopsis (Zhang et al., 2020), or using stable transgenic lines for inducible expression of *HaRxL94b* (Zuo et al., 2000). Inducible transgenic lines may be applicable to studying HaRxL94b function more broadly, as exemplified by the use of β-estradiol inducible expression of Phytoplasma effectors in Arabidopsis (Omenge et al., 2021).

While co-localisation experiments in *N. benthamiana* revealed a complete overlap between ChEC153.1 and HaRxL94b puncta (Section 3.3.1), and between ChEC153.1 and MRC puncta (Section 3.3.4), when *HaRxL94b* and *MRC* were co-expressed, many chloroplasts appeared to display a greater number of HaRxL94b puncta than MRC puncta. Expression of *HaRxL94b* was also not seen to induce the formation of GBSS-RFP puncta indicative of starch granules (Section 3.3.7), suggesting that ChEC153.1 and HaRxL94b function differently. The impacts of HaRxL94b on host starch remain to be tested. While strong evidence is presented here for an interaction between ChEC153.1 and CRBIC, and both MRC and TaCSP41a were also able to recruit CRBIC to puncta, HaRxL94b did not alter the CRBIC localisation during co-expression in *N. benthamiana* (Section 4.3.6). Thus, while the sub-chloroplastic localisations of ChEC153.1 and HaRxL94b appear indistinguishable, their host targets, aside from MRC, seem to differ. The majority of the work in this thesis exploited the genetic malleability and easy axenic culturing and of *C. higginsianum* to investigate ChEC153.1 and ChEC153.3, and a number of questions therefore remain regarding the *H. arabidopsidis* putative effector HaRxL94b. For example, assaying for host targets and any impact of *HaRxL94b* expression on total starch and sugar content of plants, and susceptibility, remain avenues to be explored.

6.3 Links between starch granule initials and proteins relating to RNA-processing and plastid transcription

Throughout this project, a number of proteins with RNA-binding functions or predictions, and links to plastidial transcription appear. For example, I observed co-localisation of both putative effector proteins with known RNA-binding protein TaCSP41a, which is also associated with the PEP-

complex. Further, I identified CRBIC, annotated as an RNA-binding protein due to its putative S4-domain, as a likely host interactor of ChEC153.1. Other candidate interactors of ChEC153.1 include PEP-associated proteins such as FSD3, MurE and FLN2, as highlighted in Section 4.4.4. I also observed partial co-localisation of the putative effectors with the nucleoid marker PEND. While these proteins were studied here in the context of the putative effectors and from an infection/immunity perspective, their correlation with starch granule initiation proteins raises questions outside of the infection context. For example, TaCSP41a was seen to co-localise with MRC, and co-expression with MRC is sufficient to recruit CRBIC to chloroplastic puncta.

Whether starch granule initiation takes place at the sites of RNA processing or plastidial translation at nucleoids is unclear, but taken together, these results suggest at least some overlap, perhaps transiently, between the punctate structures associated with starch granule initials (introduced in Section 1.7.2) and nucleoids (introduced in Section 1.10.1). Whether the sites of starch granule initiation are spatially distinct from those of plastidial transcription remains to be determined. However, the work I present here indicates some likely link between these processes which could be investigated further. A first step toward elucidating this potential link could be the co-localisation of key starch granule initiation proteins and known nucleoid markers. Assaying multiple starch granule initiation proteins, for example MRC, SS4, PTST2, and MFP1, may help to disentangle whether certain aspects of granule initiation correlate with nucleoids/PEP-associated proteins. The localisations of these granule initial proteins is believed to be subtly different – for example, while MFP1 is thought to be primarily associated with the thylakoid membrane, PTST2 is partially associated with the thylakoid, and both MRC and SS4 are predominantly soluble/stromal (Seung et al., 2018). Therefore, it is possible that some granule initiation proteins correlate more or less strongly with nucleoids. While evidence for associations of these proteins with nucleoids are currently unclear, it is notable that MFP1 was first described to be nucleoid-associated by Jeong and colleagues, before it was known to be involved in granule initiation (Jeong et al., 2003). Jeong *et al*. describe the DNA-binding capacity of MFP1 and cite the appearance of MFP1 in nucleoid fractions of chloroplasts as evidence for its nucleoid association (Jeong et al., 2003). MFP1 is not likely to be required for the functionality of plastidial transcription, suggested by its abundance not correlating with transcriptional activity, its localisation not being solely associated with nucleoids (Melonek et al., 2010), and no evidence that *mfp1* mutants have any defect in plastid transcription (Seung et al., 2018). Investigating these potential links may help to expand our understanding of the intricacies of granule initiation.

6.4 Importance of gene model validation

The work contained within Chapter 5 of this thesis highlights discrepancies seen in gene model annotations between genome assemblies. Where resources such as highly curated genome sequences are limited for organisms of interest, *in silico* tools are invaluable in predicting gene models. However, from my experience, I would encourage the experimental confirmation of gene models predicted *in silico* as far as practicably possible. This may be especially pertinent if published gene models disagree, but even where gene model predictions agree these should not be treated as infallible.

The gold standard of gene model validation remains direct cloning and sequencing of genes of interest from cDNA, particularly if strains in use are not well annotated or differ from published sequences. While gene model validation in this way is labour-intensive and impracticable in the context of large-scale screening projects, once genes of interest are selected, ideally gene models would be confirmed experimentally. Where RNA-Seq data are available, I would advocate for comparing reads mapping to the region of interest to any *in silico* predicted gene models. I found the publicly available *C. higginsianum* RNA-Seq data (O'Connell et al., 2012) to be highly accurate in indicating intron/exon boundaries for *ChEC153.3*, and so this approach may offer a comparatively quick and simple method to complement gene model predictions in lieu of direct cDNA cloning and sequence verification.

6.5 Final conclusions

While the importance of chloroplastic processes in defence responses is a wide area of research, any involvement of starch granule initiation was hitherto unexplored. The work presented here has identified potential links between host starch granule initiation and susceptibility to infection through the characterisation of two specific putative effector proteins: ChEC153 and HaRxL94b. The mechanisms by which these putative effectors function remains to be determined, particularly in light of the new gene model which I established for *ChEC153*. However, the importance of *ChEC153* in infection, and a potential link to carbohydrate targeting, is highlighted in my analysis of transcriptional differences during infection of Arabidopsis with *C. higginsianum* strains in the presence/absence of *ChEC153*. Further, questions have been raised regarding whether the seemingly functionally distinct punctate processes of starch granule initiation and of plastidial transcription are spatially separated.

CHAPTER 7

Bibliography

ABT, M. R., PFISTER, B., SHARMA, M., EICKE, S., BURGY, L., NEALE, I., SEUNG, D. & ZEEMAN, S. C. 2020. STARCH SYNTHASE5, a Noncanonical Starch Synthase-Like Protein, Promotes Starch Granule Initiation in Arabidopsis. *Plant Cell,* 32**,** 2543-2565.

ADACHI, H., DEREVNINA, L. & KAMOUN, S. 2019. NLR singletons, pairs, and networks: evolution, assembly, and regulation of the intracellular immunoreceptor circuitry of plants. *Current Opinion in Plant Biology,* 50**,** 121-131.

AHN, H. K., LIN, X., OLAVE-ACHURY, A. C., DEREVNINA, L., CONTRERAS, M. P., KOURELIS, J., WU, C. H., KAMOUN, S. & JONES, J. D. G. 2023. Effector-dependent activation and oligomerization of plant NRC class helper NLRs by sensor NLR immune receptors Rpi-amr3 and Rpi-amr1. *Embo Journal,* 42.

ALAM, S. B., READE, R., MAGHODIA, A. B., GHOSHAL, B., THEILMANN, J. & ROCHON, D. 2021. Targeting of cucumber necrosis virus coat protein to the chloroplast stroma attenuates host defense response. *Virology,* 554**,** 106-119.

ALBINIAK, A. M., BAGLIERI, J. & ROBINSON, C. 2012. Targeting of lumenal proteins across the thylakoid membrane. *Journal of Experimental Botany,* 63**,** 1689-1698.

ALLEN, R. L., BITTNER-EDDY, P. D., GRENVITTE-BRIGGS, L. J., MEITZ, J. C., REHMANY, A. P., ROSE, L. E. & BEYNON, J. L. 2004. Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science,* 306**,** 1957-1960.

ALONSO, J. M., STEPANOVA, A. N., LEISSE, T. J., KIM, C. J., CHEN, H. M., SHINN, P., STEVENSON, D. K., ZIMMERMAN, J., BARAJAS, P., CHEUK, R., GADRINAB, C., HELLER, C., JESKE, A., KOESEMA, E., MEYERS, C. C., PARKER, H., PREDNIS, L., ANSARI, Y., CHOY, N., DEEN, H., GERALT, M., HAZARI, N., HOM, E., KARNES, M., MULHOLLAND, C., NDUBAKU, R., SCHMIDT, I., GUZMAN, P., AGUILAR-HENONIN, L., SCHMID, M., WEIGEL, D., CARTER, D. E., MARCHAND, T., RISSEEUW, E., BROGDEN, D., ZEKO, A., CROSBY, W. L., BERRY, C. C. & ECKER, J. R. 2003. Genome-wide Insertional mutagenesis of *Arabidopsis thaliana*. *Science,* 301**,** 653-657.

AMARASINGHE, S. L., SU, S., DONG, X. Y., ZAPPIA, L., RITCHIE, M. E. & GOUIL, Q. 2020. Opportunities and challenges in long-read sequencing data analysis. *Genome Biology,* 21.

ANDAC, A., OZKETEN, A. C., DAGVADORJ, B. & AKKAYA, M. S. 2020. An effector of *Puccinia striiformis* f. sp. *tritici* targets chloroplasts with a novel and robust targeting signal. *European Journal of Plant Pathology,* 157**,** 751-765.

ANDREWS, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data* [Online]. Babraham Institute. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed].

ARBIBE, L., KIM, D. W., BATSCHE, E., PEDRON, T., MATEESCU, B., MUCHARDT, C., PARSOT, C. & SANSONETTI, P. J. 2007. An injected bacterial effector targets chromatin access for transcription factor NF-kappa B to alter transcription of host genes involved in immune responses. *Nature Immunology,* 8**,** 47-56.

ARMBRUSTER, U., HERTLE, A., MAKARENKO, E., ZUHLKE, J., PRIBIL, M., DIETZMANN, A., SCHLIEBNER, I., ASEEVA, E., FENINO, E., SCHARFENBERG, M., VOIGT, C. & LEISTER, D. 2009. Chloroplast Proteins without Cleavable Transit Peptides: Rare Exceptions or a Major Constituent of the Chloroplast Proteome? *Molecular Plant,* 2**,** 1325-1335.

ARMENTEROS, J. J. A., SALVATORE, M., EMANUELSSON, O., WINTHER, O., VON HEIJNE, G., ELOFSSON, A. & NIELSEN, H. 2019a. Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance,* 2.

ARMENTEROS, J. J. A., SONDERBY, C. K., SONDERBY, S. K., NIELSEN, H. & WINTHER, O. 2017. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics,* 33**,** 3387-3395.

ARMENTEROS, J. J. A., TSIRIGOS, K. D., SONDERBY, C. K., PETERSEN, T. N., WINTHER, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2019b. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology,* 37**,** 420–423.

ARSOVA, B., HOJA, U., WIMMELBACHER, M., GREINER, E., USTUN, S., MELZER, M., PETERSEN, K., LEIN, W. & BORNKE, F. 2010. Plastidial Thioredoxin z Interacts with Two Fructokinase-Like Proteins in a Thiol-Dependent Manner: Evidence for an Essential Role in Chloroplast Development in Arabidopsis and *Nicotiana benthamiana*. *Plant Cell,* 22**,** 1498-1515.

ASAI, S., CEVIK, V., JONES, J. D. G. & SHIRASU, K. 2023. Cell-specific RNA profiling reveals host genes expressed in Arabidopsis cells haustoriated by downy mildew. *Plant Physiology,* 193**,** 259–270.

ASAI, S., FURZER, O. J., CEVIK, V., KIM, D. S., ISHAQUE, N., GORITSCHNIG, S., STASKAWICZ, B. J., SHIRASU, K. & JONES, J. D. G. 2018. A downy mildew effector evades recognition by polymorphism of expression and subcellular localization. *Nature Communications,* 9.

ASAI, S., RALLAPALLI, G., PIQUEREZ, S. J. M., CAILLAUD, M. C., FURZER, O. J., ISHAQUE, N., WIRTHMUELLER, L., FABRO, G., SHIRASU, K. & JONES, J. D. G. 2014. Expression Profiling during Arabidopsis/Downy Mildew Interaction Reveals a Highly-Expressed Effector That Attenuates Responses to Salicylic Acid. *Plos Pathogens,* 10.

AUSTIN, J. R., FROST, E., VIDI, P. A., KESSLER, F. & STAEHELIN, L. A. 2006. Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. *Plant Cell,* 18**,** 1693-1703.

BAILEY, K., CEVIK, V., HOLTON, N., BYRNE-RICHARDSON, J., SOHN, K. H., COATES, M., WOODS-TOR, A., AKSOY, H. M., HUGHES, L., BAXTER, L., JONES, J. D. G., BEYNON, J., HOLUB, E. B. & TOR, M. 2011. Molecular Cloning of ATR5^{Emoy2} from *Hyaloperonospora arabidopsidis*, an Avirulence Determinant That Triggers RPP5-Mediated Defense in Arabidopsis. *Molecular Plant-Microbe Interactions,* 24**,** 827-838.

BALMER, E. A. & FASO, C. 2021. The Road Less Traveled? Unconventional Protein Secretion at Parasite-Host Interfaces. *Frontiers in Cell and Developmental Biology,* 9.

BANERJEE, S., BHANDARY, P., WOODHOUSE, M., SEN, T. Z., WISE, R. P. & ANDORF, C. M. 2021. FINDER: an automated software package to annotate eukaryotic genes from RNA-Seq data and associated protein sequences. *Bmc Bioinformatics,* 22.

BAROJA-FERNÁNDEZ, E., MUÑOZ, F. J., LI, J., BAHAJI, A., ALMAGRO, G., MONTERO, M., ETXEBERRIA, E., HIDALGO, M., SESMA, M. T. & POZUETA-ROMERO, J. 2012. Sucrose synthase activity in the *sus1 sus2 sus3 sus4* Arabidopsis mutant is sufficient to support normal cellulose and starch production. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 321-326.

BAUD, S., VAULTIER, M. N. & ROCHAT, C. 2004. Structure and expression profile of the sucrose synthase multigene family in Arabidopsis. *Journal of Experimental Botany,* 55**,** 397-409.

BAXTER, L., TRIPATHY, S., ISHAQUE, N., BOOT, N., CABRAL, A., KEMEN, E., THINES, M., AH-FONG, A., ANDERSON, R., BADEJOKO, W., BITTNER-EDDY, P., BOORE, J. L., CHIBUCOS, M. C., COATES, M., DEHAL, P., DELEHAUNTY, K., DONG, S. M., DOWNTON, P., DUMAS, B., FABRO, G., FRONICK, C., FUERSTENBERG, S. I., FULTON, L., GAULIN, E., GOVERS, F., HUGHES, L., HUMPHRAY, S., JIANG, R. H. Y., JUDELSON, H., KAMOUN, S., KYUNG, K., MEIJER, H., MINX, P., MORRIS, P., NELSON, J., PHUNTUMART, V., QUTOB, D., REHMANY, A., ROUGON-CARDOSO, A., RYDEN, P., TORTO-ALALIBO, T., STUDHOLME, D., WANG, Y. C., WIN, J., WOOD, J., CLIFTON, S. W., ROGERS, J., VAN DEN ACKERVEKEN, G., JONES, J. D. G.,

MCDOWELL, J. M., BEYNON, J. & TYLER, B. M. 2010. Signatures of Adaptation to Obligate Biotrophy in the *Hyaloperonospora arabidopsidis* Genome. *Science,* 330**,** 1549-1551.

BELHAJ, K., LIN, B. Q. & MAUCH, F. 2009. The chloroplast protein RPH1 plays a role in the immune response of Arabidopsis to *Phytophthora brassicae*. *Plant Journal,* 58**,** 287-298.

BENDTSEN, J. D., NIELSEN, H., VON HEIJNE, G. & BRUNAK, S. 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology,* 340**,** 783-795.

BERGER, S., SINHA, A. K. & ROITSCH, T. 2007. Plant physiology meets phytopathology: plant primary metabolism and plant-pathogen interactions. *Journal of Experimental Botany,* 58**,** 4019-4026.

BI, G. Z., SU, M., LI, N., LIANG, Y., DANG, S., XU, J. C., HU, M. J., WANG, J. Z., ZOU, M. X., DENG, Y. A., LI, Q. Y., HUANG, S. J., LI, J. J., CHAI, J. J., HE, K. M., CHEN, Y. H. & ZHOU, J. M. 2021. The ZAR1 resistosome is a calcium-permeable channel triggering plant immune signaling. *Cell,* 184**,** 3528-3541.

BIENIAWSKA, Z., BARRATT, D. H. P., GARLICK, A. P., THOLE, V., KRUGER, N. J., MARTIN, C., ZRENNER, R. & SMITH, A. M. 2007. Analysis of the sucrose synthase gene family in Arabidopsis. *Plant Journal,* 49**,** 810-828.

BILIR, Ö., TELLI, O., NORMAN, C., BUDAK, H., HONG, Y. G. & TÖR, M. 2019. Small RNA inhibits infection by downy mildew pathogen *Hyaloperonospora arabidopsidis*. *Molecular Plant Pathology,* 20**,** 1523-1534.

BOBIK, K., MCCRAY, T. N., ERNEST, B., FERNANDEZ, J. C., HOWELL, K. A., LANE, T., STATON, M. & BURCH-SMITH, T. M. 2017. The chloroplast RNA helicase ISE2 is required for multiple chloroplast RNA processing steps in *Arabidopsis thaliana*. *Plant Journal,* 91**,** 114-131.

BOEVINK, P. C., WANG, X. D., MCLELLAN, H., HE, Q., NAQVI, S., ARMSTRONG, M. R., ZHANG, W., HEIN, I., GILROY, E. M., TIAN, Z. D. & BIRCH, P. R. J. 2016. A *Phytophthora infestan*s RXLR effector targets plant PP1c isoforms that promote late blight disease. *Nature Communications,* 7.

BOLLENBACH, T. J., SHARWOOD, R. E., GUTIERREZ, R., LERBS-MACHE, S. & STERN, D. B. 2009. The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in Arabidopsis. *Plant Molecular Biology,* 69**,** 541-552.

BOS, J. I. B., KANNEGANTI, T. D., YOUNG, C., CAKIR, C., HUITEMA, E., WIN, J., ARMSTRONG, M. R., BIRCH, P. R. J. & KAMOUN, S. 2006. The C-terminal half of Phytophthora infestans RXLR effector AVR3a is sufficient to trigger R3a-mediated hypersensitivity and suppress INF1-induced cell death in Nicotiana benthamiana. *Plant Journal,* 48**,** 165-176.

BOS, J. I. B., PRINCE, D., PITINO, M., MAFFEI, M. E., WIN, J. & HOGENHOUT, S. A. 2010. A Functional Genomics Approach Identifies Candidate Effectors from the Aphid Species *Myzus persicae* (Green Peach Aphid). *Plos Genetics,* 6.

BOUCHNAK, I., BRUGIÈRE, S., MOYET, L., LE GALL, S., SALVI, D., KUNTZ, M., TARDIF, M. & ROLLAND, N. 2019. Unraveling Hidden Components of the Chloroplast Envelope Proteome: Opportunities and Limits of Better MS Sensitivity. *Molecular & Cellular Proteomics,* 18**,** 1285-1306.

BOUTEMY, L. S., KING, S. R. F., WIN, J., HUGHES, R. K., CLARKE, T. A., BLUMENSCHEIN, T. M. A., KAMOUN, S. & BANFIELD, M. J. 2011. Structures of Phytophthora RXLR Effector Proteins: A Conserved but Adaptable Fold Underpins Functional Diversity *Journal of Biological Chemistry,* 286**,** 35834-35842.

BRADLEY, D. J., KJELLBOM, P. & LAMB, C. J. 1992. Elicitor-induced and wound-induced oxidative cross-linking of a proline-rich plant-cell wall protein - a novel, rapid defense response. *Cell,* 70**,** 21-30.

BREEN, J. & BELLGARD, M. 2010. Germin-like proteins (GLPs) in cereal genomes: gene clustering and dynamic roles in plant defence. *Functional & Integrative Genomics,* 10**,** 463-476.

BREEZE, E. & MULLINEAUX, P. M. 2022. The Passage of H_2O_2 from Chloroplasts to Their Associated Nucleus during Retrograde Signalling: Reflections on the Role of the Nuclear Envelope. *Plants-Basel,* 11.

BRUNA, T., HOFF, K. J., LOMSADZE, A., STANKE, M. & BORODOVSKY, M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP plus and AUGUSTUS supported by a protein database. *Nar Genomics and Bioinformatics,* 3.

BRUNKARD, J. O., RUNKEL, A. M. & ZAMBRYSKI, P. C. 2015. Chloroplasts extend stromules independently and in response to internal redox signals. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** 10044-10049.

BUJA, I., SABELLA, E., MONTEDURO, A. G., CHIRIACÒ, M. S., DE BELLIS, L., LUVISI, A. & MARUCCIO, G. 2021. Advances in Plant Disease Detection and Monitoring: From Traditional Assays to In-Field Diagnostics. *Sensors,* 21.

BURCH-SMITH, T. M., BRUNKARD, J. O., CHOI, Y. G. & ZAMBRYSKI, P. C. 2011. Organelle-nucleus cross-talk regulates plant intercellular communication via plasmodesmata. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** E1451-E1460.

BÜRGY, L., EICKE, S., KOPP, C., JENNY, C., LU, K. J., ESCRIG, S., MEIBOM, A. & ZEEMAN, S. C. 2021. Coalescence and directed anisotropic growth of starch granule initials in subdomains of *Arabidopsis thaliana* chloroplasts. *Nature Communications,* 12.

CAI, Q., QIAO, L. L., WANG, M., HE, B. Y., LIN, F. M., PALMQUIST, J., HUANG, S. N. D. & JIN, H. L. 2018. Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science,* 360**,** 1126-1129.

CAILLAUD, M. C., ASAI, S., RALLAPALLI, G., PIQUEREZ, S., FABRO, G. & JONES, J. D. G. 2013. A Downy Mildew Effector Attenuates Salicylic Acid-Triggered Immunity in Arabidopsis by Interacting with the Host Mediator Complex. *Plos Biology,* 11.

CAILLAUD, M. C., PIQUEREZ, S. J. M., FABRO, G., STEINBRENNER, J., ISHAQUE, N., BEYNON, J. & JONES, J. D. G. 2012. Subcellular localization of the Hpa RxLR effector repertoire identifies a tonoplast-associated protein HaRxL17 that confers enhanced plant susceptibility. *Plant Journal,* 69**,** 252-265.

CAMPO, S., GILBERT, K. B. & CARRINGTON, J. C. 2016. Small RNA-Based Antiviral Defense in the Phytopathogenic Fungus *Colletotrichum higginsianum*. *Plos Pathogens,* 12.

CANNON, P. F., DAMM, U., JOHNSTON, P. R. & WEIR, B. S. 2012. *Colletotrichum* - current status and future directions. *Studies in Mycology***,** 181-213.

CANTAREL, B. L., KORF, I., ROBB, S. M. C., PARRA, G., ROSS, E., MOORE, B., HOLT, C., ALVARADO, A. S. & YANDELL, M. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research,* 18**,** 188-196.

CAO, Y. R., LIANG, Y., TANAKA, K., NGUYEN, C. T., JEDRZEJCZAK, R. P., JOACHIMIAK, A. & STACEY, G. 2014. The kinase LYK5 is a major chitin receptor in Arabidopsis and forms a chitin-induced complex with related kinase CERK1. *Elife,* 3.

CAPLAN, J. L., KUMAR, A. S., PARK, E., PADMANABHAN, M. S., HOBAN, K., MODLA, S., CZYMMEK, K. & DINESH-KUMAR, S. P. 2015. Chloroplast Stromules Function during Innate Immunity. *Developmental Cell,* 34**,** 45-57.

CARREÓN-ANGUIANO, K. G., ISLAS-FLORES, I., VEGA-ARREGUÍN, J., SÁENZ-CARBONELL, L. & CANTO-CANCHÉ, B. 2020. EffHunter: A Tool for Prediction of Effector Protein Candidates in Fungal Proteomic Databases. *Biomolecules,* 10**,** 712.

CARRIE, C., KÜHN, K., MURCHA, M. W., DUNCAN, O., SMALL, I. D., O'TOOLE, N. & WHELAN, J. 2009. Approaches to defining dual-targeted proteins in Arabidopsis. *Plant Journal,* 57**,** 1128-1139.

CARRIE, C. & WHELAN, J. 2013. Widespread dual targeting of proteins in land plants: when, where, how and why. *Plant Signal Behav,* 8.

CASPAR, T., HUBER, S. C. & SOMERVILLE, C. 1985. Alterations in growth, photosynthesis, and respiration in a starchless mutant of *Arabidopsis thaliana* (L.) deficient in chloroplast phosphoglucomutase activity. *Plant Physiology,* 79**,** 11-17.

CASPAR, T., LIN, T. P., KAKEFUDA, G., BENBOW, L., PREISS, J. & SOMERVILLE, C. 1991. Mutants of Arabidopsis with altered regulation of starch degradation. *Plant Physiology,* 95**,** 1181-1188.

CHAMBERS, M. C., MACLEAN, B., BURKE, R., AMODEI, D., RUDERMAN, D. L., NEUMANN, S., GATTO, L., FISCHER, B., PRATT, B., EGERTSON, J., HOFF, K., KESSNER, D., TASMAN, N., SHULMAN, N., FREWEN, B., BAKER, T. A., BRUSNIAK, M. Y., PAULSE, C., CREASY, D., FLASHNER, L., KANI, K., MOULDING, C., SEYMOUR, S. L., NUWAYSIR, L. M., LEFEBVRE, B., KUHLMANN, F., ROARK, J., RAINER, P., DETLEV, S., HEMENWAY, T., HUHMER, A., LANGRIDGE, J., CONNOLLY, B., CHADICK, T., HOLLY, K., ECKELS, J., DEUTSCH, E. W.,

MORITZ, R. L., KATZ, J. E., AGUS, D. B., MACCOSS, M., TABB, D. L. & MALLICK, P. 2012. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology,* 30**,** 918-920.

