UNIVERSITY OF CALIFORNIA
RIVERSIDE

A Chromosome-Scale Reference Genome Provides Insight into Genome Biology and Fungicide Resistance in *Phytophthora infestans*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Genetics Genomics & Bioinformatics

by

Michael Edmund Hans Matson

September 2018

Dissertation Committee:
Dr. Howard S Judelson, Chairperson
Dr. Timothy J Close
Dr. Thomas Girke

The Dissertation of Michael Edmund Hans Matson is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

This work was supported by grant 2011-68004-30154 from the United States Department of Agriculture National Institute of Food and Agriculture to Howard S. Judelson. We thank members of the USAblight team for providing many of the isolates used in this study. We are especially grateful to Steve Whisson for generously sharing unpublished data and to Kevin Myers and Dr. William Fry at Cornell for helping to maintain and ship isolates, and for their contributions to the metalaxyl field experiments.

I would like to thank Dr. Audrey Ah Fong of the Judelson laboratory for the preparation of the DNA used for third generation sequencing, the laboratory of Dr. Stefano Londardi for development and assistance with the Novo&Stitch script to prepare the input assembly for chromosome scaffolding, and Qihua Liang from the Londardi lab for assisting wtih gene annotations.

Finally, I am eternally grateful for the advice and support given by my friends during my time at UCR, from the countless anonymous contributions to my code-writing skills and programming scripts from users of the many bioinformatics help websites, and for the guidance of Dr. Judelson, who brought me to the finish line.

Chapter I of this dissertation appears in full in *Phytophthology* (2015) with Kevin Myers and Dr. William Fry and Dr. Howard Judelson as co-authors.

Dedication

I would like dedicate this dissertation to my brother, Conor, who inspires me with his dedication, to my parents, Michael and Eileen, who have never faltered in their support for my scientific dreams, and to my other halves in Riverside, Kelley and Kepler, who not only endured the writing of this dissertation, but provided the best support a stressed-out graduate student could have ever asked for.

ABSTRACT OF THE DISSERTATION

A Chromosome-Scale Reference Genome Provides Insight into Genome Biology and Fungicide Resistance in *Phytophthora infestans*

by

Michael Edmund Hans Matson

Doctor of Philosophy, Genetics Genomics & Bioinformatics
University of California, Riverside, September 2018
Dr. Howard S. Judelson, Chairperson

The oomycetes are a group of filamentous diploid organisms that superficially resemble true fungi, but instead belong to the kingdom Stramenopiles. Many oomycetes are responsible for devastating diseases of plants, including *Phytophthora infestans,* the cause of late blight on potato and tomato, and the notorious agent behind the Irish potato famine in the mid 1840s. *P. infestans* remains as destructive today as it was in the past due to its ability to spread through explosively through susceptible fields given ideal conditions and rapid adaptability to environmental stresses, host defenses, and chemical control. To understand this remarkable adaptability, we examined laboratory-generated sexual population and observed a major genetic locus conferring resistance to the popular fungicide metalaxyl segregating through the progeny. We found that this locus was genetically separate from a previously described

locus shown to confer resistance to metalaxyl. Identification of the genes underlying our new major locus was unsuccessful due to the unassembled nature of the currently available reference genome for *P. infestans*. We sought to construct a new reference genome using a combination of third generation sequencing and scaffolding technology, as well as utilizing two sexual populations of *P. infestans* to construct genetic maps to assembly of chromosome-sized scaffolds. A large interval of markers strongly associated with metalaxyl resistance was identified at the end of chromosome 3. 130 genes were identified in this interval, several of which may play a role in the resistance phenotype. This new reference genome also allowed us to examine the extent of structural variation and genome plasticity within our sexual populations, and we observed considerable plasticity in the forms of whole genome triploidy, tetraploidy, and individual aneuploid chromosomes. Transcriptomic analysis similarly revealed extensive variation among progeny, and our results conclude that sexual reproduction has the potential to induce massive levels of variation in a short period of time in *P. infestans.*

Table of contents

Introduction 1

References 13

Chapter I Metalaxyl resistance in *Phytophthora infestans*: assessing role of RPA190 gene and diversity within clonal lineages