CHANCLUD, E., KISIALA, A., EMERY, N. R. J., CHALVON, V., DUCASSE, A., ROMITI-MICHEL, C., GRAVOT, A., KROJ, T. & MOREL, J. B. 2016. Cytokinin Production by the Rice Blast Fungus Is a Pivotal Requirement for Full Virulence. *Plos Pathogens,* 12.

CHANDA, B., VENUGOPAL, S. C., KULSHRESTHA, S., NAVARRE, D. A., DOWNIE, B., VAILLANCOURT, L., KACHROO, A. & KACHROO, P. 2008. Glycerol-3-phosphate levels are associated with basal resistance to the hemibiotrophic fungus *Colletotrichum higginsianum* in Arabidopsis. *Plant Physiology,* 147**,** 2017-2029.

CHEN, J. 2022. *Characterisation of SS4 and MRC in starch granule initiation.* Doctor of Philosophy, University of East Anglia.

CHEN, J., CHEN, Y., WATSON-LAZOWSKI, A., HAWKINS, E., BARCLAY, J. E., FAHY, B., BOWERS, R. D., CORBIN, K., WARREN, F. J., BLENNOW, A., UAUY, C. & SEUNG, D. 2022. The plastidial protein MRC promotes starch granule initiation in wheat leaves but delays B-type granule initiation in the endosperm. *bioRxiv***,** 2022.10.07.511297.

CHEN, J. W., HAWKINS, E. & SEUNG, D. 2021a. Towards targeted starch modification in plants. *Current Opinion in Plant Biology,* 60.

CHEN, L. Q., HOU, B. H., LALONDE, S., TAKANAGA, H., HARTUNG, M. L., QU, X. Q., GUO, W. J., KIM, J. G., UNDERWOOD, W., CHAUDHURI, B., CHERMAK, D., ANTONY, G., WHITE, F. F., SOMERVILLE, S. C., MUDGETT, M. B. & FROMMER, W. B. 2010. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature,* 468**,** 527-U199.

CHEN, S. Y., MA, T., SONG, S. R., LI, X. L., FU, P. N., WU, W., LIU, J. Q., GAO, Y., YE, W. X., DRY, I. B. & LU, J. 2021b. Arabidopsis downy mildew effector HaRxLL470 suppresses plant immunity by attenuating the DNA-binding activity of bZIP transcription factor HY5. *New Phytologist,* 230**,** 1562-1577.

CHEN, X. M., PAN, S., BAI, H. M., FAN, J. X., BATOOL, W., SHABBIR, A., HAN, Y. J., ZHENG, H. K., LU, G. D., LIN, L. L., TANG, W. & WANG, Z. H. 2023. A nonclassically secreted effector of *Magnaporthe oryza*e targets host nuclei and plays important roles in fungal growth and plant infection. *Molecular Plant Pathology,* 24**,** 1093–1106.

CHIA, T., CHIRICO, M., KING, R., RAMIREZ-GONZALEZ, R., SACCOMANNO, B., SEUNG, D., SIMMONDS, J., TRICK, M., UAUY, C., VERHOEVEN, T. & TRAFFORD, K. 2020. A carbohydrate-binding protein, B-GRANULE CONTENT 1, influences starch granule size distribution in a dose-dependent manner in polyploid wheat. *Journal of Experimental Botany,* 71**,** 105-115.

CHINCHILLA, D., BAUER, Z., REGENASS, M., BOLLER, T. & FELIX, G. 2006. The Arabidopsis receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell,* 18**,** 465-476.

CHONG, J., PIRON, M. C., MEYER, S., MERDINOGLU, D., BERTSCH, C. & MESTRE, P. 2014. The SWEET family of sugar transporters in grapevine: VvSWEET4 is involved in the interaction with *Botrytis cinerea*. *Journal of Experimental Botany,* 65**,** 6589-6601.

CHOUDHARY, D. K., PRAKASH, A. & JOHRI, B. N. 2007. Induced systemic resistance (ISR) in plants: mechanism of action. *Indian J Microbiol,* 47**,** 289-297.

CMERO, M., SCHMIDT, B., MAJEWSKI, I. J., EKERT, P. G., OSHLACK, A. & DAVIDSON, N. M. 2021. MINTIE: identifying novel structural and splice variants in transcriptomes using RNA-seq data. *Genome Biology,* 22.

COATES, M. E. & BEYNON, J. L. 2010. Hyaloperonospora arabidopsidis as a Pathogen Model. *In:* VANALFEN, N. K., BRUENING, G. & LEACH, J. E. (eds.) *Annual Review of Phytopathology, Vol 48.*

COLEMAN, A. W. 1979. USE OF THE FLUOROCHROME 4'6-DIAMIDINO-2-PHENYLINDOLE IN GENETIC AND DEVELOPMENTAL STUDIES OF CHLOROPLAST DNA. *Journal of Cell Biology,* 82**,** 299-305.

COLLEMARE, J., O'CONNELL, R. & LEBRUN, M. H. 2019. Nonproteinaceous effectors: the terra incognita of plant-fungal interactions. *New Phytologist,* 223**,** 590-596.

CONTRERAS, M. P., PAI, H., TUMTAS, Y., DUGGAN, C., YUEN, E. L. H., CRUCES, A. V., KOURELIS, J., AHN, H. K., LEE, K. T., WU, C. H., BOZKURT, T. O., DEREVNINA, L. & KAMOUN, S. 2023. Sensor NLR immune proteins activate oligomerization of their NRC helpers in response to plant pathogens. *Embo Journal,* 42.

CORREDOR-MORENO, P., BADGAMI, R., JONES, S. & SAUNDERS, D. G. O. 2022. Temporally coordinated expression of nuclear genes encoding chloroplast proteins in wheat promotes *Puccinia striiformis* f. sp. *tritici* infection. *Communications Biology,* 5.

CRUMPTON-TAYLOR, M., GRANDISON, S., PNG, K. M. Y., BUSHBY, A. J. & SMITH, A. M. 2012. Control of Starch Granule Numbers in Arabidopsis Chloroplasts. *Plant Physiology,* 158**,** 905-916.

CRUMPTON-TAYLOR, M., PIKE, M., LU, K. J., HYLTON, C. M., FEIL, R., EICKE, S., LUNN, J. E., ZEEMAN, S. C. & SMITH, A. M. 2013. Starch synthase 4 is essential for coordination of starch granule formation with chloroplast division during Arabidopsis leaf expansion. *New Phytologist,* 200**,** 1064-1075.

DALLERY, J. F., LAPALU, N., ZAMPOUNIS, A., PIGNE, S., LUYTEN, I., AMSELEM, J., WITTENBERG, A. H. J., ZHOU, S. G., DE QUEIROZ, M. V., ROBIN, G. P., AUGER, A., HAINAUT, M., HENRISSAT, B., KIM, K. T., LEE, Y. H., LESPINET, O., SCHWARTZ, D. C., THON, M. R. & O'CONNELL, R. J. 2017. Gapless genome assembly of *Colletotrichum higginsianum* reveals chromosome structure and association of transposable elements with secondary metabolite gene clusters. *Bmc Genomics,* 18.

DAMM, U., O'CONNELL, R. J., GROENEWALD, J. Z. & CROUS, P. W. 2014. The *Colletotrichum destructivum* species complex - hemibiotrophic pathogens of forage and field crops. *Studies in Mycology***,** 49-84.

DANECEK, P., BONFIELD, J. K., LIDDLE, J., MARSHALL, J., OHAN, V., POLLARD, M. O., WHITWHAM, A., KEANE, T., MCCARTHY, S. A., DAVIES, R. M. & LI, H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience,* 10.

DE GROOT, M. J., BUNDOCK, P., HOOYKAAS, P. J. & BEIJERSBERGEN, A. G. 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat Biotechnol,* 16**,** 839-842.

DE GUILLEN, K., ORTIZ-VALLEJO, D., GRACY, J., FOURNIER, E., KROJ, T. & PADILLA, A. 2015. Structure Analysis Uncovers a Highly Diverse but Structurally Conserved Effector Family in Phytopathogenic Fungi. *Plos Pathogens,* 11.

DE JONGE, R., VAN ESSE, H. P., KOMBRINK, A., SHINYA, T., DESAKI, Y., BOURS, R., VAN DER KROL, S., SHIBUYA, N., JOOSTEN, M. & THOMMA, B. 2010. Conserved Fungal LysM Effector Ecp6 Prevents Chitin-Triggered Immunity in Plants. *Science,* 329**,** 953-955.

DEAN, R., VAN KAN, J. A. L., PRETORIUS, Z. A., HAMMOND-KOSACK, K. E., DI PIETRO, A., SPANU, P. D., RUDD, J. J., DICKMAN, M., KAHMANN, R., ELLIS, J. & FOSTER, G. D. 2012. The Top 10 fungal pathogens in molecular plant pathology *Molecular Plant Pathology,* 13**,** 414-430.

DEB, D., ANDERSON, R. G., HOW-YEW-KIN, T., TYLER, B. M. & MCDOWELL, J. M. 2018a. Conserved RxLR Effectors From Oomycetes *Hyaloperonospora arabidopsidis* and *Phytophthora sojae* Suppress PAMP- and Effector-Triggered Immunity in Diverse Plants. *Molecular Plant-Microbe Interactions,* 31**,** 374-385.

DEB, D., MACKEY, D., OPIYO, S. O. & MCDOWELL, J. M. 2018b. Application of alignment-free bioinformatics methods to identify an oomycete protein with structural and functional similarity to the bacterial AvrE effector protein. *Plos One,* 13.

DÉJARDIN, A., SOKOLOV, L. N. & KLECZKOWSKI, L. A. 1999. Sugar/osmoticum levels modulate differential abscisic acid-independent expression of two stress-responsive sucrose synthase genes in Arabidopsis. *Biochemical Journal,* 344**,** 503-509.

DELANEY, T. P., UKNES, S., VERNOOIJ, B., FRIEDRICH, L., WEYMANN, K., NEGROTTO, D., GAFFNEY, T., GUTRELLA, M., KESSMANN, H., WARD, E. & RYALS, J. 1994. A central role of salicylic-acid in plant-disease resistance. *Science,* 266**,** 1247-1250.

DELGADO-BAQUERIZO, M., GUERRA, C. A., CANO-DIAZ, C., EGIDI, E., WANG, J. T., EISENHAUER, N., SINGH, B. K. & MAESTRE, F. T. 2020. The proportion of soil-borne pathogens increases with warming at the global scale. *Nature Climate Change,* 10**,** 550–554.

DELIC, M., VALLI, M., GRAF, A. B., PFEFFER, M., MATTANOVICH, D. & GASSER, B. 2013. The secretory pathway: exploring yeast diversity. *Fems Microbiology Reviews,* 37**,** 872-914.

DENG, Y. W., NING, Y. S., YANG, D. L., ZHAI, K. R., WANG, G. L. & HE, Z. H. 2020. Molecular Basis of Disease Resistance and Perspectives on Breeding Strategies for Resistance Improvement in Crops. *Molecular Plant,* 13**,** 1402-1419.

DESIKAN, R., LAST, K., HARRETT-WILLIAMS, R., TAGLIAVIA, C., HARTER, K., HOOLEY, R., HANCOCK, J. T. & NEILL, S. J. 2006. Ethylene-induced stomatal closure in Arabidopsis occurs via AtrbohF-mediated hydrogen peroxide synthesis. *Plant Journal,* 47**,** 907-916.

DESVEAUX, D., SUBRAMANIAM, R., DESPRÉS, C., MESS, J. N., LÉVESQUE, C., FOBERT, P. R., DANGL, J. L. & BRISSON, N. 2004. A "whirly" transcription factor is required for salicylic acid-dependent disease resistance in Arabidopsis. *Developmental Cell,* 6**,** 229-240.

DING, X., JIMENEZ-GONGORA, T., KRENZ, B. & LOZANO-DURAN, R. 2019. Chloroplast clustering around the nucleus is a general response to pathogen perception in *Nicotiana benthamiana*. *Molecular Plant Pathology,* 20**,** 1298-1306.

DJAMEI, A., SCHIPPER, K., RABE, F., GHOSH, A., VINCON, V., KAHNT, J., OSORIO, S., TOHGE, T., FERNIE, A. R., FEUSSNER, I., FEUSSNER, K., MEINICKE, P., STIERHOF, Y. D., SCHWARZ, H., MACEK, B., MANN, M. & KAHMANN, R. 2011. Metabolic priming by a secreted fungal effector. *Nature,* 478**,** 395–398.

DONG, O. X. & RONALD, P. C. 2019. Genetic Engineering for Disease Resistance in Plants: Recent Progress and Future Perspectives. *Plant Physiology,* 180**,** 26-38.

DU, J., VERZAUX, E., CHAPARRO-GARCIA, A., BIJSTERBOSCH, G., KEIZER, L. C. P., ZHOU, J., LIEBRAND, T. W. H., XIE, C. H., GOVERS, F., ROBATZEK, S., VAN DER VOSSEN, E. A. G., JACOBSEN, E., VISSER, R. G. F., KAMOUN, S. & VLEESHOUWERS, V. 2015. Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. *Nature Plants,* 1.

DUAN, G., CHRISTIAN, N., SCHWACHTJE, J., WALTHER, D. & EBENHÖH, O. 2013. The Metabolic Interplay between Plants and Phytopathogens. *Metabolites,* 3**,** 1-23.

DUNKER, F., OBERKOFLER, L., LEDERER, B., TRUTZENBERG, A. & WEIBERG, A. 2021. An Arabidopsis downy mildew non-RxLR effector suppresses induced plant cell death to promote biotroph infection. *Journal of Experimental Botany,* 72**,** 718-732.

ENGELSDORF, T., HORST, R. J., PROLS, R., PROSCHEL, M., DIETZ, F., HUCKELHOVEN, R. & VOLL, L. M. 2013. Reduced Carbohydrate Availability Enhances the Susceptibility of Arabidopsis toward *Colletotrichum higginsianum*. *Plant Physiology,* 162**,** 225-238.

ENGELSDORF, T., WILL, C., HOFMANN, J., SCHMITT, C., MERRITT, B. B., RIEGER, L., FRENGER, M. S., MARSCHALL, A., FRANKE, R. B., PATTATHIL, S. & VOLL, L. M. 2017. Cell wall composition and penetration resistance against the fungal pathogen *Colletotrichum higginsianum* are affected by impaired starch turnover in Arabidopsis mutants. *Journal of Experimental Botany,* 68**,** 701-713.

ENGLER, C., KANDZIA, R. & MARILLONNET, S. 2008. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *Plos One,* 3.

ENGLER, C., YOULES, M., GRUETZNER, R., EHNERT, T. M., WERNER, S., JONES, J. D. G., PATRON, N. J. & MARILLONNET, S. 2014. A Golden Gate Modular Cloning Toolbox for Plants. *Acs Synthetic Biology,* 3**,** 839-843.

ESCH, L., NGAI, Q. Y., BARCLAY, J. E., MCNELLY, R., HAYTA, S., SMEDLEY, M. A., SMITH, A. M. & SEUNG, D. 2023. Increasing amyloplast size in wheat endosperm through mutation of PARC6 affects starch granule morphology. *New Phytologist,* 240**,** 224-241.

FABREGAS, N. & FERNIE, A. R. 2022. The reliance of phytohormone biosynthesis on primary metabolite precursors. *Journal of Plant Physiology,* 268.

FANG, Y. F. & TYLER, B. M. 2016. Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Molecular Plant Pathology,* 17**,** 127-139.

FANNON, J. E., HAUBER, R. J. & BEMILLER, J. N. 1992. Surface pores of starch granules. *Cereal Chemistry,* 69**,** 284-288.

FAULKNER, C., PETUTSCHNIG, E., BENITEZ-ALFONSO, Y., BECK, M., ROBATZEK, S., LIPKA, V. & MAULE, A. J. 2013. LYM2-dependent chitin perception limits molecular flux via

plasmodesmata. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 9166-9170.

FEKETE, E., KARAFFA, L., SEIBOTH, B., ÉVA, F., KUBICEK, C. P. & FLIPPHI, M. 2012. Identification of a permease gene involved in lactose utilisation in Aspergillus nidulans. *Fungal Genetics and Biology,* 49**,** 415-425.

FERNÁNDEZ-CALVINO, L., GUZMÁN-BENITO, I., DEL TORO, F. J., DONAIRE, L., CASTRO-SANZ, A. B., RUÍZ-FERRER, V. & LLAVE, C. 2016. Activation of senescence-associated *Dark-inducible* (*DIN*) genes during infection contributes to enhanced susceptibility to plant viruses. *Molecular Plant Pathology,* 17**,** 3-15.

FERNANDEZ, J., MARROQUIN-GUZMAN, M. & WILSON, R. A. 2014. Evidence for a Transketolase-Mediated Metabolic Checkpoint Governing Biotrophic Growth in Rice Cells by the Blast Fungus *Magnaporthe oryzae*. *Plos Pathogens,* 10.

FERNANDEZ, J., WRIGHT, J. D., HARTLINE, D., QUISPE, C. F., MADAYIPUTHIYA, N. & WILSON, R. A. 2012. Principles of Carbon Catabolite Repression in the Rice Blast Fungus: Tps1, Nmr1-3, and a MATE-Family Pump Regulate Glucose Metabolism during Infection. *Plos Genetics,* 8.

FESEL, P. H. & ZUCCARO, A. 2016. β-glucan: Crucial component of the fungal cell wall and elusive MAMP in plants. *Fungal Genetics and Biology,* 90**,** 53-60.

FLOR, H. H. 1971. Current status of gene-for-gene concept. *Annual Review of Phytopathology,* 9**,** 275-296.

FLORES-PÉREZ, U. & JARVIS, P. 2013. Molecular chaperone involvement in chloroplast protein import. *Biochimica Et Biophysica Acta-Molecular Cell Research,* 1833**,** 332-340.

FONES, H. N., BEBBER, D. P., CHALONER, T. M., KAY, W. T., STEINBERG, G. & GURR, S. J. 2020. Threats to global food security from emerging fungal and oomycete crop pathogens *Nature Food,* 1**,** 332–342.

FORDERER, A., LI, E., LAWSON, A. W., DENG, Y. N., SUN, Y., LOGEMANN, E., ZHANG, X. X., WEN, J., HAN, Z. F., CHANG, J. B., CHEN, Y. H., SCHULZE-LEFERT, P. & CHAI, J. J. 2022. A wheat resistosome defines common principles of immune receptor channels. *Nature,* 610**,** 532–539.

FORMELA-LUBOINSKA, M., CHADZINIKOLAU, T., DRZEWIECKA, K., JELEN, H., BOCIANOWSKI, J., KESY, J., LABUDDA, M., JEANDET, P. & MORKUNAS, I. 2020. The Role of Sugars in the Regulation of the Level of Endogenous Signaling Molecules during Defense Response of Yellow Lupine to *Fusarium oxysporum*. *International Journal of Molecular Sciences,* 21.

FOSTER, A. J., MARTIN-URDIROZ, M., YAN, X., WRIGHT, H. S., SOANES, D. M. & TALBOT, N. J. 2018. CRISPR-Cas9 ribonucleoprotein-mediated co-editing and counterselection in the rice blast fungus. *Scientific Reports,* 8.

FUJIKI, Y., YOSHIKAWA, Y., SATO, T., INADA, N., ITO, M., NISHIDA, I. & WATANABE, A. 2001. Dark-inducible genes from *Arabidopsis thaliana* are associated with leaf senescence and repressed by sugars. *Physiologia Plantarum,* 111**,** 345-352.

FUKADA, F., KODAMA, S., NISHIUCHI, T., KAJIKAWA, N. & KUBO, Y. 2019. Plant pathogenic fungi *Colletotrichum* and *Magnaporthe* share a common G(1) phase monitoring strategy for proper appressorium development. *New Phytologist,* 222**,** 1909-1923.

FURZER, O. J., CEVIK, V., FAIRHEAD, S., BAILEY, K., REDKAR, A., SCHUDOMA, C., MACLEAN, D., HOLUB, E. B. & JONES, J. D. G. 2022. An Improved Assembly of the *Albugo candida* Ac2V Genome Reveals the Expansion of the "CCG" Class of Effectors. *Molecular Plant-Microbe Interactions,* 35**,** 39-48.

GABRIEL, L., BRŮNA, T., HOFF, K. J., EBEL, M., LOMSADZE, A., BORODOVSKY, M. & STANKE, M. 2023. BRAKER3: Fully Automated Genome Annotation Using RNA-Seq and Protein Evidence with GeneMark-ETP, AUGUSTUS and TSEBRA. *bioRxiv*.

GAMEZ-ARJONA, F. M., RAYNAUD, S., RAGEL, P. & MERIDA, A. 2014. Starch synthase 4 is located in the thylakoid membrane and interacts with plastoglobule-associated proteins in Arabidopsis. *Plant Journal,* 80**,** 305-316.

GANESAN, I., SHI, L. X., LABS, M. & THEG, S. M. 2018. Evaluating the Functional Pore Size of Chloroplast TOC and TIC Protein Translocons: Import of Folded Proteins. *Plant Cell,* 30**,** 2161-2173.

GAO, Q. M., ZHU, S. F., KACHROO, P. & KACHROO, A. 2015. Signal regulators of systemic acquired resistance. *Frontiers in Plant Science,* 6.

GARCIA, M., MYOUGA, F., TAKECHI, K., SATO, H., NABESHIMA, K., NAGATA, N., TAKIO, S., SHINOZAKI, K. & TAKANO, H. 2008. An Arabidopsis homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant Journal,* 53**,** 924-934.

GASPAROTO, M. C. G., LOURENCO, S. A., TANAKA, F. A. O., SPOSITO, M. B., MARCHINI, L. C., SILVA, G. J. & AMORIM, L. 2017. Honeybees can spread *Colletotrichum acutatum* and *C. gloeosporioides* among citrus plants. *Plant Pathology,* 66**,** 777-782.

GEBAUER, P., KORN, M., ENGELSDORF, T., SONNEWALD, U., KOCH, C. & VOLL, L. M. 2017. Sugar Accumulation in Leaves of Arabidopsis *sweet11/sweet12* Double Mutants Enhances Priming of the Salicylic Acid-Mediated Defense Response. *Frontiers in Plant Science,* 8.

GHAREEB, H., EL-SAYED, M., POUND, M., TETYUK, O., HANIKA, K., HERRFURTH, C., FEUSSNER, I. & LIPKA, V. 2020. Quantitative Hormone Signaling Output Analyses of *Arabidopsis thaliana* Interactions With Virulent and Avirulent *Hyaloperonospora arabidopsidis* Isolates at Single-Cell Resolution. *Frontiers in Plant Science,* 11.

GIMENEZ-IBANEZ, S., BOTER, M., FERNANDEZ-BARBERO, G., CHINI, A., RATHJEN, J. P. & SOLANO, R. 2014. The Bacterial Effector HopX1 Targets JAZ Transcriptional Repressors to Activate Jasmonate Signaling and Promote Infection in Arabidopsis. *Plos Biology,* 12.

GIRALDO, M. C., DAGDAS, Y. F., GUPTA, Y. K., MENTLAK, T. A., YI, M., MARTINEZ-ROCHA, A. L., SAITOH, H., TERAUCHI, R., TALBOT, N. J. & VALENT, B. 2013. Two distinct secretion systems facilitate tissue invasion by the rice blast fungus *Magnaporthe oryzae*. *Nature Communications,* 4.

GISLASON, M. H., NIELSEN, H., ARMENTEROS, J. J. A. & JOHANSEN, A. R. 2021. Prediction of GPI-anchored proteins with pointer neural networks. *Current Research in Biotechnology,* 3**,** 6-13.

GLAZEBROOK, J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology,* 43**,** 205-227.

GODFREY, D., BOHLENIUS, H., PEDERSEN, C., ZHANG, Z. G., EMMERSEN, J. & THORDAL-CHRISTENSEN, H. 2010. Powdery mildew fungal effector candidates share N-terminal Y/F/WxC-motif. *Bmc Genomics,* 11.

GOMEZ-ARIZA, J., CAMPO, S., RUFAT, M., ESTOPA, M., MESSEGUER, J., SAN SEGUNDO, B. & COCA, M. 2007. Sucrose-mediated priming of plant defense responses and broad-spectrum disease resistance by overexpression of the maize pathogenesis-related PRms protein in rice plants. *Molecular Plant-Microbe Interactions,* 20**,** 832-842.

GOMEZ-GOMEZ, L. & BOLLER, T. 2000. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular Cell,* 5**,** 1003-1011.

GONZALEZ-FERNANDEZ, R. & JORRIN-NOVO, J. V. 2012. Contribution of Proteomics to the Study of Plant Pathogenic Fungi. *Journal of Proteome Research,* 11**,** 3-16.

GONZALEZ-MULA, A., LACHAT, J., MATHIAS, L., NAQUIN, D., LAMOUCHE, F., MERGAERT, P. & FAURE, D. 2018. The biotroph *Agrobacterium tumefaciens* thrives in tumors by exploiting a wide spectrum of plant host metabolites. *New Phytologist,* 222**,** 455-467.

GONZALEZ, P., REYES-DE-CORCUERA, J. & ETXEBERRIA, E. 2012. Characterization of leaf starch from HLB-affected and unaffected-girdled citrus trees. *Physiological and Molecular Plant Pathology,* 79**,** 71-78.

GOULD, S. J., KELLER, G. A. & SUBRAMANI, S. 1987. Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *Journal of Cell Biology,* 105**,** 2923-2931.

GRAF, A., SCHLERETH, A., STITT, M. & SMITH, A. M. 2010. Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 9458-9463.

GU, B. G., PARKES, T., RABANAL, F., SMITH, C., LU, F. H., MCKENZIE, N., DONG, H., WEIGEL, D., JONES, J. D. G., CEVIK, V. & BEVAN, M. W. 2023. The integrated LIM-peptidase domain of the CSA1-CHS3/DAR4 paired immune receptor detects changes in DA1 peptidase inhibitors in Arabidopsis. *Cell Host & Microbe,* 31**,** 949-961.

GU, Q. N., YUAN, Q. F., ZHAO, D., HUANG, J. B., HSIANG, T., WEI, Y. D. & ZHENG, L. 2019. Acetyl-coenzyme A synthetase gene *ChAcs1* is essential for lipid metabolism, carbon utilization and virulence of the hemibiotrophic fungus *Colletotrichum higginsianum*. *Molecular Plant Pathology,* 20**,** 107-123.

GUO, C. Y., WU, G. H., XING, J., LI, W. Q., TANG, D. Z. & CUI, B. M. 2013. A mutation in a coproporphyrinogen III oxidase gene confers growth inhibition, enhanced powdery mildew resistance and powdery mildew-induced cell death in Arabidopsis. *Plant Cell Reports,* 32**,** 687-702.

GUO, H. J., LIU, Y. C., LI, X., YAN, Z. H., XIE, Y. D., XIONG, H. C., ZHAO, L. S., GU, J. Y., ZHAO, S. R. & LIU, L. X. 2017. Novel mutant alleles of the starch synthesis gene *TaSSIVb-D* result in the reduction of starch granule number per chloroplast in wheat. *Bmc Genomics,* 18.

GUO, M., ZHU, X. L., LI, H. X., TAN, L. Y. & PAN, Y. M. 2016. Development of a novel strategy for fungal transformation based on a mutant locus conferring carboxin-resistance in *Magnaporthe oryzae*. *Amb Express,* 6.

GUPTA, P. K. 2020. *SWEET* Genes for Disease Resistance in Plants. *Trends in Genetics,* 36**,** 901-904.

GUPTA, P. K., BALYAN, H. S. & GAUTAM, T. 2021. *SWEET* genes and TAL effectors for disease resistance in plants: Present status and future prospects. *Molecular Plant Pathology,* 22**,** 1014-1026.

HAAS, B. J., KAMOUN, S., ZODY, M. C., JIANG, R. H. Y., HANDSAKER, R. E., CANO, L. M., GRABHERR, M., KODIRA, C. D., RAFFAELE, S., TORTO-ALALIBO, T., BOZKURT, T. O., AH-FONG, A. M. V., ALVARADO, L., ANDERSON, V. L., ARMSTRONG, M. R., AVROVA, A., BAXTER, L., BEYNON, J., BOEVINK, P. C., BOLLMANN, S. R., BOS, J. I. B., BULONE, V., CAI, G. H., CAKIR, C., CARRINGTON, J. C., CHAWNER, M., CONTI, L., COSTANZO, S., EWAN, R., FAHLGREN, N., FISCHBACH, M. A., FUGELSTAD, J., GILROY, E. M., GNERRE, S., GREEN, P. J., GRENVILLE-BRIGGS, L. J., GRIFFITH, J., GRUNWALD, N. J., HORN, K., HORNER, N. R., HU, C. H., HUITEMA, E., JEONG, D. H., JONES, A. M. E., JONES, J. D. G., JONES, R. W., KARLSSON, E. K., KUNJETI, S. G., LAMOUR, K., LIU, Z. Y., MA, L. J., MACLEAN, D., CHIBUCOS, M. C., MCDONALD, H., MCWALTERS, J., MEIJER, H. J. G., MORGAN, W., MORRIS, P. F., MUNRO, C. A., O'NEILL, K., OSPINA-GIRALDO, M., PINZON, A., PRITCHARD, L., RAMSAHOYE, B., REN, Q. H., RESTREPO, S., ROY, S., SADANANDOM, A., SAVIDOR, A., SCHORNACK, S., SCHWARTZ, D. C., SCHUMANN, U. D., SCHWESSINGER, B., SEYER, L., SHARPE, T., SILVAR, C., SONG, J., STUDHOLME, D. J., SYKES, S., THINES, M., VAN DE VONDERVOORT, P. J. I., PHUNTUMART, V., WAWRA, S., WEIDE, R., WIN, J., YOUNG, C., ZHOU, S. G., FRY, W., MEYERS, B. C., VAN WEST, P., RISTAINO, J., GOVERS, F., BIRCH, P. R. J., WHISSON, S. C., JUDELSON, H. S. & NUSBAUM, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature,* 461**,** 393-398.

HALLGREN, J., TSIRIGOS, K. D., PEDERSEN, M. D., ARMENTEROS, J. J. A., MARCATILI, P., NIELSEN, H., KROGH, A. & WINTHER, O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*.

HAN, E. K., COTTY, F., SOTTAS, C., JIANG, H. & MICHELS, C. A. 1995. Characterization of *agt1* encoding a general α-glucoside transporter from *Saccharomyces*. *Molecular Microbiology,* 17**,** 1093-1107.

HANDFORD, M. G. & CARR, J. P. 2007. A defect in carbohydrate metabolism ameliorates symptom severity in virus-infected *Arabidopsis thaliana*. *Journal of General Virology,* 88**,** 337-341.

HANSON, J., HANSSEN, M., WIESE, A., HENDRIKS, M. & SMEEKENS, S. 2008. The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of *ASPARAGINE SYNTHETASE1* and *PROLINE DEHYDROGENASE2*. *Plant Journal,* 53**,** 935-949.

HARVEY, S. E. & CHENG, C. 2016. Methods for Characterization of Alternative RNA Splicing. *Methods Mol Biol,* 1402**,** 229-241.

HAWKINS, E., CHEN, J. W., WATSON-LAZOWSKI, A., AHN-JARVIS, J., BARCLAY, J. E., FAHY, B., HARTLEY, M., WARREN, F. J. & SEUNG, D. 2021. STARCH SYNTHASE 4 is required for normal starch granule initiation in amyloplasts of wheat endosperm. *New Phytologist,* 230**,** 2371-2386.

HE, C., HOLME, J. & ANTHONY, J. 2014. SNP genotyping: the KASP assay. *Methods in molecular biology (Clifton, N.J.),* 1145**,** 75-86.

HE, J. Q., YE, W. W., CHOI, D. S., WU, B. X., ZHAI, Y., GUO, B. D., DUAN, S. Y., WANG, Y. C., GAN, J. H., MA, W. B. & MA, J. B. 2019. Structural analysis of *Phytophthora* suppressor of RNA silencing 2 (PSR2) reveals a conserved modular fold contributing to virulence. *Proceedings of the National Academy of Sciences of the United States of America,* 116**,** 8054-8059.

HILL, J. E. & HEMMINGSEN, S. M. 2001. *Arabidopsis thaliana* type I and II chaperonins. *Cell Stress & Chaperones,* 6**,** 190-200.

HIRUMA, K., ONOZAWA-KOMORI, M., TAKAHASHI, F., ASAKURA, M., BEDNAREK, P., OKUNO, T., SCHULZE-LEFERT, P. & TAKANO, Y. 2010. Entry mode-dependent function of an indole glucosinolate pathway in Arabidopsis for nonhost resistance against anthracnose pathogens. *Plant Cell,* 22**,** 2429-2443.

HOFF, K. J., LANGE, S., LOMSADZE, A., BORODOVSKY, M. & STANKE, M. 2016. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics,* 32**,** 767-769.

HOFFMAN, T., SCHMIDT, J. S., ZHENG, X. Y. & BENT, A. F. 1999. Isolation of ethylene-insensitive soybean mutants that are altered in pathogen susceptibility and gene-for-gene disease resistance. *Plant Physiology,* 119**,** 935-949.

HOLT, C. & YANDELL, M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *Bmc Bioinformatics,* 12.

HOLUB, E. B., BEYNON, L. J. & CRUTE, I. R. 1994. Phenotypic and genotypic characterization of interactions between isolates of *Peronospora parasitica* and accessions of *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions,* 7**,** 223-239.

HONG, Y. H., CAI, R. L., GUO, J. Y., ZHONG, Z. H., BAO, J. D., WANG, Z. H., CHEN, X. F., ZHOU, J. & LU, G. D. 2021. Carbon catabolite repressor MoCreA is required for the asexual development and pathogenicity of the rice blast fungus. *Fungal Genetics and Biology,* 146.

HORIE, H., SUGATA, S. & ABE, Z. 1988. Studies on anthracnose of komatsuna, *Brassica rapa*. *Bull. Tokyo Metr. Agr. Exp. Sta.,* 21**,** 189-237.

HU, Y., FANG, L., CHEN, X. L., ZHONG, J. F., LI, M. Y. & WANG, K. 2021. LIQA: long-read isoform quantification and analysis. *Genome Biology,* 22.

HUANG, C., YU, Q. B., LV, R. H., YIN, Q. Q., CHEN, G. Y., XU, L. & YANG, Z. N. 2013. The Reduced Plastid-Encoded Polymerase-Dependent Plastid Gene Expression Leads to the Delayed Greening of the Arabidopsis *fln2* Mutant. *Plos One,* 8.

HUOT, B., YAO, J., MONTGOMERY, B. L. & HE, S. Y. 2014. Growth-Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness. *Molecular Plant,* 7**,** 1267-1287.

HUSER, A., TAKAHARA, H., SCHMALENBACH, W. & O'CONNELL, R. 2009. Discovery of Pathogenicity Genes in the Crucifer Anthracnose Fungus *Colletotrichum higginsianum*, Using Random Insertional Mutagenesis. *Molecular Plant-Microbe Interactions,* 22**,** 143-156.

ICHIKAWA, S., SAKATA, M., OBA, T. & KODAMA, Y. 2023. Fluorescein staining of chloroplast starch granules in living plants. *Plant Physiology*.

INOUE-KASHINO, N., KASHINO, Y., ORII, H., SATOH, K., TERASHIMA, I. & PAKRASI, H. B. 2011. S4 Protein Sll1252 Is Necessary for Energy Balancing in Photosynthetic Electron Transport in *Synechocystis* sp. PCC 6803. *Biochemistry,* 50**,** 329-339.

IRIEDA, H., INOUE, Y., MORI, M., YAMADA, K., OSHIKAWA, Y., SAITOH, H., UEMURA, A., TERAUCHI, R., KITAKURA, S., KOSAKA, A., SINGKARAVANIT-OGAWA, S. & TAKANO, Y. 2019. Conserved fungal effector suppresses PAMP-triggered immunity by targeting plant immune kinases. *Proceedings of the National Academy of Sciences of the United States of America,* 116**,** 496-505.

IRIEDA, H., MAEDA, H., AKIYAMA, K., HAGIWARA, A., SAITOH, H., UEMURA, A., TERAUCHI, R. & TAKANO, Y. 2014. *Colletotrichum orbiculare* Secretes Virulence Effectors to a Biotrophic Interface at the Primary Hyphal Neck via Exocytosis Coupled with SEC22-Mediated Traffic. *Plant Cell,* 26**,** 2265-2281.

ISEMER, R., MULISCH, M., SCHÄFER, A., KIRCHNER, S., KOOP, H. U. & KRUPINSKA, K. 2012. Recombinant Whirly1 translocates from transplastomic chloroplasts to the nucleus. *Febs Letters,* 586**,** 85-88.

ISHIKAWA, H., YASUZAWA, M., KOIKE, N., SANJAYA, A., MORIYAMA, S., NISHIZAWA, A., MATSUOKA, K., SASAKI, S., KAZAMA, Y., HAYASHI, Y., ABE, T., FUJIWARA, M. T. & ITOH, R. D. 2020. Arabidopsis PARC6 Is Critical for Plastid Morphogenesis in Pavement, Trichome, and Guard Cells in Leaf Epidermis. *Frontiers in Plant Science,* 10.

JACOBS, K. A., COLLINSRACIE, L. A., COLBERT, M., DUCKETT, M., GOLDENFLEET, M., KELLEHER, K., KRIZ, R., LAVALLIE, E. R., MERBERG, D., SPAULDING, V., STOVER, J., WILLIAMSON, M. J. & MCCOY, J. M. 1997. A genetic selection for isolating cDNAs encoding secreted proteins. *Gene,* 198**,** 289-296.

JANKIEWICZ, L. S., GUZICKA, M. & MARASEK-CIOLAKOWSKA, A. 2021. Anatomy and Ultrastructure of Galls Induced by *Neuroterus quercusbaccarum* (Hymenoptera: Cynipidae) on Oak Leaves (*Quercus robu*r). *Insects,* 12.

JAYAWARDENA, R. S., HYDE, K. D., DAMM, U., CAI, L., LIU, M., LI, X. H., ZHANG, W., ZHAO, W. S. & YAN, J. Y. 2016. Notes on currently accepted species of *Colletotrichum*. *Mycosphere,* 7**,** 1192-1260.

JELENSKA, J., VAN HAL, J. A. & GREENBERG, J. T. 2010. *Pseudomonas syringae* hijacks plant stress chaperone machinery for virulence. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 13177-13182.

JELENSKA, J., YAO, N., VINATZER, B. A., WRIGHT, C. M., BRODSKY, J. L. & GREENBERG, J. T. 2007. A J domain virulence effector of *Pseudomonas syringae* remodels host chloroplasts and suppresses defenses. *Current Biology,* 17**,** 499-508.

JENNINGS, J. 2021. *Colletotrichum higginsianum exploitation of host cell-to-cell connectivity.* Doctor of Philosophy, University of East Anglia.

JEONG, S. Y., ROSE, A. & MEIER, I. 2003. MFP1 is a thylakoid-associated, nucleoid-binding protein with a coiled-coil structure. *Nucleic Acids Research,* 31**,** 5175-5185.

JIANG, D. Y., TANG, R. J., SHI, Y. F., KE, X. S., WANG, Y. T., CHE, Y. F., LUAN, S. & HOU, X. 2020. Arabidopsis Seedling Lethal 1 Interacting With Plastid-Encoded RNA Polymerase Complex Proteins Is Essential for Chloroplast Development. *Frontiers in Plant Science,* 11.

JIANG, R. H. Y., TRIPATHY, S., GOVERS, F. & TYLER, B. M. 2008. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 4874-4879.

JIANG, S. S., YAO, J., MA, K. W., ZHOU, H. B., SONG, J. K., HE, S. Y. & MA, W. B. 2013. Bacterial Effector Activates Jasmonate Signaling by Directly Targeting JAZ Transcriptional Repressors. *Plos Pathogens,* 9.

JOHNSON, J. M., REICHELT, M., VADASSERY, J., GERSHENZON, J. & OELMÜLLER, R. 2014. An Arabidopsis mutant impaired in intracellular calcium elevation is sensitive to biotic and abiotic stress. *Bmc Plant Biology,* 14.

JOHNSTON, M. G. & FAULKNER, C. 2021. A bootstrap approach is a superior statistical method for the comparison of non-normal data with differing variances. *New Phytologist,* 230**,** 23-26.

JONES, P., BINNS, D., CHANG, H. Y., FRASER, M., LI, W. Z., MCANULLA, C., MCWILLIAM, H., MASLEN, J., MITCHELL, A., NUKA, G., PESSEAT, S., QUINN, A. F., SANGRADOR-VEGAS, A., SCHEREMETJEW, M., YONG, S. Y., LOPEZ, R. & HUNTER, S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics,* 30**,** 1236-1240.

JORGENSEN, R. A. & DORANTES-ACOSTA, A. E. 2012. Conserved peptide upstream open reading frames are associated with regulatory genes in angiosperms. *Frontiers in Plant Science,* 3.

JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ZIDEK, A., POTAPENKO, A., BRIDGLAND, A., MEYER, C., KOHL, S. A. A., BALLARD, A. J., COWIE, A., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., ADLER, J., BACK, T., PETERSEN, S., REIMAN, D., CLANCY, E., ZIELINSKI, M., STEINEGGER, M., PACHOLSKA, M., BERGHAMMER, T., BODENSTEIN, S., SILVER, D., VINYALS, O., SENIOR,

A. W., KAVUKCUOGLU, K., KOHLI, P. & HASSABIS, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature,* 596**,** 583–589.

KAGDA, M. S., MARTINEZ-SOTO, D., AH-FONG, A. M. V. & JUDELSON, H. S. 2020. Invertases in *Phytophthora infestans* Localize to Haustoria and Are Programmed for Infection-Specific Expression. *Mbio,* 11.

KALE, S. D., GU, B. A., CAPELLUTO, D. G. S., DOU, D. L., FELDMAN, E., RUMORE, A., ARREDONDO, F. D., HANLON, R., FUDAL, I., ROUXEL, T., LAWRENCE, C. B., SHAN, W. X. & TYLER, B. M. 2010. External Lipid PI3P Mediates Entry of Eukaryotic Pathogen Effectors into Plant and Animal Host Cells. *Cell,* 142**,** 284-295.

KAMAL, H., LYNCH-HOLM, V., PAPPU, H. & TANAKA, K. 2023. Starch plays a key role in sporosorus formation by the powdery scab pathogen *Spongospora subterranea*. *Phytopathology*.

KAMBLE, N. U., MAKHAMADJONOV, F., FAHY, B., MARTINS, C., SAALBACH, G. & SEUNG, D. 2023. Initiation of B-type starch granules in wheat endosperm requires the plastidial α-glucan phosphorylase PHS1. *The Plant Cell,* 35**,** 4091-4110.

KAMIMURA, Y., TANAKA, H., KOBAYASHI, Y., SHIKANAI, T. & NISHIMURA, Y. 2018. Chloroplast nucleoids as a transformable network revealed by live imaging with a microfluidic device. *Communications Biology,* 1.

KANNEGANTI, T. D., HUITEMA, E., CAKIR, C. & KAMOUN, S. 2006. Synergistic interactions of the plant cell death pathways induced by *Phytophthora infestans* Nep1-like protein PiNPP1.1 and INF1 elicitin. *Molecular Plant-Microbe Interactions,* 19**,** 854-863.

KARIMI, M., INZÉ, D. & DEPICKER, A. 2002. GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci,* 7**,** 193-195.

KARLIK, E. 2021. Potential stress tolerance roles of barley germins and GLPs. *Development Genes and Evolution,* 231**,** 109-118.

KE, X. L., ZOU, W. X., REN, Y. F., WANG, Z. Q., LI, J., WU, X. & ZHAO, J. 2017. Functional divergence of chloroplast Cpn60 α subunits during Arabidopsis embryo development. *Plos Genetics,* 13.

KEELEY, M., ROWLAND, D. & VINCENT, C. 2022. Citrus photosynthesis and morphology acclimate to phloem-affecting huanglongbing disease at the leaf and shoot levels. *Physiologia Plantarum,* 174.

KELLEY, B. S., LEE, S. J., DAMASCENO, C. M. B., CHAKRAVARTHY, S., KIM, B. D., MARTIN, G. B. & ROSE, J. K. C. 2010. A secreted effector protein (SNE1) from *Phytophthora infestans* is a broadly acting suppressor of programmed cell death. *Plant Journal,* 62**,** 357-366.

KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. E. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols,* 10**,** 845-858.

KEMEN, E., GARDINER, A., SCHULTZ-LARSEN, T., KEMEN, A. C., BALMUTH, A. L., ROBERT-SEILANIANTZ, A., BAILEY, K., HOLUB, E., STUDHOLME, D. J., MACLEAN, D. & JONES, J. D. G. 2011. Gene Gain and Loss during Evolution of Obligate Parasitism in the White Rust Pathogen of *Arabidopsis thaliana*. *Plos Biology,* 9.

KERSHAW, M. J. & TALBOT, N. J. 2009. Genome-wide functional analysis reveals that infection-associated fungal autophagy is necessary for rice blast disease. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 15967-15972.

KEYMER, A., PIMPRIKAR, P., WEWER, V., HUBER, C., BRANDS, M., BUCERIUS, S. L., DELAUX, P. M., KLINGL, V., VON ROPENACK-LAHAYE, E., WANG, T. L., EISENREICH, W., DORMANN, P., PARNISKE, M. & GUTJAHR, C. 2017. Lipid transfer from plants to arbuscular mycorrhiza fungi. *Elife,* 6.

KHANG, C. H., BERRUYER, R., GIRALDO, M. C., KANKANALA, P., PARK, S. Y., CZYMMEK, K., KANG, S. & VALENT, B. 2010. Translocation of *Magnaporthe oryzae* Effectors into Rice Cells and Their Subsequent Cell-to-Cell Movement. *Plant Cell,* 22**,** 1388-1403.

KHODADADI, F., GONZÁLEZ, J. B., MARTIN, P. L., GIROUX, E., BILODEAU, G. J., PETER, K. A., DOYLE, V. P. & ACIMOVIC, S. G. 2020. Identification and characterization of *Colletotrichum* species causing apple bitter rot in New York and description of *C. noveboracense* sp. nov. *Scientific Reports,* 10.

KILARU, S., SCHUSTER, M., LATZ, M., DAS GUPTA, S., STEINBERG, N., FONES, H., GURR, S. J., TALBOT, N. J. & STEINBERG, G. 2015. A gene locus for targeted ectopic gene integration in *Zymoseptoria tritici*. *Fungal Genetics and Biology,* 79**,** 118-124.

KIM, D., PAGGI, J. M., PARK, C., BENNETT, C. & SALZBERG, S. L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology,* 37**,** 907–915.

KIM, S. G., WANG, Y., LEE, K. H., PARK, Z. Y., PARK, J., WU, J., KWON, S. J., LEE, Y. H., AGRAWAL, G. K., RAKWAL, R., KIM, S. T. & KANG, K. Y. 2013. In-depth insight into in vivo apoplastic secretome of rice-*Magnaporthe oryzae* interaction. *Journal of Proteomics,* 78**,** 58-71.

KLASEK, L., INOUE, K. & THEG, S. M. 2020. Chloroplast Chaperonin-Mediated Targeting of a Thylakoid Membrane Protein. *Plant Cell,* 32**,** 3884-3901.

KLEEMANN, J., RINCON-RIVERA, L. J., TAKAHARA, H., NEUMANN, U., VAN THEMAAT, E. V. L., VAN DER DOES, H. C., HACQUARD, S., STUBER, K., WILL, I., SCHMALENBACH, W., SCHMELZER, E. & O'CONNELL, R. J. 2012. Sequential Delivery of Host-Induced Virulence Effectors by Appressoria and Intracellular Hyphae of the Phytopathogen *Colletotrichum higginsianum*. *Plos Pathogens,* 8.

KOBAYASHI, K., OTEGUI, M. S., KRISHNAKUMAR, S., MINDRINOS, M. & ZAMBRYSKI, P. 2007. *INCREASED SIZE EXCLUSION LIMIT2* encodes a putative DEVH box RNA helicase involved in plasmodesmata function during Arabidopsis embryogenesis. *Plant Cell,* 19**,** 1885-1897.

KOCH, E. & SLUSARENKO, A. 1990. Arabidopsis is susceptible to infection by a downy mildew fungus. *Plant Cell,* 2**,** 437-445.

KONG, L., QIU, X. F., KANG, J. G., WANG, Y., CHEN, H., HUANG, J., QIU, M., ZHAO, Y., KONG, G. H., MA, Z. C., WANG, Y., YE, W. W., DONG, S. M., MA, W. B. & WANG, Y. C. 2017. A *Phytophthora* Effector Manipulates Host Histone Acetylation and Reprograms Defense Gene Expression to Promote Infection. *Current Biology,* 27**,** 981-991.

KORN, M., SCHMIDPETER, J., DAHL, M., MULLER, S., VOLL, L. M. & KOCH, C. 2015. A Genetic Screen for Pathogenicity Genes in the Hemibiotrophic Fungus *Colletotrichum higginsianum* Identifies the Plasma Membrane Proton Pump Pma2 Required for Host Penetration. *Plos One,* 10.

KOTTING, O., SANTELIA, D., EDNER, C., EICKE, S., MARTHALER, T., GENTRY, M. S., COMPAROT-MOSS, S., CHEN, J., SMITH, A. M., STEUP, M., RITTE, G. & ZEEMAN, S. C. 2009. STARCH-EXCESS4 Is a Laforin-Like Phosphoglucan Phosphatase Required for Starch Degradation in *Arabidopsis thaliana*. *Plant Cell,* 21**,** 334-346.

KOUSSEVITZKY, S., NOTT, A., MOCKLER, T. C., HONG, F., SACHETTO-MARTINS, G., SURPIN, M., LIM, I. J., MITTLER, R. & CHORY, J. 2007. Signals from chloroplasts converge to regulate nuclear gene expression. *Science,* 316**,** 715-719.

KRASILEVA, K. V., ZHENG, C., LEONELLI, L., GORITSCHNIG, S., DAHLBECK, D. & STASKAWICZ, B. J. 2011. Global Analysis of Arabidopsis/Downy Mildew Interactions Reveals Prevalence of Incomplete Resistance and Rapid Evolution of Pathogen Recognition. *Plos One,* 6.

KRAUSE, D. R., WOOD, C. J. & MACLEAN, D. J. 1991. Glucoamylase (exo-1,4-α-d-glucan glucanohydrolase, EC 3.2.1.3) is the major starch-degrading enzyme secreted by the phytopathogenic fungus *Colletotrichum gloeosporioides*. *Journal of General Microbiology,* 137**,** 2463-2468.

KRIEGER, N., PASTRYK, K. F., FORCHHAMMER, K. & KOLUKISAOGLU, U. 2021. Arabidopsis PII Proteins Form Characteristic Foci in Chloroplasts Indicating Novel Properties in Protein Interaction and Degradation. *International Journal of Molecular Sciences,* 22.

KROJ, T., CHANCLUD, E., MICHEL-ROMITI, C., GRAND, X. & MOREL, J. B. 2016. Integration of decoy domains derived from protein targets of pathogen effectors into plant immune receptors is widespread. *New Phytologist,* 210**,** 618-626.

KRUEGER, F. 2015. *Trim Galore!: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data* [Online]. Babraham Institute. Available: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed].

KUMAKURA, N., UENO, A. & SHIRASU, K. 2019. Establishment of a selection marker recycling system for sequential transformation of the plant-pathogenic fungus *Colletotrichum orbiculare*. *Molecular Plant Pathology,* 20**,** 447-459.

KUNZE, M. & BERGER, J. 2015. The similarity between N-terminal targeting signals for protein import into different organelles and its evolutionary relevance. *Frontiers in Physiology,* 6.

KUPPIREDDY, V. S., UVERSKY, V. N., TOH, S. S., TSAI, M. C., BECKERSON, W. C., CAHILL, C., CARMAN, B. & PERLIN, M. H. 2017. Identification and Initial Characterization of the Effectors of an Anther Smut Fungus and Potential Host Target Proteins. *International Journal of Molecular Sciences,* 18.

KUROIWA, T., SUZUKI, T., OGAWA, K. & KAWANO, S. 1981. The chloroplast nucleus - distribution, number, size, and shape, and a model for the multiplication of the chloroplast genome during chloroplast development. *Plant and Cell Physiology,* 22**,** 381-396.

KUZNETSOVA, A., BROCKHOFF, P. B. & CHRISTENSEN, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software,* 82**,** 1-26.

KUZNIAK, E. & KOPCZEWSKI, T. 2020. The Chloroplast Reactive Oxygen Species-Redox System in Plant Immunity and Disease. *Frontiers in Plant Science,* 11.

LAI, C. C., QUE, Q. X., PAN, R., WANG, Q., GAO, H. Y., GUAN, X. F., CHE, J. M. & LAI, G. T. 2022. The Single-Stranded DNA-Binding Gene *Whirly* (*Why1*) with a Strong Pathogen-Induced Promoter from *Vitis pseudoreticulata* Enhances Resistance to *Phytophthora capsici*. *International Journal of Molecular Sciences,* 23.

LAI, Z. B., LI, Y., WANG, F., CHENG, Y., FAN, B. F., YU, J. Q. & CHEN, Z. X. 2011. Arabidopsis Sigma Factor Binding Proteins Are Activators of the WRKY33 Transcription Factor in Plant Defense. *Plant Cell,* 23**,** 3824-3841.