Abstract 23

Introduction 24

Materials and methods 27

Results 31

Discussion 51

References 55

Chapter II Assembly of a new *Phytophthora infestans* reference genome

Abstract 60

Introduction 61

Materials and methods 67

Results 79

Discussion 96

References 111

Chapter III Functional consequences of genome plasticity and transcriptional variation in *Phytophthora infestans* sexual populations

Abstract 120

Introduction 121

Materials and methods 130

Results 144

Discussion 176

References 200

Conclusions 212

References 223

List of tables

Chapter I

Table 1.1	RNA190 genotype in 22 isolates of *P. infestans*	39
Table 1.2	Subclsses within lineages based on RPA190 SNPs	44

Chapter II

Table 2.1	Summary and statistics of the input assemblies for stitched7	80
Table 2.2	Summary of genetic markers and individual genotypes curated in preparation for the final genetic maps	85
Table 2.3	Genetic size of parent 1306 over four rounds of genotype correction and curation	86
Table 2.4	Summary of the markers and linkage groups contributing to the assembly of the pseudochromosomes	87
Table 2.5	Genetic sizes of linkage groups for each parent	89
Table 2.6	Gene count and density along each pseudochromosome assembly	99
Table 2.7	Variant numbers and variant density for each pseudochromosome and unassembled contigs	103

Chapter III

Table 3.1	Phenotypic summary of all sequenced CrossM progeny	147
Table 3.2	Variant numbers of selected individuals per chromosome	148
Table 3.3	Mapping locations of all RNA Polymerase I subunits	152
Table 3.4	Summary of qPCR and HRM markers	153
Table 3.5	List and annotation of all genes in the chromosome 3 deletion interval defined by fast-growing isolate M64F	154
Table 3.6	Numbers of genes differentially expressed in relation to metalaxyl	175
Table 3.7	GO term enrichments for genes differentially expressed in response to metalaxyl treatment	177
Table 3.8	GO term enrichment for genes differentially expressed between metalaxyl resistant and sensitive progeny	178

List of figures

Chapter I

Figure 1.1	Effect of metalaxyl in field assay	32
Figure 1.2.	Methods for genotyping RPA190	36
Figure 1.3	Characteristics of US-23 and US-24 isolates	41
Figure 1.4	Metalaxyl resistance and 1145 genotype of progeny from 8811 × E13a cross	46
Figure 1.5	Metalaxyl resistance and genotypes of 618 × 1306 cross	49

Chapter II

Figure 2.1	Flowchart of the process used to curate markers used as input for the genetic maps	74
Figure 2.2	Examples of genotypes to be corrected in contig20	83
Figure 2.3	ALLMAPS representation of the genetic and physical structure of pseudochromosome 14	91
Figure 2.4	Fifteen pseudochromosome assemblies	93
Figure 2.5	Chromosome comparisons and annotations	97

Chapter III

Figure 3.1	Graphs of the four main phenotypes measured for the CrossM progeny	145
Figure 3.2	Genome-wide analysis of markers linked to metalaxyl resistance for all 82 sequenced progeny	150
Figure 3.3	QTL analysis of other phenotypes	156
Figure 3.4	Confirmation of increased metalaxyl resistance in fast growing single-zoospore derivatives progeny	158
Figure 3.5	The sectors growing fast on metalaxyl media share a deletion at the end of chromosome 3	161
Figure 3.6	qPCR and High Resolution Melt (HRM) analysis of the deleted region at the end of chromosome 3	164
Figure 3.7	Progeny from CrossM and Cross20 show a diversity of karyotypes.	167
Figure 3.8	Diversity of CrossM transcriptome profiles	173

Introduction

Phytophthora infestans and its populations

The oomycete pathogen *Phytophthora infestans* is a major pathogen of potato and tomato production and annually inflicts billions of dollars in crop loss and damage globally (Haverkort et al. 2008). While phylogenetically unrelated to true fungi (Gunderson et al. 1987), oomycetes nonetheless grow in a similar manner to fungi through filamentous hyphae and the production of vast numbers of spores which either directly germinate, or release swimming zoospores. These asexual sporangia and zoospores, as opposed to sexual oospores, make up the vast majority of spore types produced and given optimal cool and humid environmental conditions contribute to the potential for explosive epidemics overrunning crop fields (Mizubuti and Fry 1998; Seidl Johnson et al. 2015).