LAL, N. K., NAGALAKSHMI, U., HURLBURT, N. K., FLORES, R., BAK, A., SONE, P., MA, X. Y., SONG, G. Y., WALLEY, J., SHAN, L. B., HE, P., CASTEEL, C., FISHER, A. J. & DINESH-KUMAR, S. P. 2018. The Receptor-like Cytoplasmic Kinase BIK1 Localizes to the Nucleus and Regulates Defense Hormone Expression during Plant Innate Immunity. *Cell Host & Microbe,* 23**,** 485-497.

LAMB, C. & DIXON, R. A. 1997. The oxidative burst in plant disease resistance. *Annual Review of Plant Physiology and Plant Molecular Biology,* 48**,** 251-275.

LAMESCH, P., BERARDINI, T. Z., LI, D. H., SWARBRECK, D., WILKS, C., SASIDHARAN, R., MULLER, R., DREHER, K., ALEXANDER, D. L., GARCIA-HERNANDEZ, M., KARTHIKEYAN, A. S., LEE, C. H., NELSON, W. D., PLOETZ, L., SINGH, S., WENSEL, A. & HUALA, E. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research,* 40**,** D1202-D1210.

LEE, S., JOUNG, Y. H., KIM, J. K., CHOI, Y. D. & JANG, G. 2019. An isoform of the plastid RNA polymerase-associated protein FSD3 negatively regulates chloroplast development. *Bmc Plant Biology,* 19.

LEE, S. J. & ROSE, J. K. 2010. Mediation of the transition from biotrophy to necrotrophy in hemibiotrophic plant pathogens by secreted effector proteins. *Plant Signal Behav,* 5**,** 769-772.

LEE, S. J. & ROSE, J. K. 2012. A yeast secretion trap assay for identification of secreted proteins from eukaryotic phytopathogens and their plant hosts. *Methods Mol Biol,* 835**,** 519-530.

LEI, J., TENG, X., WANG, Y. F., JIANG, X. K., ZHAO, H. H., ZHENG, X. M., REN, Y. L., DONG, H., WANG, Y. L., DUAN, E. C., ZHANG, Y. Y., ZHANG, W. W., YANG, H., CHEN, X. L., CHEN, R. B., ZHANG, Y., YU, M. Z., XU, S. B., BAO, X. H., ZHANG, P. C., LIU, S. J., LIU, X., TIAN, Y. L., JIANG, L., WANG, Y. H. & WAN, J. M. 2022. Plastidic pyruvate dehydrogenase complex E1 component subunit α1 is involved in galactolipid biosynthesis required for amyloplast development in rice. *Plant Biotechnology Journal,* 20**,** 437-453.

LEISTER, D. 2014. Complex(iti)es of the ubiquitous RNA-binding CSP41 proteins. *Frontiers in Plant Science,* 5.

LENTH, V. R. 2023. *emmeans: Estimated Marginal Means, aka Least-Squares Means* [Online]. Available: https://CRAN.R-project.org/package=emmeans [Accessed 2023].

LI, H.-T., ZHANG, W., ZHU, H., CHAO, C. & GUO, Q. 2023. Unlocking the Potential of High-Amylose Starch for Gut Health: Not All Function the Same. *Fermentation,* 9.

LIEBERS, M., GILLET, F. X., ISRAEL, A., POUNOT, K., CHAMBON, L., CHIEB, M., CHEVALIER, F., RUEDAS, R., FAVIER, A., GANS, P., ERBA, E. B., COBESSI, D., PFANNSCHMIDT, T. & BLANVILLAIN, R. 2020. Nucleo-plastidic PAP8/pTAC6 couples chloroplast formation with photomorphogenesis. *Embo Journal,* 39.

LIN, T. P., CASPAR, T., SOMERVILLE, C. & PREISS, J. 1988. Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* (L.) heynh lacking ADPglucose pyrophosphorylase activity. *Plant Physiology,* 86**,** 1131-1135.

LINDSAY, R. J., KERSHAW, M. J., PAWLOWSKA, B. J., TALBOT, N. J. & GUDELJ, I. 2016. Harbouring public good mutants within a pathogen population can increase both fitness and virulence. *Elife,* 5.

LING, Q. H., BROAD, W., TROSCH, R., TOPEL, M., SERT, T. D., LYMPEROPOULOS, P., BALDWIN, A. & JARVIS, R. P. 2019. Ubiquitin-dependent chloroplast-associated protein degradation in plants. *Science,* 363.

LING, Q. H. & JARVIS, P. 2015. Regulation of Chloroplast Protein Import by the Ubiquitin E3 Ligase SP1 Is Important for Stress Tolerance in Plants. *Current Biology,* 25**,** 2527-2534.

LINGNER, U., MUNCH, S., DEISING, H. B. & SAUER, N. 2011. Hexose Transporters of a Hemibiotrophic Plant Pathogen: FUNCTIONAL VARIATIONS AND REGULATORY DIFFERENCES AT DIFFERENT STAGES OF INFECTION. *Journal of Biological Chemistry,* 286**,** 20913-20922.

LINKS, M. G., HOLUB, E., JIANG, R. H. Y., SHARPE, A. G., HEGEDUS, D., BEYNON, E., SILLITO, D., CLARKE, W. E., UZUHASHI, S. & BORHAN, M. H. 2011. *De novo* sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. *Bmc Genomics,* 12.

LIU, G. S., JI, Y. Y., BHUIYAN, N. H., PILOT, G., SELVARAJ, G., ZOU, J. T. & WEI, Y. D. 2010. Amino Acid Homeostasis Modulates Salicylic Acid-Associated Redox Status and Defense Responses in Arabidopsis. *Plant Cell,* 22**,** 3845-3863.

LIU, L. P., ZHAO, D., ZHENG, L., HSIANG, T., WEI, Y. D., FU, Y. P. & HUANG, J. B. 2013. Identification of virulence genes in the crucifer anthracnose fungus *Colletotrichum higginsianum* by insertional mutagenesis. *Microbial Pathogenesis,* 64**,** 6-17.

LIU, Q. T., ZHOU, Y. & FETTKE, J. 2021. Starch Granule Size and Morphology of *Arabidopsis thaliana* Starch-Related Mutants Analyzed during Diurnal Rhythm and Development. *Molecules,* 26.

LIU, T. L., SONG, T. Q., ZHANG, X., YUAN, H. B., SU, L. M., LI, W. L., XU, J., LIU, S. H., CHEN, L. L., CHEN, T. Z., ZHANG, M. X., GU, L. C., ZHANG, B. L. & DOU, D. L. 2014. Unconventionally secreted effectors of two filamentous pathogens target plant salicylate biosynthesis. *Nature Communications,* 5.

LIU, X., BELLANDI, A., JOHNSTON, M. G. & FAULKNER, C. 2022. The *Hyaloperanospora arabidopsidis* effector HaRxL77 is hypermobile between cells and manipulates host defence. *bioRxiv*.

LIU, Y. D., REN, D. T., PIKE, S., PALLARDY, S., GASSMANN, W. & ZHANG, S. Q. 2007. Chloroplast-generated reactive oxygen species are involved in hypersensitive response-like cell death mediated by a mitogen-activated protein kinase cascade. *Plant Journal,* 51**,** 941-954.

LIU, Y. X., LAN, X., SONG, S. R., YIN, L., DRY, I. B., QU, J. J., XIANG, J. & LU, J. 2018. *In Planta* Functional Analysis and Subcellular Localization of the Oomycete Pathogen *Plasmopara viticola* Candidate RXLR Effector Repertoire. *Frontiers in Plant Science,* 9.

LO PRESTI, L., ZECHMANN, B., KUMLEHN, J., LIANG, L., LANVER, D., TANAKA, S., BOCK, R. & KAHMANN, R. 2017. An assay for entry of secreted fungal effectors into plant cells. *New Phytologist,* 213**,** 956-964.

LOEWITH, R. & HALL, M. N. 2011. Target of Rapamycin (TOR) in Nutrient Signaling and Growth Control. *Genetics,* 189**,** 1177-1201.

LORRAIN, C., KAMOUN, S. & PETRE, B. 2014. *Populus trichocarpa* Coproporphyrinogen III oxidase (PtCPO) localizes in sub-chloroplastic structures. figshare.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology,* 15.

LUDWIG, N., REISSMANN, S., SCHIPPER, K., GONZALEZ, C., ASSMANN, D., GLATTER, T., MORETTI, M., MA, L. S., REXER, K. H., SNETSELAAR, K. & KAHMANN, R. 2021. A cell surface-exposed protein complex with an essential virulence function in *Ustilago maydis*. *Nature Microbiology,* 6**,** 722-730.

LUNDQUIST, P. K., MANTEGAZZA, O., STEFANSKI, A., STUHLER, K. & WEBER, A. P. M. 2017. Surveying the Oligomeric State of *Arabidopsis thaliana* Chloroplasts. *Molecular Plant,* 10**,** 197-211.

LUNDQUIST, P. K., POLIAKOV, A., BHUIYAN, N. H., ZYBAILOV, B., SUN, Q. & VAN WIJK, K. J. 2012. The Functional Network of the Arabidopsis Plastoglobule Proteome Based on Quantitative Proteomics and Genome-Wide Coexpression Analysis. *Plant Physiology,* 158**,** 1172-1192.

MA, Z. C., ZHU, L., SONG, T. Q., WANG, Y., ZHANG, Q., XIA, Y. Q., QIU, M., LIN, Y. C., LI, H. Y., KONG, L., FANG, Y. F., YE, W. W., WANG, Y., DONG, S. M., ZHENG, X. B., TYLER, B. M. & WANG, Y. C. 2017. A paralogous decoy protects *Phytophthora sojae* apoplastic effector PsXEG1 from a host inhibitor. *Science,* 355**,** 710-714.

MACHO, A. P. & ZIPFEL, C. 2014. Plant PRRs and the Activation of Innate Immune Signaling. *Molecular Cell,* 54**,** 263-272.

MACIOSZEK, V. K., JECZ, T., CIERESZKO, I. & KONONOWICZ, A. K. 2023. Jasmonic Acid as a Mediator in Plant Response to Necrotrophic Fungi. *Cells,* 12.

MADEIRA, F., PEARCE, M., TIVEY, A. R. N., BASUTKAR, P., LEE, J., EDBALI, O., MADHUSOODANAN, N., KOLESNIKOV, A. & LOPEZ, R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research,* 50**,** W276-W279.

MAJERAN, W., FRISO, G., ASAKURA, Y., QU, X., HUANG, M. S., PONNALA, L., WATKINS, K. P., BARKAN, A. & VAN WIJK, K. J. 2012. Nucleoid-Enriched Proteomes in Developing Plastids and Chloroplasts from Maize Leaves: A New Conceptual Framework for Nucleoid Functions. *Plant Physiology,* 158**,** 156-189.

MARCELINO, J. A. P., GOULI, S., PARKER, B. L., SKINNER, M. & GIORDANO, R. 2009. Entomopathogenic activity of a variety of the fungus, *Colletotrichum acutatum*, recovered from the elongate hemlock scale, Fiorinia externa. *Journal of Insect Science,* 9.

MARSHALL, R., KOMBRINK, A., MOTTERAM, J., LOZA-REYES, E., LUCAS, J., HAMMOND-KOSACK, K. E., THOMMA, B. & RUDD, J. J. 2011. Analysis of Two in Planta Expressed LysM Effector Homologs from the Fungus *Mycosphaerella graminicola* Reveals Novel Functional Properties and Varying Contributions to Virulence on Wheat. *Plant Physiology,* 156**,** 756-769.

MARTIN, C. 2017. *ggConvexHull* [Online]. Available: https://github.com/cmartin/ggConvexHull [Accessed].

MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal,* 17**,** 10-12.

MCDOWELL, J. M. 2014. *Hyaloperonospora arabidopsidis*: A Model Pathogen of Arabidopsis. *In:* DEAN, R. A., LICHENS-PARK, A. & KOLE, C. (eds.) *Genomics of Plant-Associated Fungi and Oomycetes: Dicot Pathogens.* Berlin, Heidelberg: Springer Berlin Heidelberg.

MCLELLAN, H., HARVEY, S. E., STEINBRENNER, J., ARMSTRONG, M. R., HE, Q., CLEWES, R., PRITCHARD, L., WANG, W., WANG, S. H., NUSSBAUMER, T., DOHAI, B., LUO, Q. Q., KUMARI, P., DUAN, H., ROBERTS, A., BOEVINK, P. C., NEUMANN, C., CHAMPOURET, N., HEIN, I., FALTER-BRAUN, P., BEYNON, J., DENBY, K. & BIRCH, P. R. J. 2022. Exploiting breakdown in nonhost effector-target interactions to boost host disease resistance. *Proceedings of the National Academy of Sciences of the United States of America,* 119.

MELIDA, H., SANDOVAL-SIERRA, J. V., DIEGUEZ-URIBEONDO, J. & BULONE, V. 2013. Analyses of Extracellular Carbohydrates in Oomycetes Unveil the Existence of Three Different Cell Wall Types. *Eukaryotic Cell,* 12**,** 194-203.

MELONEK, J., MULISCH, M., SCHMITZ-LINNEWEBER, C., GRABOWSKI, E., HENSEL, G. & KRUPINSKA, K. 2010. Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids. *Planta,* 232**,** 471-481.

MELOTTO, M., UNDERWOOD, W., KOCZAN, J., NOMURA, K. & HE, S. Y. 2006. Plant stomata function in innate immunity against bacterial invasion. *Cell,* 126**,** 969-980.

MÉTEIGNIER, L. V., GHANDOUR, R., ZIMMERMAN, A., KUHN, L., MEURER, J., ZOSCHKE, R. & HAMMANI, K. 2021. Arabidopsis mTERF9 protein promotes chloroplast ribosomal assembly and translation by establishing ribonucleoprotein interactions *in vivo*. *Nucleic Acids Research,* 49**,** 1114–1132.

MIMS, C. W., RICHARDSON, E. A., HOLT III, B. F. & DANGL, J. L. 2004. Ultrastructure of the host-pathogen interface in *Arabidopsis thaliana* leaves infected by the downy mildew *Hyaloperonospora parasitica Canadian Journal of Botany-Revue Canadienne De Botanique,* 82**,** 1545-1545.

MIMS, C. W. & VAILLANCOURT, L. J. 2002. Ultrastructural characterization of infection and colonization of maize leaves by *Colletotrichum graminicola*, and by a *C. graminicola* pathogenicity mutant. *Phytopathology,* 92**,** 803-812.

MIRDITA, M., SCHUTZE, K., MORIWAKI, Y., HEO, L., OVCHINNIKOV, S. & STEINEGGER, M. 2022. ColabFold: making protein folding accessible to all. *Nature Methods,* 19**,** 679–682.

MIYA, A., ALBERT, P., SHINYA, T., DESAKI, Y., ICHIMURA, K., SHIRASU, K., NARUSAKA, Y., KAWAKAMI, N., KAKU, H. & SHIBUYA, N. 2007. CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 19613-19618.

MIYAGISHIMA, S. Y., WOLK, C. P. & OSTERYOUNG, K. W. 2005. Identification of cyanobacterial cell division genes by comparative and mutational analyses. *Molecular Microbiology,* 56**,** 126-143.

MU-FORSTER, C., HUANG, R. M., POWERS, J. R., HARRIMAN, R. W., KNIGHT, M., SINGLETARY, G. W., KEELING, P. L. & WASSERMAN, B. P. 1996. Physical association of starch biosynthetic enzymes with starch granules of maize endosperm - Granule-associated forms of starch synthase I and starch branching enzyme II. *Plant Physiology,* 111**,** 821-829.

MULLINEAUX, P. M., EXPOSITO-RODRIGUEZ, M., LAISSUE, P. P., SMIRNOFF, N. & PARK, E. 2020. Spatial chloroplast-to-nucleus signalling involving plastid-nuclear complexes and stromules. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 375.

MURAKAMI, R., USHIMA, R., SUGIMOTO, R., TAMAOKI, D., KARAHARA, I., HANBA, Y., WAKASUGI, T. & TSUCHIDA, T. 2021. A new galling insect model enhances photosynthetic activity in an obligate holoparasitic plant. *Scientific Reports,* 11.

MYOUGA, F., HOSODA, C., UMEZAWA, T., IIZUMI, H., KUROMORI, T., MOTOHASHI, R., SHONO, Y., NAGATA, N., IKEUCHI, M. & SHINOZAKI, K. 2008. A Heterocomplex of Iron Superoxide Dismutases Defends Chloroplast Nucleoids against Oxidative Stress and Is Essential for Chloroplast Development in Arabidopsis. *Plant Cell,* 20**,** 3148-3162.

NAESTED, H., HOLM, A., JENKINS, T., NIELSEN, H. B., HARRIS, C. A., BEALE, M. H., ANDERSEN, M., MANT, A., SCHELLER, H., CAMARA, B., MATTSSON, O. & MUNDY, J. 2004. Arabidopsis *VARIEGATED 3* encodes a chloroplasttargeted, zinc-finger protein required for chloroplast and palisade cell development. *Journal of Cell Science,* 117**,** 4807-4818.

NARUSAKA, M., SHIRAISHI, T., IWABUCHI, M. & NARUSAKA, Y. 2010. Monitoring fungal viability and development in plants infected with *Colletotrichum higginsianum* by quantitative reverse transcription-polymerase chain reaction. *Journal of General Plant Pathology,* 76**,** 1-6.

NARUSAKA, M., SHIRASU, K., NOUTOSHI, Y., KUBO, Y., SHIRAISHI, T., IWABUCHI, M. & NARUSAKA, Y. 2009. *RRS1* and *RPS4* provide a dual *Resistance*-gene system against fungal and bacterial pathogens. *Plant Journal,* 60**,** 218-226.

NAWRATH, C., HECK, S., PARINTHAWONG, N. & METRAUX, J. P. 2002. EDS5, an essential component of salicylic acid-dependent signaling for disease resistance in Arabidopsis, is a member of the MATE transporter family. *Plant Cell,* 14**,** 275-286.

NAWRATH, C. & METRAUX, J. P. 1999. Salicylic acid induction-deficient mutants of Arabidopsis express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen inoculation. *Plant Cell,* 11**,** 1393-1404.

NEGI, N. P., PRAKASH, G., NARWAL, P., PANWAR, R., KUMAR, D., CHAUDHRY, B. & RUSTAGI, A. 2023. The calcium connection: exploring the intricacies of calcium signaling in plant-microbe interactions. *Frontiers in Plant Science,* 14.

NEVAREZ, P. A., QIU, Y. J., INOUE, H., YOO, C. Y., BENFEY, P. N., SCHNELL, D. J. & CHEN, M. 2017. Mechanism of Dual Targeting of the Phytochrome Signaling Component HEMERA/pTAC12 to Plastids and the Nucleus. *Plant Physiology,* 173**,** 1953-1966.

NGOU, B. P. M., AHN, H. K., DING, P. T. & JONES, J. D. G. 2021. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature*.

NIELSEN, H. & KROGH, A. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol,* 6**,** 122-130.

NISHIO, K., HIROHASHI, T. & NAKAI, M. 1999. Chloroplast chaperonins: Evidence for heterogeneous assembly of α and β Cpn60 polypeptides into a chaperonin oligomer. *Biochemical and Biophysical Research Communications,* 266**,** 584-587.

NIU, J. H., LIU, P., LIU, Q., CHEN, C. L., GUO, Q. X., JUNMEIYIN, GUANGSUIYANG & JIAN, H. 2016. Msp40 effector of root-knot nematode manipulates plant immunity to facilitate parasitism. *Scientific Reports,* 6.

NOMURA, H., KOMORI, T., UEMURA, S., KANDA, Y., SHIMOTANI, K., NAKAI, K., FURUICHI, T., TAKEBAYASHI, K., SUGIMOTO, T., SANO, S., SUWASTIKA, I. N., FUKUSAKI, E., YOSHIOKA, H., NAKAHIRA, Y. & SHIINA, T. 2012. Chloroplast-mediated activation of plant immune signalling in Arabidopsis. *Nature Communications,* 3.

O'CONNELL, R., HERBERT, C., SREENIVASAPRASAD, S., KHATIB, M., ESQUERRE-TUGAYE, M. T. & DUMAS, B. 2004. A novel Arabidopsis-*Colletotrichum* pathosystem for the molecular dissection of plant-fungal interactions. *Molecular Plant-Microbe Interactions,* 17**,** 272-282.

O'CONNELL, R. J., THON, M. R., HACQUARD, S., AMYOTTE, S. G., KLEEMANN, J., TORRES, M. F., DAMM, U., BUIATE, E. A., EPSTEIN, L., ALKAN, N., ALTMULLER, J., ALVARADO-BALDERRAMA, L., BAUSER, C. A., BECKER, C., BIRREN, B. W., CHEN, Z. H., CHOI, J., CROUCH, J. A., DUVICK, J. P., FARMAN, M. A., GAN, P., HEIMAN, D., HENRISSAT, B., HOWARD, R. J., KABBAGE, M., KOCH, C., KRACHER, B., KUBO, Y., LAW, A. D., LEBRUN, M. H., LEE, Y. H., MIYARA, I., MOORE, N., NEUMANN, U., NORDSTROM, K., PANACCIONE, D. G., PANSTRUGA, R., PLACE, M., PROCTOR, R. H., PRUSKY, D., RECH, G., REINHARDT, R., ROLLINS, J. A., ROUNSLEY, S., SCHARDL, C. L., SCHWARTZ, D. C., SHENOY, N., SHIRASU, K., SIKHAKOLLI, U. R., STUBER, K., SUKNO, S. A., SWEIGARD, J. A., TAKANO, Y., TAKAHARA, H., TRAIL, F., VAN DER DOES, H. C., VOLL, L. M., WILL, I., YOUNG, S., ZENG, Q. D., ZHANG, J. Z., ZHOU, S. G., DICKMAN, M. B., SCHULZE-LEFERT, P., VAN THEMAAT, E. V. L., MA, L. J. & VAILLANCOURT, L. J. 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics,* 44**,** 1060–1065.

O'KEEFFE, K. R. & JONES, C. D. 2019. Challenges and solutions for analysing dual RNA-seq data for non-model host-pathogen systems. *Methods in Ecology and Evolution,* 10**,** 401-414.

OHTSU, M., JENNINGS, J., JOHNSTON, M. G., BREAKSPEAR, A., LIU, X., STARK, K., MORRIS, R., DE KEIJZER, J. & FAULKNER, C. 2023. Assaying effector cell-to-cell mobility in plant tissues identifies hypermobility and indirect manipulation of plasmodesmata. *Mol Plant Microbe Interact*.

OLIVEIRA-GARCIA, E., TAMANG, T. M., PARK, J., DALBY, M., MARTIN-URDIROZ, M., HERRERO, C. R., VU, A. H., PARK, S., TALBOT, N. J. & VALENT, B. 2023. Clathrin-mediated endocytosis facilitates the internalization of *Magnaporthe oryzae* effectors into rice cells. *Plant Cell,* 35**,** 2527-2551.

OMENGE, K. M., RUEMPLER, F., KATHALINGAM, S. S., FURCH, A. C. U. & THEISSEN, G. 2021. Studying the Function of Phytoplasma Effector Proteins Using a Chemical-Inducible Expression System in Transgenic Plants. *International Journal of Molecular Sciences,* 22.

PARK, E., LEE, H. Y., WOO, J., CHOI, D. & DINESH-KUMAR, S. P. 2017. Spatiotemporal Monitoring of *Pseudomonas syringae* Effectors via Type III Secretion Using Split Fluorescent Protein Fragments. *Plant Cell,* 29**,** 1571-1584.

PATKAR, R. N., BENKE, P. I., QU, Z. W., CHEN, Y. Y. C., YANG, F., SWARUP, S. & NAQVI, N. I. 2015. A fungal monooxygenase-derived jasmonate attenuates host innate immunity. *Nature Chemical Biology,* 11**,** 733–740.

PATRON, N., ORZAEZ, D., MARILLONNET, S., WARZECHA, H., MATTHEWMAN, C., YOULES, M., RAITSKIN, O., LEVEAU, A., FARRÉ, G., ROGERS, C., SMITH, A., HIBBERD, J., WEBB, A. A. R.,

LOCKE, J., SCHORNACK, S., AJIOKA, J., BAULCOMBE, D. C., ZIPFEL, C., KAMOUN, S., JONES, J. D. G., KUHN, H., ROBATZEK, S., VAN ESSE, H. P., SANDERS, D., OLDROYD, G., MARTIN, C., FIELD, R., O'CONNOR, S., FOX, S., WULFF, B., MILLER, B., BREAKSPEAR, A., RADHAKRISHNAN, G., DELAUX, P. M., LOQUÉ, D., GRANELL, A., TISSIER, A., SHIH, P., BRUTNELL, T. P., QUICK, W. P., RISCHER, H., FRASER, P. D., AHARONI, A., RAINES, C., SOUTH, P. F., ANÉ, J. M., HAMBERGER, B. R., LANGDALE, J., STOUGAARD, J., BOUWMEESTER, H., UDVARDI, M., MURRAY, J. A. H., NTOUKAKIS, V., SCHÄFER, P., DENBY, K., EDWARDS, K. J., OSBOURN, A. & HASELOFF, J. 2015. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytologist,* 208**,** 13-19.

PAYSAN-LAFOSSE, T., BLUM, M., CHUGURANSKY, S., GREGO, T., PINTO, B. L., SALAZAR, G. A., BILESCHI, M. L., BORK, P., BRIDGE, A., COLWELL, L., GOUGH, J., HAFT, D. H., LETUNIC, I., MARCHLER-BAUER, A., MI, H. Y., NATALE, D. A., ORENGO, C. A., PANDURANGAN, A. P., RIVOIRE, C., SIGRIST, C. J. A., SILLITOE, I., THANKI, N., THOMAS, P. D., TOSATTO, S. C. E., WU, C. H. & BATEMAN, A. 2022. InterPro in 2022. *Nucleic Acids Research,* 51**,** D418–D427.

PERFECT, S. E., HUGHES, H. B., O'CONNELL, R. J. & GREEN, J. R. 1999. *Colletotrichum* - A model genus for studies on pathology and fungal-plant interactions. *Fungal Genetics and Biology,* 27**,** 186-198.

PERTEA, M., PERTEA, G. M., ANTONESCU, C. M., CHANG, T. C., MENDELL, J. T. & SALZBERG, S. L. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology,* 33**,** 290–295.

PETERSEN, T. N., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods,* 8**,** 785-786.

PETRE, B. & KAMOUN, S. 2014. How Do Filamentous Pathogens Deliver Effector Proteins into Plant Cells? *Plos Biology,* 12.

PETRE, B., LORRAIN, C., SAUNDERS, D. G. O., WIN, J., SKLENAR, J., DUPLESSIS, S. & KAMOUN, S. 2016a. Rust fungal effectors mimic host transit peptides to translocate into chloroplasts. *Cellular Microbiology,* 18**,** 453-465.

PETRE, B., SAUNDERS, D. G. O., SKLENAR, J., LORRAIN, C., KRASILEVA, K. V., WIN, J., DUPLESSIS, S. & KAMOUN, S. 2016b. Heterologous Expression Screens in *Nicotiana benthamiana* Identify a Candidate Effector of the Wheat Yellow Rust Pathogen that Associates with Processing Bodies. *Plos One,* 11.

PETRE, B., SAUNDERS, D. G. O., SKLENAR, J., LORRAIN, C., WIN, J., DUPLESSIS, S. & KAMOUN, S. 2015. Candidate Effector Proteins of the Rust Pathogen *Melampsora larici-populina* Target Diverse Plant Cell Compartments. *Molecular Plant-Microbe Interactions,* 28**,** 689-700.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. R. C., MENG, E. E. C., COUCH, G. S., CROLL, T. I., MORRIS, J. H. & FERRIN, T. E. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science,* 30**,** 70-82.

PFALZ, J., LIERE, K., KANDLBINDER, A., DIETZ, K. J. & OELMÜLLER, R. 2006. PTAC2,-6, and-12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell,* 18**,** 176-197.

PFISTER, B. & ZEEMAN, S. C. 2016. Formation of starch in plant cells. *Cellular and Molecular Life Sciences,* 73**,** 2781-2807.

PIERLEONI, A., MARTELLI, P. L. & CASADIO, R. 2008. PredGPI: a GPI-anchor predictor. *Bmc Bioinformatics,* 9.

PITINO, M., ALLEN, V. & DUAN, Y. P. 2018. Las Δ5315 Effector Induces Extreme Starch Accumulation and Chlorosis as *Ca*. Liberibacter asiaticus Infection in *Nicotiana benthamiana*. *Frontiers in Plant Science,* 9.

PLAUMANN, P. L., SCHMIDPETER, J., DAHL, M., TAHER, L. & KOCH, C. 2018. A Dispensable Chromosome Is Required for Virulence in the Hemibiotrophic Plant Pathogen *Colletotrichum higginsianum*. *Frontiers in Microbiology,* 9.

PUNT, P. J., DINGEMANSE, M. A., KUYVENHOVEN, A., SOEDE, R. D. M., POUWELS, P. H. & VANDENHONDEL, C. 1990. Functional elements in the promoter region of the *Aspergillus-Nidulans gpdA* gene encoding Glyceraldehyde-3-Phosphate Dehydrogenase. *Gene,* 93**,** 101-109.

PUNT, P. J., ZEGERS, N. D., BUSSCHER, M., POUWELS, P. H. & VANDENHONDEL, C. 1991. Intracellular and extracellular production of proteins in *Aspergillus* under the control of expression signals of the highly expressed *Aspergillus nidulans gpdA* gene. *Journal of Biotechnology,* 17**,** 19-33.

QI, T., GUO, J., LIU, P., HE, F. X., WAN, C. P., ISLAM, M. A., TYLER, B. M., KANG, Z. S. & GUO, J. 2019. Stripe Rust Effector PstGSRE1 Disrupts Nuclear Localization of ROS-Promoting Transcription Factor TaLOL2 to Defeat ROS-Induced Defense in Wheat. *Molecular Plant,* 12**,** 1624-1638.

R CORE TEAM. 2021. *R: A language and environment for statistical computing*. [Online]. Vienna, Austria. Available: https://www.R-project.org/ [Accessed].

RAFFAELE, S., FARRER, R. A., CANO, L. M., STUDHOLME, D. J., MACLEAN, D., THINES, M., JIANG, R. H. Y., ZODY, M. C., KUNJETI, S. G., DONOFRIO, N. M., MEYERS, B. C., NUSBAUM, C. & KAMOUN, S. 2010. Genome Evolution Following Host Jumps in the Irish Potato Famine Pathogen Lineage. *Science,* 330**,** 1540-1543.

RAFFEINER, M., USTUN, S., GUERRA, T., SPINTI, D., FITZNER, M., SONNEWALD, S., BALDERMANN, S. & BORNKE, F. 2022. The *Xanthomonas* type-III effector XopS stabilizes CaWRKY40a to regulate defense responses and stomatal immunity in pepper (*Capsicum annuum*). *Plant Cell,* 34**,** 1684-1708.

RAJARAMMOHAN, S. 2021. Redefining Plant-Necrotroph Interactions: The Thin Line Between Hemibiotrophs and Necrotrophs. *Frontiers in Microbiology,* 12.

RAMSDEN, D. A. & GELLERT, M. 1998. Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *Embo Journal,* 17**,** 609-614.

REHMANY, A. P., GORDON, A., ROSE, L. E., ALLEN, R. L., ARMSTRONG, M. R., WHISSON, S. C., KAMOUN, S., TYLER, B. M., BIRCH, P. R. J. & BEYNON, J. L. 2005. Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. *Plant Cell,* 17**,** 1839-1850.

RICCI, P., BONNET, P., HUET, J. C., SALLANTIN, M., BEAUVAISCANTE, F., BRUNETEAU, M., BILLARD, V., MICHEL, G. & PERNOLLET, J. C. 1989. Structure and activity of proteins from pathogenic fungi phytophthora eliciting necrosis and acquired-resistance in tobacco. *European Journal of Biochemistry,* 183**,** 555-563.

RICHTER, S. & LAMPPA, G. K. 1998. A chloroplast processing enzyme functions as the general stromal processing peptidase. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 7463-7468.

RITTE, G., HEYDENREICH, M., MAHLOW, S., HAEBEL, S., KOTTING, O. & STEUP, M. 2006. Phosphorylation of C6- and C3-positions of glucosyl residues in starch is catalysed by distinct dikinases. *Febs Letters,* 580**,** 4872-4876.

ROBATZEK, S., CHINCHILLA, D. & BOLLER, T. 2006. Ligand-induced endocytosis of the pattern recognition receptor FLS2 in Arabidopsis. *Genes & Development,* 20**,** 537-542.

ROBIN, G. P., KLEEMANN, J., NEUMANN, U., CABRE, L., DALLERY, J. F., LAPALU, N. & O'CONNELL, R. J. 2018. Subcellular Localization Screening of *Colletotrichum higginsianum* Effector Candidates Identifies Fungal Proteins Targeted to Plant Peroxisomes, Golgi Bodies, and Microtubules. *Frontiers in Plant Science,* 9.

ROBINSON, S. M. & BOSTOCK, R. M. 2015. β-glucans and eicosapolyenoic acids as MAMPs in plant-oomycete interactions: past and present. *Frontiers in Plant Science,* 5.

RODENBURG, S. Y. A., SEIDL, M. F., DE RIDDER, D. & GOVERS, F. 2018. Genome-wide characterization of *Phytophthora infestans* metabolism: a systems biology approach. *Molecular Plant Pathology,* 19**,** 1403-1413.

ROJAS, C. M., SENTHIL-KUMAR, M., TZIN, V. & MYSORE, K. S. 2014. Regulation of primary plant metabolism during plant-pathogen interactions and its contribution to plant defense. *Frontiers in Plant Science,* 5.

ROLDAN, I., WATTEBLED, F., LUCAS, M. M., DELVALLE, D., PLANCHOT, V., JIMENEZ, S., PEREZ, R., BALL, S., D'HULST, C. & MERIDA, A. 2007. The phenotype of soluble starch synthase IV defective mutants of Arabidopsis thaliana suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant Journal,* 49**,** 492-504.

ROLLAND, F., BAENA-GONZALEZ, E. & SHEEN, J. 2006. Sugar sensing and signaling in plants: Conserved and novel mechanisms. *Annual Review of Plant Biology,* 57**,** 675-709.

ROMER, P., HAHN, S., JORDAN, T., STRAUSS, T., BONAS, U. & LAHAYE, T. 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science,* 318**,** 645-648.

ROMER, P., RECHT, S., STRAUSS, T., ELSAESSER, J., SCHORNACK, S., BOCH, J., WANG, S. P. & LAHAYE, T. 2010. Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist,* 187**,** 1048-1057.

RUTTER, B. D., CHU, T.-T.-H., DALLERY, J.-F., ZAJT, K. K., O'CONNELL, R. J. & INNES, R. W. 2022. The development of extracellular vesicle markers for the fungal phytopathogen *Colletotrichum higginsianum*. *Journal of Extracellular Vesicles,* 11**,** e12216.

SALCEDO, A. F., PURAYANNUR, S., STANDISH, J. R., MILES, T., THIESSEN, L. & QUESADA-OCAMPO, L. M. 2021. Fantastic Downy Mildew Pathogens and How to Find Them: Advances in Detection and Diagnostics. *Plants-Basel,* 10**,** 24.

SALZBERG, S. L. 2019. Next-generation genome annotation: we still struggle to get it right. *Genome Biology,* 20.

SAUNDERS, D. G. O., AVES, S. J. & TALBOT, N. J. 2010. Cell Cycle-Mediated Regulation of Plant Infection by the Rice Blast Fungus. *Plant Cell,* 22**,** 497-507.

SAVAGE, Z., DUGGAN, C., TOUFEXI, A., PANDEY, P., LIANG, Y. X., SEGRETIN, M. E., YUEN, L. H., GABORIAU, D. C. A., LEARY, A. Y., TUMTAS, Y., KHANDARE, V., WARD, A. D., BOTCHWAY, S. W., BATEMAN, B. C., PAN, I., SCHATTAT, M., SPARKES, I. & BOZKURT, T. O. 2021. Chloroplasts alter their morphology and accumulate at the pathogen interface during infection by *Phytophthora infestans*. *Plant Journal,* 107**,** 1771-1787.

SAVARY, S., WILLOCQUET, L., PETHYBRIDGE, S. J., ESKER, P., MCROBERTS, N. & NELSON, A. 2019. The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution,* 3**,** 430-439.

SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J.-Y., WHITE, D. J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. & CARDONA, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods,* 9**,** 676-682.

SCHMIDPETER, J., DAHL, M., HOFMANN, J. & KOCH, C. 2017. ChMob2 binds to ChCbk1 and promotes virulence and conidiation of the fungal pathogen *Colletotrichum higginsianum*. *Bmc Microbiology,* 17.

SCHORNACK, S., VAN DAMME, M., BOZKURT, T. O., CANO, L. M., SMOKER, M., THINES, M., GAULIN, E., KAMOUN, S. & HUITEMA, E. 2010. Ancient class of translocated oomycete effectors targets the host nucleus. *Proceedings of the National Academy of Sciences of the United States of America,* 107**,** 17421-17426.

SEDLAZECK, F. J., RESCHENEDER, P. & VON HAESELER, A. 2013. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics,* 29**,** 2790-2791.

SEUNG, D., BOUDET, J., MONROE, J., SCHREIER, T. B., DAVID, L. C., ABT, M., LU, K. J., ZANELLA, M. & ZEEMANA, S. C. 2017. Homologs of PROTEIN TARGETING TO STARCH Control Starch Granule Initiation in Arabidopsis Leaves. *Plant Cell,* 29**,** 1657–1677.

SEUNG, D., LU, K. J., STETTLER, M., STREB, S. & ZEEMAN, S. C. 2016. Degradation of Glucan Primers in the Absence of Starch Synthase 4 Disrupts Starch Granule Initiation in Arabidopsis. *Journal of Biological Chemistry,* 291**,** 20718-20728.

SEUNG, D., SCHREIER, T. B., BURGY, L., EICKE, S. & ZEEMAN, S. C. 2018. Two Plastidial Coiled-Coil Proteins Are Essential for Normal Starch Granule Initiation in Arabidopsis. *Plant Cell,* 30**,** 1523-1542.

SEUNG, D. & SMITH, A. M. 2019. Starch granule initiation and morphogenesis-progress in Arabidopsis and cereals. *Journal of Experimental Botany,* 70**,** 771-784.

SEUNG, D., SOYK, S., COIRO, M., MAIER, B. A., EICKE, S. & ZEEMAN, S. C. 2015. PROTEIN TARGETING TO STARCH Is Required for Localising GRANULE-BOUND STARCH SYNTHASE to Starch Granules and for Normal Amylose Synthesis in Arabidopsis. *Plos Biology,* 13.

SHAO, D. D., SMITH, D. L., KABBAGE, M. & ROTH, M. G. 2021. Effectors of Plant Necrotrophic Fungi. *Frontiers in Plant Science,* 12.

SHEVCHENKO, A., TOMAS, H., HAVLIS, J., OLSEN, J. V. & MANN, M. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols,* 1**,** 2856-2860.

SHI, L. X. & THEG, S. M. 2013. Energetic cost of protein import across the envelope membranes of chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 930-935.

SHIMA, Y., ITO, Y., KANEKO, S., HATABAYASHI, H., WATANABE, Y., ADACHI, Y. & YABE, K. 2009. Identification of three mutant loci conferring carboxin-resistance and development of a novel transformation system in *Aspergillus oryzae*. *Fungal Genetics and Biology,* 46**,** 67-76.

SHINYA, T., YAMAGUCHI, K., DESAKI, Y., YAMADA, K., NARISAWA, T., KOBAYASHI, Y., MAEDA, K., SUZUKI, M., TANIMOTO, T., TAKEDA, J., NAKASHIMA, M., FUNAMA, R., NARUSAKA, M., NARUSAKA, Y., KAKU, H., KAWASAKI, T. & SHIBUYA, N. 2014. Selective regulation of the chitin-induced defense response by the Arabidopsis receptor-like cytoplasmic kinase PBL27. *Plant Journal,* 79**,** 56-66.

SMITH, A. M. & ZEEMAN, S. C. 2006. Quantification of starch in plant tissues. *Nature Protocols,* 1**,** 1342-1345.

SMITH, A. M. & ZEEMAN, S. C. 2020. Starch: A Flexible, Adaptable Carbon Store Coupled to Plant Growth. *In:* MERCHANT, S. S. (ed.) *Annual Review of Plant Biology, Vol 71, 2020.*

SMITH, S. M., FULTON, D. C., CHIA, T., THORNEYCROFT, D., CHAPPLE, A., DUNSTAN, H., HYLTON, C., ZEEMAN, S. C. & SMITH, A. M. 2004. Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in arabidopsis leaves. *Plant Physiology,* 136**,** 2687-2699.

SOYLU, E. M. & SOYLU, S. 2003. Light and electron microscopy of the compatible interaction between Arabidopsis and the downy mildew pathogen *Peronospora parasitica*. *Journal of Phytopathology,* 151**,** 300-306.

SPERSCHNEIDER, J., CATANZARITI, A. M., DEBOER, K., PETRE, B., GARDINER, D. M., SINGH, K. B., DODDS, P. N. & TAYLOR, J. M. 2017. LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Scientific Reports,* 7.

SPERSCHNEIDER, J. & DODDS, P. N. 2022. EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes. *Molecular Plant-Microbe Interactions,* 35**,** 146-156.

SPERSCHNEIDER, J., DODDS, P. N., GARDINER, D. M., MANNERS, J. M., SINGH, K. B. & TAYLOR, J. M. 2015a. Advances and Challenges in Computational Prediction of Effectors from Plant Pathogenic Fungi. *Plos Pathogens,* 11.

SPERSCHNEIDER, J., WILLIAMS, A. H., HANE, J. K., SINGH, K. B. & TAYLOR, J. M. 2015b. Evaluation of Secretion Prediction Highlights Differing Approaches Needed for Oomycete and Fungal Effectors. *Frontiers in Plant Science,* 6.

STARK, R., GRZELAK, M. & HADFIELD, J. 2019. RNA sequencing: the teenage years. *Nature Reviews Genetics,* 20**,** 631-656.

STOTZ, H. U., MITROUSIA, G. K., DE WIT, P. & FITT, B. D. L. 2014. Effector-triggered defence against apoplastic fungal pathogens. *Trends in Plant Science,* 19**,** 491-500.

STREB, S. & ZEEMAN, S. C. 2012. Starch metabolism in Arabidopsis. *The arabidopsis book,* 10**,** e0160-e0160.

STREUBEL, J., PESCE, C., HUTIN, M., KOEBNIK, R., BOCH, J. & SZUREK, B. 2013. Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytologist,* 200**,** 808-819.

STUTTMANN, J., HUBBERTEN, H. M., RIETZ, S., KAUR, J., MUSKETT, P., GUEROIS, R., BEDNAREK, P., HOEFGEN, R. & PARKER, J. E. 2011. Perturbation of Arabidopsis Amino Acid Metabolism Causes Incompatibility with the Adapted Biotrophic Pathogen *Hyaloperonospora arabidopsidis*. *Plant Cell,* 23**,** 2788-2803.

SUGIO, A., KINGDOM, H. N., MACLEAN, A. M., GRIEVE, V. M. & HOGENHOUT, S. A. 2011. Phytoplasma protein effector SAP11 enhances insect vector reproduction by

manipulating plant development and defense hormone biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America,* 108**,** E1254-E1263.

SUJKA, M. & JAMROZ, J. 2010. Characteristics of pores in native and hydrolyzed starch granules. *Starch-Starke,* 62**,** 229-235.

SUZUKI, K., NAKANISHI, H., BOWER, J., YODER, D. W., OSTERYOUNG, K. W. & MIYAGISHIMA, S. Y. 2009. Plastid chaperonin proteins Cpn60 alpha and Cpn60 beta are required for plastid division in *Arabidopsis thaliana*. *Bmc Plant Biology,* 9.

TADASHI, T., MASUDA, T. & KOYAMA, Y. 2006. Enhanced gene targeting frequency in *ku70* and *ku80* disruption mutants of *Aspergillus sojae* and *Aspergillus oryzae*. *Molecular Genetics and Genomics,* 275**,** 460-470.

TAKAHARA, H., DOLF, A., ENDL, E. & O'CONNELL, R. 2009. Flow cytometric purification of *Colletotrichum higginsianum* biotrophic hyphae from Arabidopsis leaves for stage-specific transcriptome analysis. *Plant Journal,* 59**,** 672-683.

TAKAHARA, H., HACQUARD, S., KOMBRINK, A., HUGHES, H. B., HALDER, V., ROBIN, G. P., HIRUMA, K., NEUMANN, U., SHINYA, T., KOMBRINK, E., SHIBUYA, N., THOMMA, B. & O'CONNELL, R. J. 2016. *Colletotrichum higginsianum* extracellular LysM proteins play dual roles in appressorial function and suppression of chitin-triggered plant immunity. *New Phytologist,* 211**,** 1323-1337.

TAKAHARA, H., HUSER, A. & O CONNELL, R. 2012. Two arginine biosynthesis genes are essential for pathogenicity of *Colletotrichum higginsianum* on Arabidopsis. *Mycology,* 3**,** 54 - 64.

TAKAHARA, H., YAMAGUCHI, S., OMURA, N., NAKAJIMA, S., OTOKU, K., TANAKA, S., OGURA, K., KLEEMANN, J. & O'CONNELL, R. 2021. The *Colletotrichum higginsianum* secreted effector protein ChEC91 induces plant cell death. *Journal of General Plant Pathology,* 87**,** 344-353.

TANAKA, S., BREFORT, T., NEIDIG, N., DJAMEI, A., KAHNT, J., VERMERRIS, W., KOENIG, S., FEUSSNER, K., FEUSSNER, I. & KAHMANN, R. 2014. A secreted *Ustilago maydis* effector promotes virulence by targeting anthocyanin biosynthesis in maize. *Elife,* 3.

TANG, B., FENG, L., HULIN, M. T., DING, P. & MA, W. 2023. Cell-type-specific responses to fungal infection in plants revealed by single-cell transcriptomics. *Cell Host Microbe,* 31**,** 1732-1747.e5.

TANG, L. G., YANG, G. G., MA, M., LIU, X. F., LI, B., XIE, J. T., FU, Y. P., CHEN, T., YU, Y., CHEN, W. D., JIANG, D. H. & CHENG, J. S. 2020. An effector of a necrotrophic fungal pathogen targets the calcium-sensing receptor in chloroplasts to inhibit host resistance. *Molecular Plant Pathology,* 21**,** 686-701.

TANZ, S. K., CASTLEDEN, I., SMALL, I. D. & MILLAR, A. H. 2013. Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants. *Frontiers in Plant Science,* 4.

TECSI, L. I., WANG, D., SMITH, A. M., LEEGOOD, R. C. & MAULE, A. J. 1992. Red-clover mottle virus-infection affects sink source relationships and starch accumulation in pea-plants. *Journal of Experimental Botany,* 43**,** 1409-1412.

TEE, E. E., SAMWALD, S. & FAULKNER, C. 2022. Quantification of Cell-to-Cell Connectivity Using Particle Bombardment. *Methods Mol Biol.,* 2457**,** 263-272.

TEIXEIRA, P., THOMAZELLA, D. P. D., REIS, O., DO PRADO, P. F. V., DO RIO, M. C. S., FIORIN, G. L., JOSÉ, J., COSTA, G. G. L., NEGRI, V. A., MONDEGO, J. M. C., MIECZKOWSKI, P. & PEREIRA, G. A. G. 2014. High-Resolution Transcript Profiling of the Atypical Biotrophic Interaction between *Theobroma cacao* and the Fungal Pathogen *Moniliophthora perniciosa*. *Plant Cell,* 26**,** 4245-4269.

TEO, G. C., LIU, G. M., ZHANG, J. P., NESVIZHSKII, A. I., GINGRAS, A. C. & CHOI, H. 2014. *SAINTexpress*: Improvements and additional features in Significance Analysis of INTeractome software. *Journal of Proteomics,* 100**,** 37-43.

TERASAWA, K. & SATO, N. 2005. Visualization of plastid nucleoids in situ using the PEND-GFP fusion protein. *Plant and Cell Physiology,* 46**,** 649-660.

TEUFEL, F., ARMENTEROS, J. J. A., JOHANSEN, A. R., GISLASON, M. H., PIHL, S. I., TSIRIGOS, K. D., WINTHER, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology,* 40**,** 1023–1025.

THOMMA, B., EGGERMONT, K., TIERENS, K. & BROEKAERT, W. F. 1999. Requirement of functional ethylene-insensitive 2 gene for efficient resistance of Arabidopsis to infection by *Botrytis cinerea*. *Plant Physiology,* 121**,** 1093-1101.

THORVALDSDÓTTIR, H., ROBINSON, J. T. & MESIROV, J. P. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics,* 14**,** 178-192.

THUMULURI, V., ARMENTEROS, J. J. A., JOHANSEN, A. R., NIELSEN, H. & WINTHER, O. 2022. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Research,* 50**,** W228-W234.

TORRES, M. A., DANGL, J. L. & JONES, J. D. G. 2002. Arabidopsis gp91(phox) homologues AtrbohD and AtrbohF are required for accumulation of reactive oxygen intermediates in the plant defense response. *Proceedings of the National Academy of Sciences of the United States of America,* 99**,** 517-522.

TORRES, M. F., CUADROS, D. F. & VAILLANCOURT, L. J. 2014. Evidence for a diffusible factor that induces susceptibility in the *Colletotrichum*-maize disease interaction. *Molecular Plant Pathology,* 15**,** 80-93.

TORTO, T. A., LI, S. A., STYER, A., HUITEMA, E., TESTA, A., GOW, N. A. R., VAN WEST, P. & KAMOUN, S. 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Research,* 13**,** 1675-1685.

TRUMAN, W., DE ZABALA, M. T. & GRANT, M. 2006. Type III effectors orchestrate a complex interplay between transcriptional networks to modify basal defence responses during pathogenesis and resistance. *Plant Journal,* 46**,** 14-33.

TSUSHIMA, A., GAN, P., KUMAKURA, N., NARUSAKA, M., TAKANO, Y., NARUSAKA, Y. & SHIRASU, K. 2019a. Genomic Plasticity Mediated by Transposable Elements in the Plant Pathogenic Fungus *Colletotrichum higginsianum*. *Genome Biology and Evolution,* 11**,** 1487-1500.

TSUSHIMA, A., GAN, P. & SHIRASU, K. 2019b. Method for Assessing Virulence of *Colletotrichum higginsianum* on *Arabidopsis thaliana* Leaves Using Automated Lesion Area Detection and Measurement. *Bio-Protocol,* 9.

TSUSHIMA, A., NARUSAKA, M., GAN, P., KUMAKURA, N., HIROYAMA, R., KATO, N., TAKAHASHI, S., TAKANO, Y., NARUSAKA, Y. & SHIRASU, K. 2021. The Conserved *Colletotrichum* spp. Effector Candidate CEC3 Induces Nuclear Expansion and Cell Death in Plants. *Frontiers in Microbiology,* 12.

TURNER, R. S. 2005. After the famine: Plant pathology, *Phytophthora infestans*, and the late blight of potatoes. *Historical Studies in the Physical and Biological Sciences,* 35**,** 341-370.

TZELEPIS, G., DOLFORS, F., HOLMQUIST, L. & DIXELIUS, C. 2021. Plant mitochondria and chloroplasts are targeted by the *Rhizoctonia solani* RsCRP1 effector. *Biochemical and Biophysical Research Communications,* 544**,** 86-90.

UMEMOTO, N., KAKITANI, M., IWAMATSU, A., YOSHIKAWA, M., YAMAOKA, N. & ISHIDA, I. 1997. The structure and function of a soybean β-glucan-elicitor-binding protein. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 1029-1034.

USHIMARU, T., TERADA, H., TSUBOI, K., KOGOU, Y., SAKAGUCHI, A., TSUJI, G. & KUBO, Y. 2010. Development of an efficient gene targeting system in *Colletotrichum higginsianum* using a non-homologous end-joining mutant and *Agrobacterium tumefaciens*-mediated gene transfer. *Molecular Genetics and Genomics,* 284**,** 357-371.

VAN DER BIEZEN, E. A., FREDDIE, C. T., KAHN, K., PARKER, J. E. & JONES, J. D. G. 2002. Arabidopsis RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. *Plant Journal,* 29**,** 439-451.

VAN DER ENT, S., VAN WEES, S. C. M. & PIETERSE, C. M. J. 2009. Jasmonate signaling in plant interactions with resistance-inducing beneficial microbes. *Phytochemistry,* 70**,** 1581-1588.

VANDERVERE, P. S., BENNETT, T. M., OBLONG, J. E. & LAMPPA, G. K. 1995. A chloroplast processing enzyme involved in precursor maturation shares a zinc-binding motif with a recently recognized family of metalloendopeptidases. *Proceedings of the National Academy of Sciences of the United States of America,* 92**,** 7177-7181.

VANDROMME, C., SPRIET, C., DAUVILLEE, D., COURSEAUX, A., PUTAUX, J. L., WYCHOWSKI, A., KRZEWINSKI, F., FACON, M., D'HULST, C. & WATTEBLED, F. 2019. PII1: a protein involved in starch initiation that determines granule number and size in Arabidopsis chloroplast. *New Phytologist,* 221**,** 356-370.

VARGAS, W. A., SANZ-MARTÍN, J. M., RECH, G. E., ARMIJOS-JARAMILLO, V. D., RIVERA, L. P., ECHEVERRIA, M. M., DÍAZ-MÍNGUEZ, J. M., THON, M. R. & SUKNO, S. A. 2016. A Fungal Effector With Host Nuclear Localization and DNA-Binding Properties Is Required for Maize Anthracnose Development. *Molecular Plant-Microbe Interactions,* 29**,** 83-95.

VELHO, A. C., MONDINO, P. & STADNIK, M. J. 2018. Extracellular enzymes of *Colletotrichum fructicola* isolates associated to Apple bitter rot and Glomerella leaf spot. *Mycology,* 9**,** 145-154.

VILLALBA, F., COLLEMARE, J., LANDRAUD, P., LAMBOU, K., BROZEK, V., CIRER, B., MORIN, D., BRUEL, C., BEFFA, R. & LEBRUN, M. H. 2008. Improved gene targeting in *Magnaporthe grisea* by inactivation of *MgKU80* required for non-homologous end joining. *Fungal Genetics and Biology,* 45**,** 68-75.

VIRDI, K. S., WAMBOLDT, Y., KUNDARIYA, H., LAURIE, J. D., KEREN, I., KUMAR, K. R. S., BLOCK, A., BASSET, G., LUEBKER, S., ELOWSKY, C., DAY, P. M., ROOSE, J. L., BRICKER, T. M., ELTHON, T. & MACKENZIE, S. A. 2016. MSH1 Is a Plant Organellar DNA Binding and Thylakoid Protein under Precise Spatial Regulation to Alter Development. *Molecular Plant,* 9**,** 245-260.

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 Biology,* 8.

WANG, B., SUN, Y., SONG, N., ZHAO, M., LIU, R., FENG, H., WANG, X. & KANG, Z. 2017. *Puccinia striiformis* f. sp. *tritici* microRNA-like RNA 1 (Pst-milR1), an important pathogenicity factor of *Pst*, impairs wheat resistance to Pst by suppressing the wheat pathogenesis-related 2 gene. *New Phytol,* 215**,** 338-350.

WANG, G. X., ROUX, B., FENG, F., GUY, E., LI, L., LI, N. N., ZHANG, X. J., LAUTIER, M., JARDINAUD, M. F., CHABANNES, M., ARLAT, M., CHEN, S., HE, C. Z., NOEL, L. D. & ZHOU, J. M. 2015. The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified-Self Recognition and Immunity in Plants. *Cell Host & Microbe,* 18**,** 285-295.

WANG, H., OLIVEIRA-GARCIA, E., BOEVINK, P. C., TALBOT, N. J., BIRCH, P. R. J. & VALENT, B. 2023a. Filamentous pathogen effectors enter plant cells via endocytosis. *Trends Plant Sci,* 28**,** 1214-1217.

WANG, H. X., WANG, S. M., WANG, W., XU, L., WELSH, L. R. J., GIERLINSKI, M., WHISSON, S. C., HEMSLEY, P. A., BOEVINK, P. C. & BIRCH, P. R. J. 2023b. Uptake of oomycete RXLR effectors into host cells by clathrin-mediated endocytosis. *Plant Cell,* 35**,** 2504-2526.

WANG, J. Z., HU, M. J., WANG, J., QI, J. F., HAN, Z. F., WANG, G. X., QI, Y. J., WANG, H. W., ZHOU, J. M. & CHAI, J. J. 2019. Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science,* 364.

WANG, L. Y., PATENA, W., VAN BAALEN, K. A., XIE, Y. H., SINGER, E. R., GAVRILENKO, S., WARREN-WILLIAMS, M., HAN, L. Q., HARRIGAN, H. R., HARTZ, L. D., CHEN, V. V., TON, V., KYIN, S., SHWE, H. H., CAHN, M. H., WILSON, A. T., ONISHI, M., HU, J. P., SCHNELL, D. J., MCWHITE, C. D. & JONIKAS, M. C. 2023c. A chloroplast protein atlas reveals punctate structures and spatial organization of biosynthetic pathways. *Cell,* 186**,** 3499-3518.

WANKE, A., ROVENICH, H., SCHWANKE, F., VELTE, S., BECKER, S., HEHEMANN, J. H., WAWRA, S. & ZUCCARO, A. 2020. Plant species-specific recognition of long and short β-1,3-linked glucans is mediated by different receptor systems. *Plant Journal,* 102**,** 1142-1156.

WAWRA, S., BELMONTE, R., LOBACH, L., SARAIVA, M., WILLEMS, A. & VAN WEST, P. 2012. Secretion, delivery and function of oomycete effector proteins. *Current Opinion in Microbiology,* 15**,** 685-691.

WEIBERG, A., WANG, M., LIN, F. M., ZHAO, H. W., ZHANG, Z. H., KALOSHIAN, I., HUANG, H. D. & JIN, H. L. 2013. Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science,* 342**,** 118-123.

WESTERMANN, A. J., GORSKI, S. A. & VOGEL, J. 2012. Dual RNA-seq of pathogen and host. *Nature Reviews Microbiology,* 10**,** 618-630.

WHISSON, S. C., BOEVINK, P. C., MOLELEKI, L., AVROVA, A. O., MORALES, J. G., GILROY, E. M., ARMSTRONG, M. R., GROUFFAUD, S., VAN WEST, P., CHAPMAN, S., HEIN, I., TOTH, I. K., PRITCHARD, L. & BIRCH, P. R. J. 2007. A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature,* 450**,** 115–118.

WICKHAM, H. 2009. *ggplot2: Elegant graphics for data analysis,* New York, NY, Springer.

WICKHAM, H., AVERICK, M., BRYAN, J., CHANG, W., D'AGOSTINO MCGOWAN, L., FRANÇOIS, R., GROLEMUND, G., HAYES, A., HENRY, L., HESTER, J., KUHN, M., PEDERSEN, T. L., MILLER, E., BACHE, S. M., MÜLLER, K., OOMS, J., ROBINSON, D., SEIDEL, D. P., SPINU, V., TAKAHASHI, K., VAUGHAN, D., WILKE, C., WOO, K. & YUTANI, H. 2019. Welcome to the Tidyverse. *Journal of Open Source Software,* 4

WILDERMUTH, M. C., DEWDNEY, J., WU, G. & AUSUBEL, F. M. 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature,* 414**,** 562-565.

WIN, J. & KAMOUN, S. 2004. pCB301-p19: A Binary Plasmid Vector to Enhance Transient Expression of Transgenes by Agroinfiltration. *http://www.KamounLab.net*.

WONGSURAWAT, T., JENJAROENPUN, P., TAYLOR, M. K., LEE, J., TOLARDO, A. L., PARVATHAREDDY, J., KANDEL, S., WADLEY, T. D., KAEWNAPAN, B., ATHIPANYASILP, N., SKIDMORE, A., CHUNG, D., CHAIMAYO, C., WHITT, M., KANTAKAMALAKUL, W., SUTTHENT, R., HORTHONGKHAM, N., USSERY, D. W., JONSSON, C. B. & NOOKAEW, I. 2019. Rapid Sequencing of Multiple RNA Viruses in Their Native Form. *Frontiers in Microbiology,* 10.

WONGSURAWAT, T., JENJAROENPUN, P., WANCHAI, V. & NOOKAEW, I. 2022. Native RNA or cDNA Sequencing for Transcriptomic Analysis: A Case Study on *Saccharomyces cerevisiae*. *Frontiers in Bioengineering and Biotechnology,* 10.

WORKMAN, R. E., TANG, A. D., TANG, P. S., JAIN, M., TYSON, J. R., RAZAGHI, R., ZUZARTE, P. C., GILPATRICK, T., PAYNE, A., QUICK, J., SADOWSKI, N., HOLMES, N., DE JESUS, J. G., JONES, K. L., SOULETTE, C. M., SNUTCH, T. P., LOMAN, N., PATEN, B., LOOSE, M., SIMPSON, J. T., OLSEN, H. E., BROOKS, A. N., AKESON, M. & TIMP, W. 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nature Methods,* 16**,** 1297–1305.

WU, Y. K., ZHANG, B., LIU, S. B., ZHAO, Z. W., REN, W. J., CHEN, L., YANG, L. M., ZHUANG, M., LV, H. H., WANG, Y., JI, J. L., HAN, F. Q. & ZHANG, Y. Y. 2023. A Whole-Genome Assembly for *Hyaloperonospora parasitica*, A Pathogen Causing Downy Mildew in Cabbage (*Brassica oleracea* var. *capitata* L.). *Journal of Fungi,* 9.

XIANG, T. T., ZONG, N., ZOU, Y., WU, Y., ZHANG, J., XING, W. M., LI, Y., TANG, X. Y., ZHU, L. H., CHAI, J. J. & ZHOU, J. M. 2008. *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Current Biology,* 18**,** 74-80.

XING, Y. Y., XU, N., BHANDARI, D. D., LAPIN, D., SUN, X. H., LUO, X. M., WANG, Y. Q., CAO, J. D., WANG, H. B., COAKER, G., PARKER, J. E. & LIU, J. 2021. Bacterial effector targeting of a plant iron sensor facilitates iron acquisition and pathogen colonization. *Plant Cell,* 33**,** 2015-2031.

XIONG, H. B., PAN, H. M., LONG, Q. Y., WANG, Z. Y., QU, W. T., MEI, T., ZHANG, N., XU, X. F., YANG, Z. N. & YU, Q. B. 2022. AtNusG, a chloroplast nucleoid protein of bacterial origin linking chloroplast transcriptional and translational machineries, is required for proper chloroplast gene expression in *Arabidopsis thaliana*. *Nucleic Acids Research,* 50**,** 6715-6734.

XIONG, Q., YE, W. W., CHOI, D., WONG, J., QIAO, Y. L., TAO, K., WANG, Y. C. & MA, W. B. 2014. *Phytophthora* Suppressor of RNA Silencing 2 Is a Conserved RxLR Effector that Promotes Infection in Soybean and *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions,* 27**,** 1379-1389.

XU, Y. Z., ARRIETA-MONTIEL, M. P., VIRDI, K. S., DE PAULA, W. B. M., WIDHALM, J. R., BASSET, G. J., DAVILA, J. I., ELTHON, T. E., ELOWSKY, C. G., SATO, S. J., CLEMENTE, T. E. & MACKENZIE, S. A. 2011. MutS HOMOLOG1 Is a Nucleoid Protein That Alters Mitochondrial and Plastid Properties and Plant Response to High Light. *Plant Cell,* 23**,** 3428-3441.

YAGI, Y., ISHIZAKI, Y., NAKAHIRA, Y., TOZAWA, Y. & SHIINA, T. 2012. Eukaryotic-type plastid nucleoid protein pTAC3 is essential for transcription by the bacterial-type plastid RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 7541-7546.

YAMADA, K., YAMAGUCHI, K., SHIRAKAWA, T., NAKAGAMI, H., MINE, A., ISHIKAWA, K., FUJIWARA, M., NARUSAKA, M., NARUSAKA, Y., ICHIMURA, K., KOBAYASHI, Y., MATSUI, H., NOMURA, Y., NOMOTO, M., TADA, Y., FUKAO, Y., FUKAMIZO, T., TSUDA, K., SHIRASU, K., SHIBUYA, N. & KAWASAKI, T. 2016. The Arabidopsis CERK1-associated kinase PBL27 connects chitin perception to MAPK activation. *Embo Journal,* 35**,** 2468-2483.

YAMASHITA, Y., TAKAMATSU, S., GLASBRENNER, M., BECKER, T., NAITO, S. & BECKMANN, R. 2017. Sucrose sensing through nascent peptide-meditated ribosome stalling at the stop codon of *Arabidopsis bZIP11* uORF2. *Febs Letters,* 591**,** 1266-1277.

YAN, X., TANG, B. Z., RYDER, L. S., MACLEAN, D., WERE, V. M., ESEOLA, A. B., CRUZ-MIRELES, N., MA, W. B., FOSTER, A. J., OSÉS-RUIZ, M. & TALBOT, N. J. 2023. The transcriptional landscape of plant infection by the rice blast fungus *Magnaporthe oryzae* reveals distinct families of temporally co-regulated and structurally conserved effectors. *Plant Cell,* 35**,** 1360-1385.

YAN, Y. Q., YUAN, Q. F., TANG, J. T., HUANG, J. B., HSIANG, T., WEI, Y. D. & ZHENG, L. 2018. *Colletotrichum higginsianum* as a Model for Understanding Host-Pathogen Interactions: A Review. *International Journal of Molecular Sciences,* 19.

YANG, J. Y., FANG, Y. L., WANG, P., YE, J. R. & HUANG, L. 2018. Pleiotropic Roles of ChSat4 in Asexual Development, Cell Wall Integrity Maintenance, and Pathogenicity in *Colletotrichum higginsianum*. *Frontiers in Microbiology,* 9.

YAO, D. Y., GONZALES-VIGIL, E. & MANSFIELD, S. D. 2020. Arabidopsis sucrose synthase localization indicates a primary role in sucrose translocation in phloem. *Journal of Experimental Botany,* 71**,** 1858-1869.

YE, F. H., ALBAROUKI, E., LINGAM, B., DEISING, H. B. & VON WIRÉN, N. 2014. An adequate Fe nutritional status of maize suppresses infection and biotrophic growth of *Colletotrichum graminicola*. *Physiologia Plantarum,* 151**,** 280-292.

YE, L., ZHAO, X., BAO, E. C., CAO, K. & ZOU, Z. R. 2019. Effects of Arbuscular Mycorrhizal Fungi on Watermelon Growth, Elemental Uptake, Antioxidant, and Photosystem II Activities and Stress-Response Gene Expressions Under Salinity-Alkalinity Stresses. *Frontiers in Plant Science,* 10.

YONEHARA, K., KUMAKURA, N., MOTOYAMA, T., ISHIHAMA, N., DALLERY, J. F., O'CONNELL, R. & SHIRASU, K. 2023. Efficient multiple gene knockout in *Colletotrichum higginsianum* via CRISPR/Cas9 ribonucleoprotein and URA3-based marker recycling. *Molecular Plant Pathology*.

YOU, M. P., AKHATAR, J., MITTAL, M., BARBETTI, M. J., MAINA, S. & BANGA, S. S. 2021. Comparative analysis of draft genome assemblies developed from whole genome sequences of two *Hyaloperonospora brassicae* isolate samples differing in field virulence on *Brassica napus*. *Biotechnol Rep (Amst),* 31**,** e00653.

YOUNG, M. E., KEEGSTRA, K. & FROEHLICH, J. E. 1999. GTP promotes the formation of early-import intermediates but is not required during the translocation step of protein import into chloroplasts. *Plant Physiology,* 121**,** 237-243.

YTTERBERG, A. J., PELTIER, J. B. & VAN WIJK, K. J. 2006. Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiology,* 140**,** 984-997.

YUAN, Q. F., YAN, Y. Q., SOHAIL, M. A., LIU, H., HUANG, J. B., HSIANG, T. & ZHENG, L. 2021. A Novel Hexose Transporter ChHxt6 Is Required for Hexose Uptake and Virulence in *Colletotrichum higginsianum*. *International Journal of Molecular Sciences,* 22.

ZABALA, M. D. T., LITTLEJOHN, G., JAYARAMAN, S., STUDHOLME, D., BAILEY, T., LAWSON, T., TILLICH, M., LICHT, D., BOLTER, B., DELFINO, L., TRUMAN, W., MANSFIELD, J., SMIRNOFF, N. & GRANT, M. 2015. Chloroplasts play a central role in plant defence and are targeted by pathogen effectors. *Nature Plants,* 1.

ZAMPOUNIS, A., PIGNE, S., DALLERY, J. F., WITTENBERG, A. H. J., ZHOU, S. G., SCHWARTZ, D. C., THON, M. R. & O'CONNELL, R. J. 2016. Genome Sequence and Annotation of *Colletotrichum higginsianum*, a Causal Agent of Crucifer Anthracnose Disease. *Microbiology Resource Announcements,* 4.

ZEEMAN, S. C., KOSSMANN, J. & SMITH, A. M. 2010. Starch: Its Metabolism, Evolution, and Biotechnological Modification in Plants. *In:* MERCHANT, S., BRIGGS, W. R. & ORT, D. (eds.) *Annual Review of Plant Biology, Vol 61.* Palo Alto: Annual Reviews.

ZEEMAN, S. C., NORTHROP, F., SMITH, A. M. & AP REES, T. 1998. A starch-accumulating mutant of *Arabidopsis thaliana* deficient in a chloroplastic starch-hydrolysing enzyme. *Plant Journal,* 15**,** 357-365.

ZEEMAN, S. C., TIESSEN, A., PILLING, E., KATO, K. L., DONALD, A. M. & SMITH, A. M. 2002. Starch synthesis in Arabidopsis. Granule synthesis, composition, and structure. *Plant Physiology,* 129**,** 516-529.

ZHAI, Y. S., YUAN, Q., QIU, S. Y., LI, S. S., LI, M. M., ZHENG, H. Y., WU, G. W., LU, Y. W., PENG, J. J., RAO, S. F., CHEN, J. P. & YAN, F. 2021. Turnip mosaic virus impairs perinuclear chloroplast clustering to facilitate viral infection. *Plant Cell and Environment,* 44**,** 3681-3699.

ZHANG, M. X., LI, Q., LIU, T. L., LIU, L., SHEN, D. Y., ZHU, Y., LIU, P. H., ZHOU, J. M. & DOU, D. L. 2015. Two Cytoplasmic Effectors of *Phytophthora sojae* Regulate Plant Cell Death via Interactions with Plant Catalases. *Plant Physiology,* 167**,** 164-175.

ZHANG, T., ZHAO, Y. L., ZHAO, J. H., WANG, S., JIN, Y., CHEN, Z. Q., FANG, Y. Y., HUA, C. L., DING, S. W. & GUO, H. S. 2016. Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nature Plants,* 2.

ZHANG, X. L., SZYDLOWSKI, N., DELVALLE, D., D'HULST, C., JAMES, M. G. & MYERS, A. M. 2008. Overlapping functions of the starch synthases SSII and SSIII in amylopectin biosynthesis in Arabidopsis. *Bmc Plant Biology,* 8.

ZHANG, Y., CHEN, M., SIEMIATKOWSKA, B., TOLECO, M. R., JING, Y., STROTMANN, V., ZHANG, J., STAHL, Y. & FERNIE, A. R. 2020. A Highly Efficient *Agrobacterium*-Mediated Method for Transient Gene Expression and Functional Studies in Multiple Plant Species. *Plant Communications,* 1.

ZHANG, Y., GAO, Y. H., WANG, H. L., KAN, C. C., LI, Z., YANG, X. F., YIN, W. L., XIA, X. L., NAM, H. G., LI, Z. H. & GUO, H. W. 2021a. *Verticillium dahliae* secretory effector PevD1 induces leaf senescence by promoting ORE1-mediated ethylene biosynthesis. *Molecular Plant,* 14**,** 1901-1917.

ZHANG, Y., ZHANG, X., CUI, H., MA, X., HU, G., WEI, J., HE, Y. & HU, Y. 2021b. Residue 49 of AtMinD1 Plays a Key Role in the Guidance of Chloroplast Division by Regulating the ARC6-AtMinD1 Interaction. *Frontiers in Plant Science,* 12.

ZHAO, W. Y., WANG, L., LIU, M. Z., ZHANG, D., ANDIKA, I. B., ZHU, Y. & SUN, L. Y. 2022. A Reduced Starch Level in Plants at Early Stages of Infection by Viruses Can Be Considered a Broad-Range Indicator of Virus Presence. *Viruses-Basel,* 14.

ZHOU, Z. Y., WU, Y. J., YANG, Y. Q., DU, M. M., ZHANG, X. J., GUO, Y., LI, C. Y. & ZHOU, J. M. 2015. An Arabidopsis Plasma Membrane Proton ATPase Modulates JA Signaling and Is Exploited by the *Pseudomonas syringae* Effector Protein AvrB for Stomatal Invasion. *Plant Cell,* 27**,** 2032-2041.

ZHU, D., XIONG, H. B., WU, J. H., ZHENG, C. H., LU, D. D., ZHANG, L. X. & XU, X. M. 2022. Protein Targeting Into the Thylakoid Membrane Through Different Pathways. *Frontiers in Physiology,* 12.

ZHU, Y. M., DUAN, L. T., ZHU, C. Q., WANG, L., HE, Z. R., YANG, M. & ZHOU, E. R. 2023. Dual Transcriptome Analysis Reveals That *ChATG8* Is Required for Fungal Development, Melanization and Pathogenicity during the Interaction between *Colletotrichum higginsianum* and *Arabidopsis thaliana*. *International Journal of Molecular Sciences,* 24.

ZUO, J. R., NIU, Q. W. & CHUA, N. H. 2000. An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant Journal,* 24**,** 265-273.

viii) Appendices

Appendix 1

Appendix 1 (**A1**) details the plasmids, primers and gBlocks used in this work.

A1 Tab. A: List of Level 0 Golden Gate parts used in this work. Pos. = position: P = promoter, NT = N-terminal tag, CDS = coding sequence, CT = C-terminal tag, T = terminator, Fv = fungal vector positions 1-5.

Part name	Details	Pos.	Source
L0M 35S P	Promoter (0.4 kb), 35S Short (Cauliflower Mosaic Virus) + 5'UTR, Ω (Tobacco Mosaic Virus) Engler et al., 2014	P	TSL SynBio
L0M ChEC153	pAGM1287, ChEC153 CDS without SP: Δ1-20	CDS	J. Jennings
L0M HaRxL94b	pAGM1287, HaRxL94b CDS without SP: Δ1-23	CDS	X. Liu
L0M HaRxL94	pAGM1287, HaRxL94 CDS without SP: Δ1-63	CDS	X. Liu
L0M HaRxL94b$_{\Delta 1\text{-}63}$	pAGM1287, HaRxL94b CDS truncation, without SP: Δ1-63 (Cloned from L0M HaRxL94b with 033/034)	CDS	This work
L0M HaRxL94$_{\Delta 1\text{-}23}$	pAGM1287, HaRxL94 CDS extension, without SP: Δ1-23 *	CDS	This work
L0M AtCPNB3	pAGM1287, AtCPNB3 (AT5G56500) CDS	CDS	This work
L0M AtEIF2-A2	pAGM1287, AtEIF2-A2 (AT5G05470), CDS	CDS	This work
L0M AtCRBIC	pAGM1287, CRBIC (AT1G53120) CDS	CDS	This work
L0M cTP	pAGM1287, cTP: chloroplast transit peptide of RuBisCO small subunit	CDS	D. Seung
L0M eGFP	pICSL01003 (pAGM1301), eGFP C-tag	CT	TSL SynBio
L0M mCherry	pICSL50004, mCherry C-tag, Engler et al., 2014	CT	TSL SynBio
L0M 35S T	pICH41414, polyadenylation signal/terminator, 35S (Cauliflower Mosaic Virus), Engler et al., 2014	T	TSL SynBio

L0M *ChEC153* 5′ HR	pAGM1251, IMI 349061 Chr8: 4,301,441 – 4,300,561	Fv1	This work
L0M cos1gpdA P	pAGM1276, cos1gpdA promoter from fv1 (amplified with primers 023 and 024)	Fv2	This work
L0M gpdA P	pAGM1276, gpdA promoter from fv *HygR* (amplified with primers 017 and 018)	Fv2	This work
L0M *HygR* CDS	pICH41308, *HygR* (*HPH*) CDS from fv *HygR*	Fv3	This work
L0M *ChEC153* 3′ HR	pICH53399, IMI 349061 Chr8: 4,298,712 – 4,297,750	Fv4	This work
L0M trpC T	pICH53388, From TSL Synbio pICSL60013	Fv5	This work

* Cloned in two parts: residues 24-63 (5′ part) amplified from HaRxL94b using primers 055/056, and residues 64-496 (3′ part) amplified from HaRxL94 using 057/059. Parts purified from agarose gel following electrophoresis and assembled into pAGM1287 acceptor plasmid via BpiI-driven Golden Gate reaction to assemble the parts in a scar-free assembly.

A1 Tab. B: **List of Golden Gate acceptor plasmids used in this work**. Numbers in brackets indicate the cloning positions referred to in **Figure 4.2** for Level 1 fungal vector construction from Level 0 parts cloned into these acceptor plasmids. Aside from fungal vectors, all acceptor plasmids are described by Engler et al., 2014.

Acceptor name	Level	Position	5′ overhang	3′ overhang	Source
pICH47732	1	1-F	GGAG	CGCT	TSL SynBio
fv *HygR*	1	1-F	GGAG	CGCT	J. Jennings
fv *ΔHygR*	1	1-F	GGAG	CGCT	This work
pAGM1287 (pICSL01005)	0	CDS1ns	AATG	TTCG	TSL SynBio
pAGM1251 (pICSL01008)	0	P5Uf (1)	GGAG	CCAT	TSL SynBio
pAGM1276 (pICSL01002)	0	NTAG (2)	CCAT	AATG	TSL SynBio
pICH41308	0	CDS (3)	AATG	GCTT	TSL SynBio
pICH53388	0	3UTR (4)	GCTT	GGTA	TSL SynBio
pICH53399	0	TERM (5)	GGTA	CGCT	TSL SynBio

A1 Tab. C: List of Level 1 Golden Gate plasmids used in this work. P = promoter, CDS = coding sequence, Tag = C-terminal tag, T = terminator.

L1 Plasmid name	Backbone	P	CDS	Tag	T	Source
35S::ChEC153-eGFP	pICH47732	35S	L0M ChEC153	eGFP	35S	J. Jennings
35S::HaRxL94b-eGFP	pICH47732	35S	L0M HaRxL94b	eGFP	35S	X. Liu
35S::ChEC153-mCherry	pICH47732	35S	L0M ChEC153	mCherry	35S	This work
35S::HaRxL94b-mCherry	pICH47732	35S	L0M HaRxL94b	mCherry	35S	This work
35S::HaRxL94-eGFP	pICH47732	35S	L0M HaRxL94	eGFP	35S	X. Liu
35S::HaRxL94b$_{\Delta 1\text{-}63}$-eGFP	pICH47732	35S	L0M HaRxL94b$_{\Delta 1\text{-}63}$	eGFP	35S	This work
35S::HaRxL94$_{\Delta 1\text{-}23}$-eGFP	pICH47732	35S	L0M HaRxL94$_{\Delta 1\text{-}23}$	eGFP	35S	This work
35S::CPNB3-mCherry	pICH47732	35S	L0M AtCPNB3	mCherry	35S	This work
35S::EIF2-A2-mCherry	pICH47732	35S	L0M AtEIF2-A2	mCherry	35S	This work
35S::CRBIC-mCherry	pICH47732	35S	L0M AtCRBIC	mCherry	35S	This work
35S::cTP-eGFP	pICH47732	35S	L0M cTP	eGFP	35S	This work
35S::Mit-GFP (pICSL11247)	pICH47742	35S	CoxIV Mit	GFP	Nos	TSL SynBio
35S::ChEC153.2a-eGFP	pICH47732	35S	L0M ChEC153.2a	eGFP	35S	This work
35S::ChEC153.2a-mCherry	pICH47732	35S	L0M ChEC153.2a	mCherry	35S	This work

A1 Tab. D: **List of Gateway plasmids used in this work.** These constructs were provided by the sources listed: A. Breakspear (Faulkner lab, JIC), D. Seung (JIC), A. Korolev (Saunders lab, JIC) and N. Kamble (Seung lab, JIC). Gateway cloning entry backbones: Karimi et al., 2002.

Plasmid name	Backbone	CDS	Tag	Source
35S::MRC-RFP	pB7RWG2	AtMRC	RFP	A. Breakspear
35S::PEND$_{1-88}$-RFP	pB7RWG2	AtPEND$_{1-88}$	RFP	D. Seung
35S::TaCSP41a-eGFP	p7FWG2	TaCSP41a	eGFP	A. Korolev
35S::GBSS-RFP	pB7RWG2	NbGBSSa E490Q	RFP	N. Kamble

A1 Tab. E: **List of primers used in this work.** Where Golden Gate cloning extensions are present, the Type IIS recognition sequences are coloured green, with Golden Gate overhangs coloured orange. F or R indicates forward or reverse primer, with primer sequences presented from 5′ to 3′. Genomic locations listed as binding sites for *Ch* primers refer to the ASM167251v1 genome assembly (Zampounis et al., 2016), Chr = chromosome.

Primer	F/R	Purpose	Sequence
Sequencing primers			
0015	F	Level 0 sequencing	CGTTATCCCCTGATTCTGTGGATAAC
0016	R	Level 0 sequencing	GTCTCATGAGCGGATACATATTTGAATG
0229	F	Level 1 sequencing	GAACCCTGTGGTTGGCATGCACATAC
0230	R	Level 1 sequencing	CTGGTGGCAGGATATATTGTGGTG
Cloning primers			
015	F	fv *ΔHygR* backbone cloning	ctGAAGACgaCGCTCACAGAGTGGGGTCAGATTGTC
016	R	fv *ΔHygR* backbone cloning	atGAAGACgaTCAAATGAGTTTTGATTTAATTTC
017	F	gpdA promoter cloning	ctGAAGACgaCATTATGTCTGCTCAAGCGGGGTAG
018	R	gpdA promoter cloning	atGAAGACgaCCATGCTCAAGCTGCTCTAGCATTC
019	F	*HygR* CDS cloning	ctGAAGACgaAAGCCACTATTCCTTTGCCCTCG
021	F	trpC terminator cloning	ctGAAGACgaGCTTCGGGGGATGAAATCATC

022	R	trpC terminator cloning	atGAAGACgaTACCAGTGGAGATGTGGAGTGG
023	F	cos1gpdA promoter cloning	ctGAAGACgaCCATCGATACCTGCGTCATAATTG
024	R	cos1gpdA promoter cloning	atGAAGACttCATTGGTGATGTCTGCTCAAGC
025	R	*HygR* CDS cloning	atGAAGACgaAATGCCTGAACTCACCGCGAC
029	F	*ChEC153* 5′ HR cloning from Chr8: 4,301,441	ctGAAGACgaGGAGCGAATCCGAGCATACGTGCAACC
030	R	*ChEC153* 5′ HR cloning from Chr8: 4,300,561	atGAAGACgaATGGGTCAACGCTTGTTGAGGACTGATG
031	F	*ChEC153* 3′ HR cloning from Chr8: 4,298,712	ctGAAGACgaGGTATCTGCCTTCTGTTTCCGCTGC
032	R	*ChEC153* 3′ HR cloning from Chr8:4,297,750	atGAAGACgaAGCGAGAGTACTTCGCGATGGTGGAC
033	F	HaRxL94b$_{\Delta 1\text{-}63}$ cloning	ctGAAGACgaAATGATGCATAGCTCGAACTCTACCAG
034	R	HaRxL94b$_{\Delta 1\text{-}63}$ cloning	atGAAGACgaCGAACCCGGGACTTTCTCG
055	F	HaRxL94$_{\Delta 1\text{-}23}$ cloning, 5′ part	ctGAAGACgaAATGATGGCGAGTGAGGCTTC
056	R	HaRxL94$_{\Delta 1\text{-}23}$ cloning, 5′ part	ctGAAGACgaTATGTGCGAATGTCGGTGTTGAAG
057	F	HaRxL94$_{\Delta 1\text{-}23}$ cloning, 3′ part	ctGAAGACgaCATAGCTCGAACTCTACCAGAATAC
059	R	HaRxL94$_{\Delta 1\text{-}23}$ cloning, 3′ part	ctGAAGACgaCGAACCTTGCGGGACTTTC
097	F	Amplify ChEC153 from Chr8: 8: 4,300,557	ctGAAGACgaAATGATGTCTTGGATTATCACACTCAGTG
099	R	Amplify ChEC153 from Chr8: 4,298,750	tGAAGACgaCGAActGACTTGGAACCCATGGAATGGGCG
100	F	Amplify ChEC153 from Chr8: 4,299,682	ctGAAGACgaAATGGTTTCCAGATTGCCCTAGGCTA
101	R	Amplify ChEC153 from Chr8: 4,299,282	tGAAGACgaCGAActTTTGTAGATTCACCGTCGATCTCAC

102	F	Amplify CH63R_13188 from Chr9: 3,252,456	ctGAAGACgaAATGAAGTCCGCCATTCTTGCCATC
103	R	Amplify CH63R_13188 from Chr9: 3,252,725	tGAAGACgaCGAActGCCGAGGACTTGCTTGGG
Genotyping primers			
SALK_LB	F	SALK T-DNA specific LBb1.3	ATTTTGCCGATTTCGGAAC
069	F	SALK_099429.36.25.x LP	AACTATCAGGACTTGGGCTCC
070	R	SALK_099429.36.25.x RP	TAGAGGCTTCTCCAAGCACAG
071	F	SALK_041981.40.50.x LP	ATGGGTTTGGACTGTTTCCTC
072	R	SALK_041981.40.50.x RP	TGCCCTAACCTCTAAATTCCC
5′ RACE primers			
C1	R	5′ RACE kit control neo1	CAGGCATCGCCATGGGTCAC
C2	R	5′ RACE kit control neo2	GCTGCCTCGTCCTGCAGTTC
C3	F	5′ RACE kit control neo3	GATTGCACGCAGGTTCTCCG
dT-a	F	5′ RACE kit oligo(dT)-Anchor	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV * * V = A, C, or G
Anchor	F	5′ RACE Anchor primer + cloning extension	ctGAAGACgaAATGGACCACGCGTATCGATGTCGAC
GSP1	R	5′ RACE cDNA synthesis	AGTCTGAACCTTGGGTACCT
GSP2	R	5′ RACE PCR1	tGAAGACgaCGAActTGCCTGCTATTGCGTGCACC
GSP3	R	5′ RACE PCR2	tGAAGACgaCGAActTTGCCTCTTCGGAAGGTTTCTGG
qPCR primers			
UBQ_Fwd	F	*AtUBQ10* qPCR	AGTCTACTCTTCACTTGGTCCTGC
UBQ_Rev	R	*AtUBQ10* qPCR	GCCCCAAAACACAAACCACCAAAG

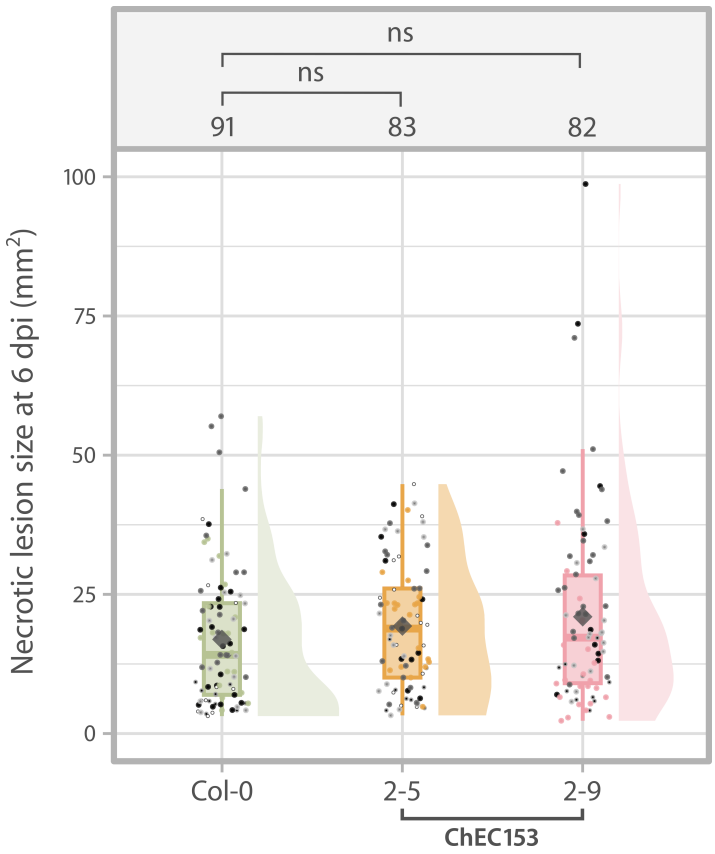
A1 Tab. F: **List of gBlocks synthesised during this work.** Synthesis by IDT (*AtCPNB3*, *AtEIF2-A2* and *AtCRBIC*) and GENEWIZ (*ChEC153.2a*). Sequences were synthesised with Golden Gate cloning extensions to facilitate cloning to Level 0 vectors. Sequences were domesticated by the removal of internal BpiI and BsaI cut sites by introduction of synonymous SNPs where these were deemed to interfere with cloning. Type IIS recognition sequences are coloured green, with Golden Gate overhangs coloured orange.

Name	Sequence
AtCPNB3	tGAAGACgaAATGATGGCATCAACCTTTAGCGCAACGTCGTCAATGGGTTCATCTTTGGCTC CTCCTTCGAATCGATTGTCATCTTTTGTTTCGATCTCATCAAGCTCTTTTGGAAGGACTCAGA GTATTGCTCAGAGGAAAGCAAGGTTTCCTAAAATATATGCGGCTAAGCAATTGCATTTCAAT AAGGACGGGACTGCGATTAAGAAGCTTCAAGCTGGTGTGAATAAACTTGCGGATTTAGTT GGGGTTACTTTAGGTCCTAAAGGCAGGAATGTTGTTCTAGAGAGCAAATACGGTTCCCCTA GAATCGTTAATGATGGTGTTACGGTCGCTAGAGAGGTTGAGCTTGAGGATCCAGTTGAGA ACATTGGTGCTAAGCTAGTGAGACAAGCTGCTTCCAAGACTAATGACTTAGCTGGTGATGG AACAACGACTTCTGTTGTTCTTGCTCAAGGTCTTATTGCTGAAGGTGTAAAGGTGGTAGCTG CTGGTGCAAATCCTGTTTTGATTACCAGAGGTATTGAGAAAACCACCAAAGCTCTTGTGGCT GAGTTGAAGAAAATGTCAAAGGAGGTTGAGGACAGTGAACTAGCAGATGTTGCAGCAGTT AGTGCAGGAAACAATTACGAAGTAGGCAATATGATTGCAGAAGCAATGGCAAAAGTTGGT CGTAAAGGTGTTGTCACTCTAGAGGAAGGCAAAAGCGCTGAGAACAGCCTTTACGTTGTTG AAGGAATGCAATTTGATCGTGGTTATATCTCTCCTTACTTTGTCACCGACAGTGAGAAAATG TGTGCAGAATATGAGAACTGCAAGTTGTTTCTTGTTGACAAAAAGATAACAAATGCTAGAG ATATTATTAGCATTCTGGAAGATGCGATTAAAGGTGGATACCCGCTTTTGATCATTGCTGAA GATATTGAGCAAGAGCCATTAGCGACTCTTGTTGTTAACAAGCTTCGTGGGACAATAAAAG TTGCTGCTTTGAAAGCTCCTGGATTTGGAGAGAGGAAAAGCCAGTACCTTGATGACATTGC TGCGCTTACAGGAGCTACTGTGATTAGGGAAGAAGTTGGACTTCAGTTGGAGAAAGTTGG ACCAGAAGTTTTGGGTAATGCTGGTAAAGTGGTTCTCACAAAAGATACTACAACGATAGTT GGTGATGGAAGTACTGAAGAAGTTGTGAAGAAGAGAGTTGAACAGATCAAGAATCTTATC GAGGCTGCTGAACAAGACTATGAAAAGGAAAAGCTTAATGAGAGAATTGCTAAATTATCA GGCGGTGTTGCTGTAATTCAGGTTGGAGCACAAACTGAGACAGAGCTTAAGGAGAAGAAA TTAAGAGTTGAAGATGCTCTTAATGCTACAAAGGCTGCTGTGGAAGAAGGTATTGTTGTTG GCGGAGGATGTACTCTGCTAAGACTTGCATCGAAAGTGGACGCTATTAAAGAGACTCTAGC AAACGATGAAGAAAAGGTTGGAGCTGACATTGTTAAGAAAGCATTGAGCTACCCACTGAA GTTAATCGCCAAGAACGCTGGTGTCAATGGCAGCGTTGTCAGTGAGAAGGTTCTTTCTAGT GACAACCCTAAACATGGTTACAATGCTGCAACTGGCAAATACGAAGATCTCATGGCTGCAG GAATCATTGATCCAACCAAAGTTGTGAGATGTTGCCTAGAGCATGCTTCATCTGTGGCAAA GACATTCTTGATGTCTGATTGTGTTGTTGTGGAGATCAAGGAGCCTGAATCAGCAGCTCCT GCTGGTAACCCTATGGACAATTCAGGTTACGGCTTCagTTCGtcGTCTTCa
AtEIF2-A2	tGAAGACgaAATGATGGCGAATCCTGCTCCGAATCTAGAATGTCGTATGTACGAATCGAGAT ACCCTGATGTAGACATGGCGGTGATGATTCAGGTCAAGACCATCGCTGACATGGGAGCTTA CGTATCTCTCCTTGAATACAACAACATCGAAGGAATGATCCTGTTCTCCGAGCTCTCTCGCC GTCGGATTCGTAGTATCAGTAGCTTAATCAAGGTCGGTCGTACCGAGCCTGTTATGGTCCTT CGTGTCGATAGAGAGAGAGGTTACATTGATCTCAGTAAACGTAGGGTTAGTGATGAGGAC AAAGAGGCTTGTGAGGAGAGGTATAATAAGAGCAAGCTTGTTCACTCTATCATGCGTCATG TTGCTGAGACTGTTGGTGTCGATTTGGAGGAGCTATACGTAAACATCGGTTGGCCATTGTA TAAGAAGCATGGACATGCTTTTGAGGCTTTCAAAATTGTTGTCACTGATCCTGATTCAGTTT

	TCGATGCTCTTACCCGAGAAGTTAAAGAAACTGGACCTGATGGTGTGGAGGTGACCAAAG TTGTCCCGGCTGTGTCTGAAGAATTGAAAGATGCATTTTTGAAGGACATTAGGAGGAGAAT GACACCACAGCCAATGAAGATTCGTGCTGATATTGAATTGAAGTGTTTTCAGTTTGATGGA GTTCTCCACATCAAGGAAGCCATGAAGAAGGCAGAGGCTGTAGGTACTGATGATTGTCCA GTCAAAATCAAGCTCGTTGCTCCACCACTTTATGTACTCACAACTCACACCCATTACAAGGA AAAAGGAATAGTGACTCTGAATAAAGCAATTGAAGCATGCATTACTGCAATTGAGGAACAC AAGGGTAAACTTGTCGTTAAAGAAGGTGCTCGTGCGGTGAGTGAGCGTGATGACAAATTG CTTGCTGAGCACATGGCTAAGCTTAGAATGGATAATGAAGAAATGAGTGGTGATGAGGGA AGCGAAGATGAAGAAGAGGACACTGGAATGGGAGAAGTCGATATCGATGGAGGTAGCGG GATAATTGAAagTTCGtcGTCTTCa
AtCRBIC	tGAAGACgaAATGATGGCTGTCACAAGCTTGGCTCCTCCATGGGTCATCTTGAGACTAGCTT TCCGGTCAGTAGCAGCTTCTTCTTGTCTCCACACAAATCAAAAAACCCTAATCACGAATCTCT CCATTCCCACTTCGTTTCCTCTCCGACAGTCAGCTTTGAGAAGATGTTACAGTGCAGAAGCT ATAAAAGGAGATGTAGATTTTCTCCTCAAAGGAGTTGGAGACCAAGCTGTTGCCAAAGAAG TCAAGCAAATTCTTGAAATGGCAAGACGTGCATCATCAAAAAGAGAAGTTCTTCATACAGA TTTTCTCACACCACCTATTGTTAAGGAATCAGTTTCACTATTGGAAAAATTTGCTGATGTTAA GATAGTTGCTCAAGGAGGTTACCCTGAGGCTGAACGGTGTAGGATCTCGATTGGACATCCT GATGTCCTAACTAGTGATCCAGATATAGTTGCTGCTTTGAGTATCACAGGGAATTTCGGGTT TCAACCTTGTTCTCACGGTGACTTCCTTGGTGCTATTCTTGGCACGGGAATTTCCAGGGAAA AACTTGGAGATATCTTAATTCAGGAAGAAAAGGGAGCCCAAGTCCTGATAGTTCCTGAACT AGTTGACTTTGTTGTTACCGCTCTTGACAAGGTTGGAAATGTTGGTGTAACTTGTAGTAAGA TACCTTTGCTTGCTCTTGAATACGAACCGCCTAGGACTAATTCCTTTAAAACCGTGGAGGCC TCGTTGAGAATTGATGCAGTAGCTAGTGCTGGTTTCAAGATTTCGCGGTCAAAGCTAGTTG ATTTGATTAGTAGCAAGGATGTTCGGGTTAACTGGGCAACCGTTACAAAGAACGGAACCAT AGTCAAGACTGGTGATGTTGTCTCCGTTAGCGGGAAAGGGAGACTCAAGATTGGAGAGAT CAATGAAACGAAGAAAGGTAAATTTGCAGTTGAAATCATCAGATATCTTagTTCGtcGTCTTC a
ChEC153.2a	cactctgtGGTCTCaAATGATGTGGCTTGCTTTTGACGTGATTGCGGATCTCTCCTTTGGCGAG GAGCTTGGCACAATCGAGATAGGCGAAGGAAACTACTGGATGCACATGCTGGCTAACAGT GGTTTCCAGATTGCCCTAGGCTACGTGGTCCGTCGACGATGGAAGGCTTTCCAAGACCTCG TGAGGTATTGCCTCGTGAATGAGAAAAGCAAGAGGATGCGCAACAAGTACCTCGCCAACG CCCGCCAGGCGGCTTCTCAGCGACTTCAGAGGGGTGCAGACGTCGATCGTTTCGATTTCTT CAGTCATCTACTGCGTGAAAAGGCCCCAGAAGCCAACATTGAATTCTTCGCCTCGCAGGGT AGCACATTGGTTGCTGCTGGAACTGAAACGACATCCACCTTCATGTCCGCGCTCACGTACC ATCTCTTGCAGCAGCCTGACTGCCTGAAGCACCTTCAGGATGAGTTACACTGCACGTTCCG ACAGCACAGTGAGATCGACGGTGAATCTACAAAGCCACTGAAGTATCTCAATGCGGTGAT AGAAGAGGGGATGCGCATTTTCAACCCTGCAGCCTTTGGGTTGCCTAGGGTTAGTCCAGG AGCTAATGTTACTGGAGAATGGATTCCGAAAGGGACGGTTATAGCCACGGCAACTCATGT AACTTCTCGTGACGAGAGATGGTTCTGTAAAGCGAAAGAGTTCCATCCGGAAAGATGGCTT CCCGCGGATCACCCATGCTACGATGCTACATTCGAAAATGACCAAAAAGAGGCTTCAAAGC CGTTCTCCATCGGCCCAAGGTCTTGCATTGGTATCCACTTGTCTTACATAGAGGTACGAATA TGTATTGCAAAGTTGGCCTGGAGCTTTGACTGGGAACAGGTCAACAAAAGCGAAGATTTC GTCAGAGATGCGCGCTTGTTAGGTCTTTGGAAGGCTTCTCCGTTCCACGTGCGTTATCGTC CGTTCCATGGGTTCCAAGTCagTTCGtGAGACCacgaagtg

Appendix 2

Appendix 2 (**A2**) contains supplementary infection assay data and statistical analyses of these data, as well as statistical analysis of *C. higginsianum in vitro* colony sizes.



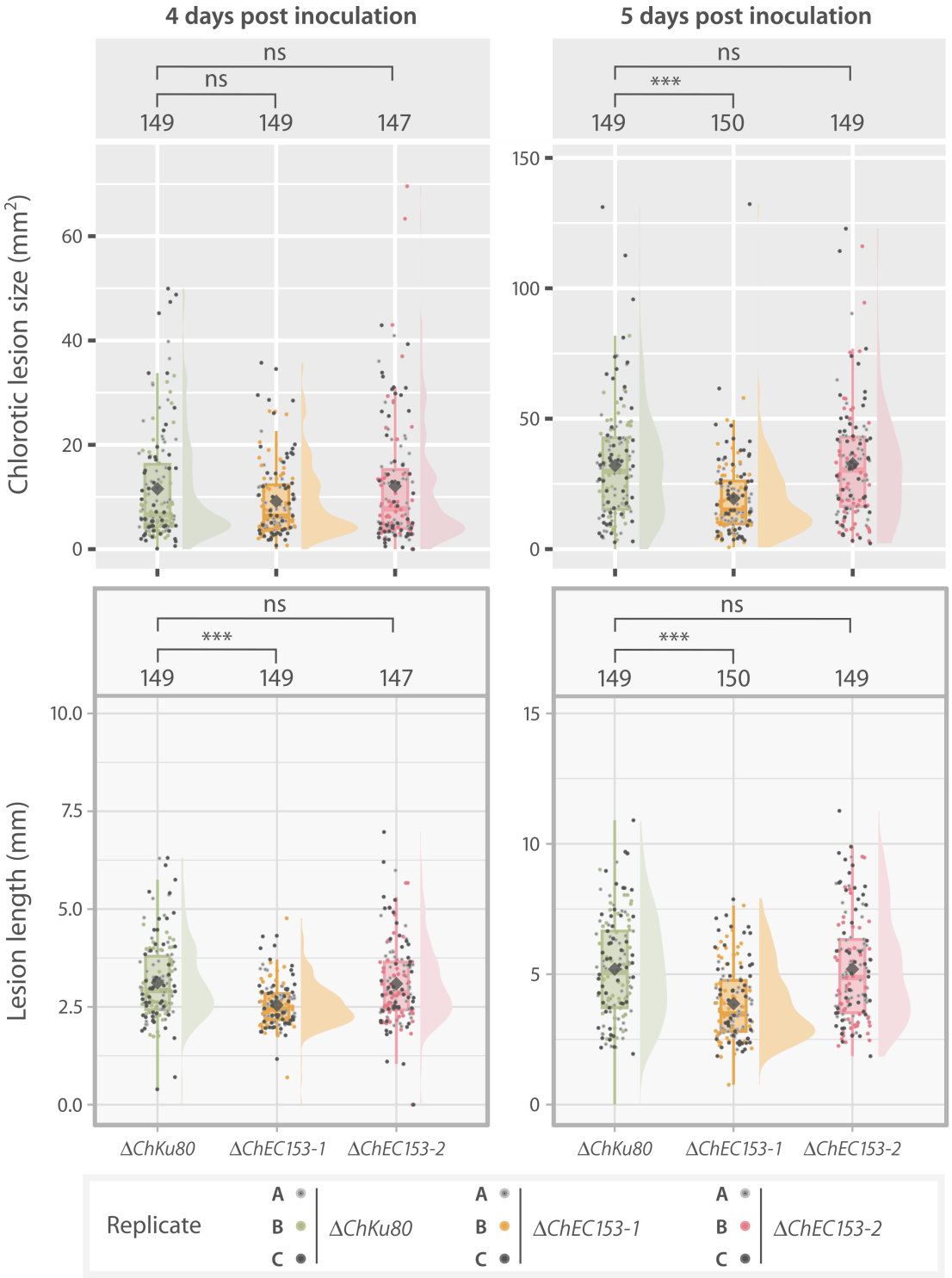
A2 Fig. A: ***35S::ChEC153-eGFP* lines show no difference in lesion size relative to Col-0 following infection with *C. higginsianum* at six days post inoculation.** Necrotic lesion areas measured at 6 dpi for *35S::ChEC153-eGFP* Arabidopsis lines 2-5 and 2-9. Data were pooled from six independent replicates of the experiment (represented by differently coloured points), with values of n stated above each plot. Mean lesion sizes are represented by grey diamonds. Differences between genotype lesion sizes were evaluated using an ANOVA on a linear mixed effects model, with the independent replicate of the experiment as included as a random factor, and Tukey HSD post hoc tests: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$ (not significant). Data for lesion necrotic lesion sizes at four- and five-days post inoculation are presented in **Figure 4.2**. Results of statistical tests are presented in **Appendix 2 Tab. A**.

A2 Tab. A: **Summary of statistical test results for *C. higginsianum* infection assay comparing Col-0 and *35S::ChEC153-eGFP* Arabidopsis lines.** For each time-point from four to six days post inoculation (dpi) lesion sizes were quantified as necrotic lesion areas and chlorotic lesion areas, and analysed using an ANOVA on a linear mixed effects model including independent replicates of the experiment as a random factor. ANOVA F values, degrees of freedom (df) and p-values reported. Tukey HSD adjusted p-values are shown for each time-point for comparisons between the either *35S::ChEC153-eGFP* line and the Col-0 control. P values < 0.05 are highlighted in yellow (statistically significant differences seen), and > 0.05 are highlighted in grey (not statistically significantly different). Data are presented in **Figure 4.2** and **Appendix 2 Fig. A**.

ChEC153-eGFP	dpi	ANOVA			Tukey HSD adjusted p-values	
		F	df	p	2-5 vs Col-0	2-9 vs Col-0
Necrotic area	4	7.9401	2	< 0.001	0.001	0.008
	5	5.3696	2	0.005	0.155	0.004
	6	2.3410	2	0.098	0.348	0.090
Chlorotic area	4	6.4088	2	0.002	0.007	0.011
	5	2.5628	2	0.079	0.325	0.071
	6	1.3929	2	0.250	0.542	0.235

A2 Tab. B: **Summary of statistical test results for *C. higginsianum* strain growth assay.** For each time-point (two to eleven days of growth), the colony size data were analysed using an ANOVA on a linear model, with ANOVA F values, degrees of freedom (df) and p-values reported. Tukey HSD adjusted p-values are shown for each time-point for comparisons between either *ΔChEC153* strain and the *ΔChKu80* background strain. P values < 0.05 are highlighted in yellow (statistically significant differences seen), and > 0.05 are highlighted in grey (not statistically significantly different). Corresponding data are presented in **Figure 4.6**.

Statistical analysis		Days of growth following sub-culturing				
		2	4	6	9	11
ANOVA	F	23.723	227.67	128.07	175.57	136.22
	df	2	2	2	2	2
	p	0.001	< 0.001	< 0.001	< 0.001	< 0.001
Tukey HSD p value	*ΔChEC153-1* vs *ΔChKu80*	0.011	< 0.001	< 0.001	< 0.001	< 0.001
	ΔChEC153-2 vs *ΔChKu80*	0.119	0.003	0.450	0.082	0.790

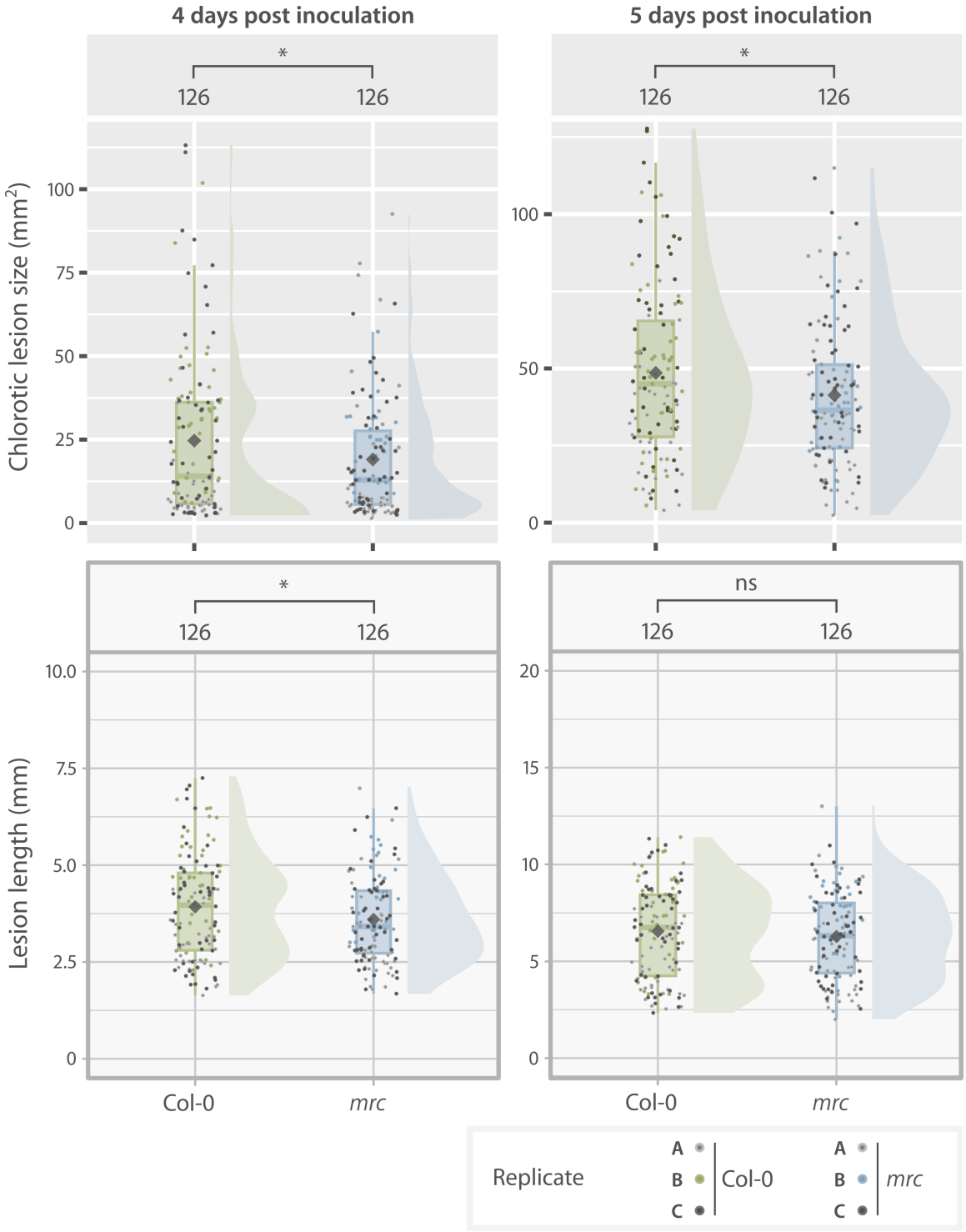


A2 Fig. B: ***ΔChEC153*** **knockout strains show no clear virulence phenotype *in planta* relative to *ΔChKu80* that cannot be accounted for by *in vitro* growth defects.** Chlorotic lesion areas and lesion lengths at 4 and 5 dpi. Data were pooled from three independent replicates of the experiment (represented by differently coloured points), with values of n stated above each plot. Mean lesion sizes are represented by grey diamonds. Differences between knockout and background strain lesions evaluated using an ANOVA on a linear mixed effects model, with the independent replicate of the experiment included as a random factor, and Tukey HSD post hoc tests: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$ (not significant). Corresponding data for necrotic lesion sizes are presented in **Figure 4.9**. Results of statistical tests are presented in **Appendix 2 Tab. C**.

A2 Tab. C: **Summary of statistical test results for *C. higginsianum* infection assay comparing *ΔChKu80* and *ΔChEC153* strains.** At four-, five-, and six-days post inoculation (dpi) lesion sizes were quantified as necrotic lesions, chlorotic lesions or lesion lengths, and analysed using an ANOVA on a linear mixed effects model including independent replicates of the experiment as a random factor. ANOVA F values, degrees of freedom (df) and p-values reported. Tukey HSD adjusted p-values are shown for each time-point for comparisons between the either *ΔChEC153* strain and the *ΔChKu80* control. Further, statistical analysis of the individual replicates of the experiment (A, B, and C) are shown, analysed using an ANOVA on a linear model. ANOVA F values, degrees of freedom (df), and p-values reported, alongside Tukey HSD adjusted p-values for each comparison of knockout strain relative to background strain. P values < 0.05 are highlighted in yellow (statistically significant differences seen), and > 0.05 are highlighted in grey (not statistically significantly different). Corresponding data are presented in **Figure 4.9** and **Appendix 2 Fig. B**.

Data pooled from three individual replicates for analysis							
Lesion measurement		**dpi**	**ANOVA**			**Tukey HSD adjusted p-values**	
			F	df	p	*ΔChEC153-1* vs *ΔChKu80*	*ΔChEC153-2 vs ΔChKu80*
Necrotic area		4	13.524	2	< 0.001	< 0.001	0.877
		5	22.642	2	< 0.001	< 0.001	0.997
		6	35.701	2	< 0.001	< 0.001	0.9131
Chlorotic area		4	3.5452	2	0.029	0.102	0.892
		5	21.558	2	< 0.001	< 0.001	0.946
		6	37.260	2	< 0.001	< 0.001	0.935
Lesion length		4	17.781	2	< 0.001	< 0.001	0.967
		5	26.114	2	< 0.001	< 0.001	1.000
		6	41.013	2	< 0.001	< 0.001	0.962
Individual replicates (A, B, and C) analysed independently							
Lesion measurement and replicate		**dpi**	**ANOVA**			**p-values**	
			F	df	p	*ΔChEC153-1* vs *ΔChKu80*	*ΔChEC153-2 vs ΔChKu80*
Necrotic area	A	4	8.1213	2	< 0.001	< 0.001	0.897
	B	4	3.5442	2	0.031	0.089	0.940
	C	4	3.6226	2	0.029	0.029	0.818
	A	5	15.652	2	< 0.001	< 0.001	0.796
	B	5	4.3622	2	0.014	0.020	0.938
	C	5	6.7608	2	0.002	0.003	0.971
	A	6	33.557	2	< 0.001	< 0.001	0.941
	B	6	7.4443	2	< 0.001	0.002	0.944
	C	6	7.7815	2	< 0.001	0.008	0.775

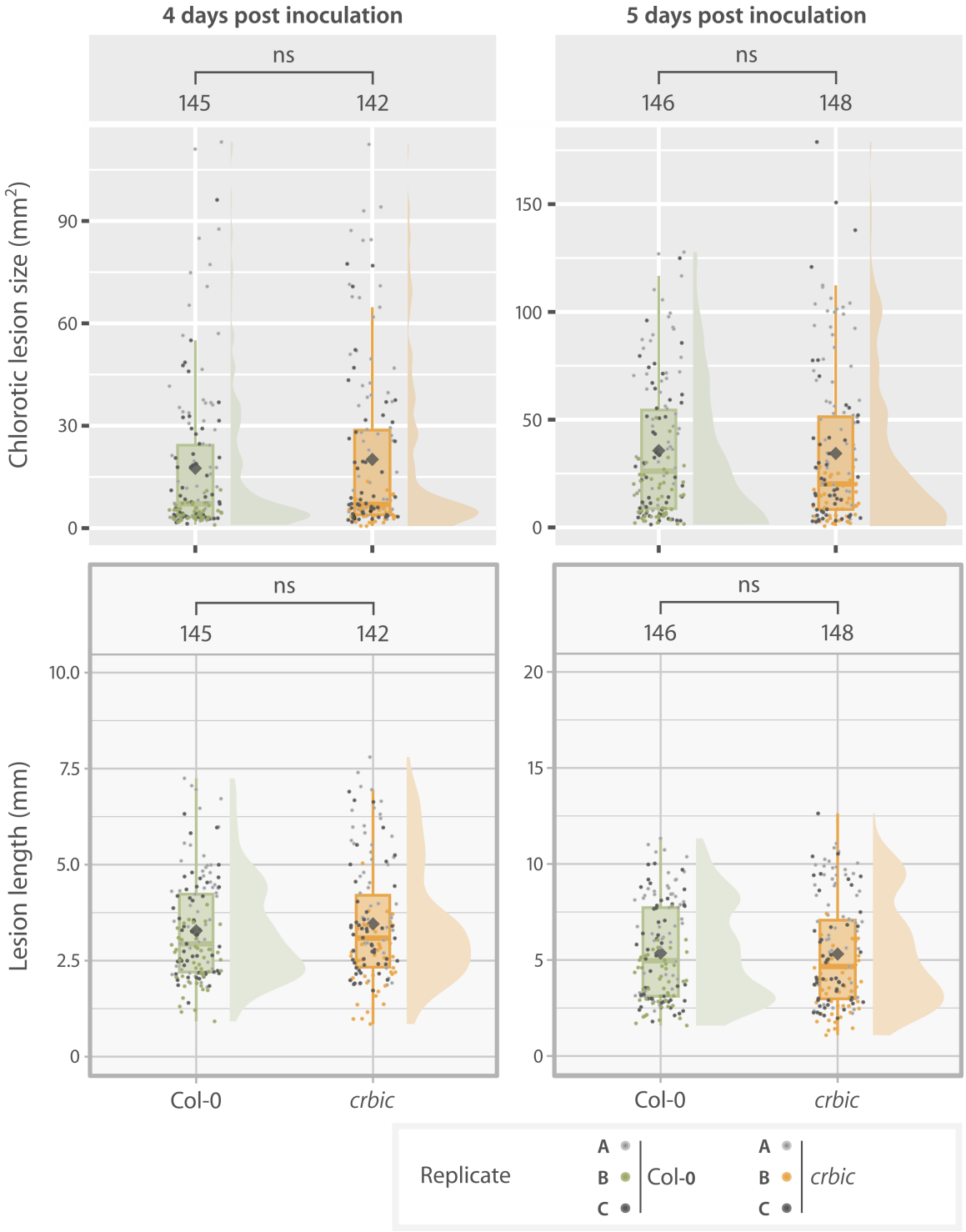
Chlorotic area	A	4	5.0774	2	0.007	0.009	0.850
	B	4	1.4613	2	0.235	0.908	0.440
	C	4	0.2720	2	0.762	0.790	1.000
	A	5	20.788	2	< 0.001	< 0.001	0.646
	B	5	6.0080	2	0.003	0.006	0.968
	C	5	4.1148	2	0.018	0.044	0.990
	A	6	37.190	2	< 0.001	< 0.001	0.608
	B	6	8.6630	2	< 0.001	< 0.001	0.587
	C	6	7.3788	2	< 0.001	0.011	0.750
Lesion length	A	4	10.600	2	< 0.001	< 0.001	0.952
	B	4	6.4773	2	0.002	0.004	0.963
	C	4	3.6525	2	0.028	0.055	0.999
	A	5	19.938	2	< 0.001	< 0.001	0.941
	B	5	5.9264	2	0.003	0.003	0.569
	C	5	6.4854	2	0.002	0.017	0.805
	A	6	34.493	2	< 0.001	< 0.001	0.915
	B	6	10.735	2	< 0.001	< 0.001	0.232
	C	6	8.5082	2	< 0.001	0.005	0.774



A2 Fig. C: ***mrc*** **Arabidopsis mutants are more resistant to *C. higginsianum* infection than Col-0 at 4 dpi.** Chlorotic lesion areas and lesion lengths at 4 and 5 dpi. Individual replicates of the experiment are plotted as differently coloured points. Values of n are shown above each plot, and mean lesion sizes for each are represented by grey diamonds. Asterisks denote significant differences between genotypes, determined using a t-test using Satterthwaite's method on a linear mixed effects model, with the independent replicates of the experiment included as a random factor: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns $p > 0.05$ (not significant). Corresponding data for necrotic lesion sizes are presented in **Figure 4.12**. Results of statistical tests are presented in **Appendix 2 Tab. D**.

A2 Tab. D: **Summary of statistical test results for *C. higginsianum* infection assay comparing *mrc* and Col-0 susceptibility.** At four- and five-days post inoculation (dpi) lesion sizes were quantified as necrotic lesions, chlorotic lesions or lesion lengths, and analysed using a t-test using Satterthwaite's method on a linear mixed effects model, with the independent replicates of the experiment included as a random factor (full dataset). Additionally, comparisons between genotypes for each individual replicate (A, B, and C) are shown: p-values are shown for each time-point for comparisons between *mrc* and Col-0, determined using t-test on a linear model. P values < 0.05 are highlighted in yellow (statistically significant differences seen), and > 0.05 are highlighted in grey (not statistically significantly different). Corresponding data are presented in **Figure 4.12** and **Appendix 2 Fig. C**.

mrc vs Col-0	dpi	Full dataset	Individual replicates		
			A	B	C
Necrotic area	4	< 0.001	0.006	0.011	0.017
	5	0.010	0.070	0.278	0.083
Chlorotic area	4	0.017	0.043	0.384	0.047
	5	0.018	0.043	0.696	0.064
Lesion length	4	0.021	0.091	0.301	0.138
	5	0.285	0.221	0.903	0.385



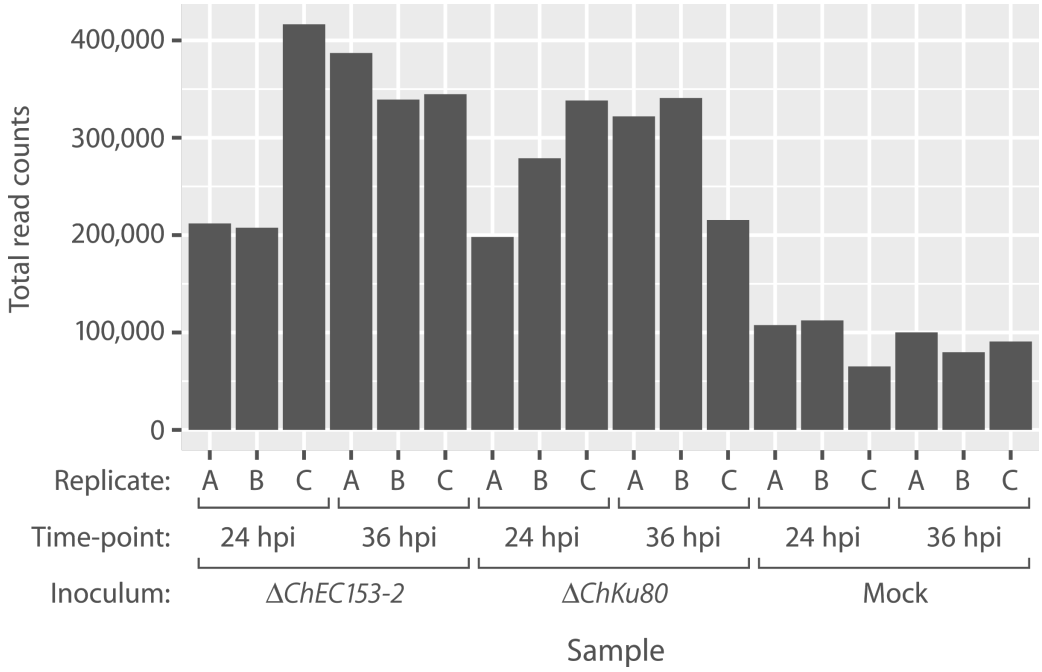
A2 Fig. D: ***crbic* Arabidopsis mutants do not have altered susceptibility to infection by *C. higginsianum* at 4 or 5 dpi.** Infection assay data for three individual replicates of the experiment are pooled and plotted, with no significant differences (ns, $p > 0.05$) seen between genotypes in terms of chlorotic lesion area, or lesion length, determined by a t-test using Satterthwaite's method on a linear mixed effects model with the individual replicates of the experiment included as a random factor. Values of n are shown above each plot for each of the three replicates. Mean lesion sizes for each are represented by grey diamonds. Data for necrotic lesion sizes are presented in **Figure 4.19**. Results of statistical tests are presented in **Appendix 2 Tab. E**.

A2 Tab. E: **Summary of statistical test results for *C. higginsianum* infection assay comparing *crbic* and Col-0 susceptibility.** At four- and five-days post inoculation (dpi) lesion sizes were quantified as necrotic lesions, chlorotic lesions or lesion lengths, and analysed using a t-test using Satterthwaite's method on a linear mixed effects model, with the independent replicates of the experiment included as a random factor (full dataset). Additionally, comparisons between genotypes for each individual replicate (A, B, and C) are shown: p-values are shown for each time-point for comparisons between *crbic* and Col-0, determined using a t-test on a linear model. p values < 0.05 are highlighted in yellow (statistically significant differences seen), and > 0.05 are highlighted in grey (not statistically significantly different). Corresponding data are presented in **Figure 4.19** and **Appendix 2 Fig. D**.

crbic vs Col-0	dpi	Full dataset, p =	Individual replicates, p =		
			A	B	C
Necrotic area	4	0.056	0.066	0.707	0.380
	5	0.524	0.307	0.295	0.765
Chlorotic area	4	0.306	0.286	0.736	0.849
	5	0.729	0.974	0.147	0.934
Lesion length	4	0.287	0.205	0.934	0.856
	5	0.898	0.498	0.426	0.675

Appendix 3

Appendix 3 (A3) contains supplementary RNA-Seq analysis data.



A3 Fig. A: **Total read counts (without normalisation) mapping to the *C. higginsianum* genome.** Read counts for three biological replicates for Arabidopsis treated with ΔChEC153-2, ΔChKu80, or mock (water) at 24- or 36-hours post inoculation. Biological replicates are indicated by letters A, B, and C. Reference genome ASM492035v1 (Tsushima et al., 2019a).

A3 Tab. A: **Differentially expressed Arabidopsis genes at 24- or 36-hours post inoculation with *ΔChKu80* relative to mock-inoculation.** AGI codes and names/descriptions (TAIR) of genes with different expression levels are listed, alongside their log_2 fold change (log_2 FC) in *ΔChKu80*-inoculated samples relative to mock-inoculated samples. Adjusted Wald test p-values (p-adj) are also listed. Differential expression defined as log_2 FC > 0.5 or <-0.5, p-adj < 0.1. Genes more highly expressed in samples inoculated with *ΔChKu80* relative to mock are highlighted in pink, while those that are less highly expressed in *ΔChKu80* samples are shown in green. Genes for which p-adj < 0.05 are highlighted in yellow.

Arabidopsis genes differentially expressed at 24 hpi with *ΔChKu80* vs. mock				
Gene ID	Gene name / description	Log_2 FC		p-adj
AT3G53400	*CPuORF46*	25.00		4.44E-08
AT4G11470	*CYSTEINE-RICH RLK 31* (*CRK31*)	4.50		8.99E-08
AT5G51030	NAD(P)-binding Rossmann-fold superfamily protein	-5.12		0.000905
AT4G34588	*CPuORF2*	25.02		0.000905
AT2G38995	O-acyltransferase (WSD1-like) family protein	7.90		0.000905
AT2G30540	CC-type glutaredoxin 7 (ROXY7)	2.78		0.000989

AT5G28630	Glycine-rich protein	1.83		0.00233
AT2G05812	Antisense long non-coding RNA	9.55		0.00309
AT2G43590	PR-3 like gene, chitinase, induced by infection	1.75		0.00856
AT1G27045	*ATHB54*, a member of the homeodomain leucine zipper (HD-Zip) family protein (*HB-54*)	-6.38		0.0144
AT3G26790	*FUSCA3* (*FUS3*) transcription factor	2.97		0.0156
AT3G52748	*SHORT OPEN READING FRAME 4* (*SORF4*)	1.73		0.0156
AT5G26270	Transmembrane protein	-4.19		0.0220
AT2G43880	Pectin lyase-like superfamily protein	-6.44		0.0413
AT1G10155	*PHLOEM PROTEIN 2-A10* (*PP2-A10*)	-1.62		0.0630
AT5G66640	*DA1-RELATED PROTEIN 3* (*DAR3*)	-1.64		0.0630
AT2G10930	Transmembrane protein	5.16		0.0630
AT1G01530	*AGAMOUS-LIKE 28* (*AGL28*)	6.35		0.0773
Arabidopsis genes differentially expressed at 36 hpi with *ΔChKu80* vs. mock				
Gene ID	Gene name / description	Log_2 FC		p-adj
AT3G49110	*PEROXIDASE 33* (*PRX33*)	18.02		3.21E-14
AT2G15220	Plant basic secretory protein (BSP) family protein	3.99		5.95E-12
AT3G53400	*CPuORF46*	-25.32		7.73E-09
AT4G13505	Antisense long non-coding RNA	-5.49		0.000124
AT1G07540	Telomere-binding protein, putative (*TRFL2*)	19.26		0.000920
AT2G43590	PR-3 like gene, chitinase, induced by infection	1.95		0.000920
AT2G40370	*LACCASE 5* (*LAC5*)	7.05		0.00137
AT3G49620	*DARK INDUCIBLE 11* (*DIN11*)	-18.41		0.00546
AT3G52748	*SHORT OPEN READING FRAME 4* (*SORF4*)	1.84		0.00608
AT5G44310	Late embryogenesis abundant protein (LEA) family protein	5.99		0.00668
AT4G11470	*CYSTEINE-RICH RLK 31* (*CRK31*)	3.01		0.00773
AT4G32510	HCO3- transporter family	9.07		0.00773
AT2G44400	Cysteine/Histidine-rich C1 domain family protein	-6.46		0.00878
AT5G17760	P-loop containing nucleoside triphosphate hydrolases superfamily protein	-1.10		0.00878
AT2G23910	NAD(P)-binding Rossmann-fold superfamily protein	-3.52		0.0147
AT1G16120	*WALL ASSOCIATED KINASE-LIKE 1* (*WAKL1*)	8.39		0.0179
AT1G79400	*CATION/H+ EXCHANGER 2* (*CHX2*)	-7.03		0.0245
AT2G15120	Pseudogene of Plant basic secretory protein (BSP) family protein	2.58		0.0288
AT3G22240	*CYSTEINE-RICH TRANSMEMBRANE MODULE 9* (*CYSTM9*)	-2.37		0.0288
AT3G01960	Hypothetical protein	0.96		0.0370
AT5G14960	*DP-E2F-like 2* (*DEL2*)	2.21		0.0370
AT1G18200	*RAB GTPASE HOMOLOG A6B* (*RABA6B*)	-1.81		0.0370
AT1G12940	*NITRATE TRANSPORTER2.5* (*NRT2.5*)	-8.29		0.0370

AT1G50770	Aminotransferase-like, plant mobile domain family protein	6.05		0.0370
AT3G60470	Transmembrane protein, putative (*DUF247*)	-4.90		0.0379
AT3G09940	*MONODEHYDROASCORBATE REDUCTASE 3* (*MDAR3*)	-1.76		0.0394
AT3G16670	Pollen Ole e 1 allergen and extensin family protein	-3.04		0.0394
AT3G02480	*ABA-RESPONSE PROTEIN* (*ABR*)	6.58		0.0394
AT3G15536	Unknown gene	-2.25		0.0439
AT4G36260	*SHI RELATED SEQUENCE 2* (*SRS2*)	-1.11		0.0499
AT2G43580	Chitinase family protein	2.78		0.0530
AT1G13440	*GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C2* (*GAPC2*)	-0.53		0.0584
AT4G17030	*EXPANSIN-LIKE B1* (*EXLB1*)	2.23		0.0584
AT5G07640	RING/U-box superfamily protein	1.47		0.0642
AT1G13550	Hypothetical protein (*DUF1262*)	-4.95		0.0642
AT3G61198	N/A	-3.91		0.0739
AT5G10760	*APOPLASTIC, EDS1-DEPENDENT 1* (*AED1*)	-1.78		0.0739
AT4G10500	*DMR6-LIKE OXYGENASE 1* (*DLO1*)	-2.25		0.0744
AT4G01140	Transmembrane protein, putative (*DUF1191*)	6.36		0.0744
AT1G29680	Histone acetyltransferase (*DUF1264*)	19.54		0.0810
AT3G19550	Glutamate racemase	-1.52		0.0842
AT1G07120	*IPGA1-LIKE1, CHUP1-LIKE PROTEIN* (*IPGAL1*)	1.66		0.0995