Due to this remarkable potential to spread, oomycetes often persist over large geographical areas as closely related asexual clonal lineages (Danies et al. 2013; Fry et al. 2013; Goodwin et al. 1994; Hu et al. 2012; Li et al. 2012; Yoshida et al. 2013; Zhu et al. 2015). The first major pandemic lineage of *P. infestans,* "HERB-1/FAM-1", was identified and resequenced from herbarium samples archived from around the time of the great famine in Ireland (Saville et al. 2016; Yoshida et al. 2013). HERB-1, however, was displaced as the dominant global lineage by another, named US-1, in the mid 20th century

(Goodwin et al. 1994; Yoshida et al. 2013). US-1 persisted well into the late 20th century as the dominant global genotype. Since then, multiple migrations out of the center of origin in Mexico (Wang et al. 2017), and the occasional sexual population elsewhere (Danies et al. 2014), have given rise to many lineages, each with their own particular characteristics such as preferred environmental conditions (Mizubuti and Fry 1998; Seidl Johnson et al. 2015), preferred hosts (Marcin Nowicki 2012; Matson et al. 2015), and fungicide sensitivity (Childers et al. 2014; Danies et al. 2013; Matson et al. 2015; Saville 2014). Additionally, variants in the mitochondrial genome have lead to the existence of four major mtDNA categories; Ia, Ib, IIa, and IIb (Griffith and Shaw 1998). The majority of recent North American lineages, as well as a recent aggressive European lineage, 13_A2, contain the Ia genotype (Danies et al. 2013; Yoshida et al. 2013), while the earlier dominant US-1 lineage contained the Ib genotype (Yoshida et al. 2013). While oomycetes have a diploid genome in the vegetative life stage, unlike most true fungi, most of the surveyed clonal lineages have shifted their genomes to a triploid state (Li et al. 2017; Therrien 1989; Yoshida et al. 2013). It is theorized that extra chromosome copies act as a buffer for the acquisition of deleterious variants, while allowing for rapid experimentation of new, possibly beneficial, variants (Li et al. 2017). Noticeable exceptions to this trend include the first major lineage HERB-1 and a recent North American lineage US-22, which were both diploid (Yoshida et al. 2013).

While virtually every aggressive epidemic of *P. infestans* can be traced to a dominant asexual lineage, sexual recombination does occur infrequently in certain geographic locations and environments. *P. infestans* is a heterothallic organism, and isolates of both A1 and A2 mating types must interact in order for the sexual life stage to occur and generate a thick-walled sexual oospore. Oospores differ from their asexual spore counterparts in that the former survives much longer in adverse conditions, allowing overwintering and survival for many years in fields. Evidence of a sexual population include measuring the level of genetic diversity, the closeness of genotype frequencies to Hardy-Weinberg equilibrium, the ratio of mating types, and the frequency of oospores in the field. One such location of persistent sexual populations is the center of origin for *P. infestans* in central Mexico (Goodwin et al. 1992b; Goss et al. 2014; Wang et al. 2017). Another major sexual population lies in the Nordic countries as well as an occasional population in The Netherlands (Drenth et al. 1994; Li et al. 2012), the hardiness of the sexual oospore, compared to the asexual sporangia zoospores, allow for an enhanced survival in the face of colder weather (Brurberg et al. 2011). Elsewhere in the world, sexual populations tend to exist only temporarily, such as in the northeastern United States in 2010-11 (Danies et al. 2014).

Early history of *Phytophthora infestans* research and control

The study of oomycetes first gained importance to the research community at the time following the devastating famine throughout Ireland in the mid 1840s. In fact, it can be argued that the work of German botanist Anton de Bary on the cause of the famine and blight disease effectively started the entire discipline of plant pathology (de Bary 1861, 1863, 1876). He demonstrated that inoculation of the fungus itself could solely be responsible for the disease symptoms and named this organism "The Plant Destroyer", or *Phytophthora infestans.* This cause and effect approach was famously elaborated by Robert Koch in the late 19th century, and "Koch's Postulates" earned him the 1905 Nobel prize in Physiology and Medicine. Research into the development and nature of *P. infestans* progressed throughout the early 20th century, with much work uncovering relationships between strains infecting tomatoes and/or potatoes (Small 1938), the biology of spore germination and development (Blackwell and Waterhouse 1931; Glendinning et al. 1963; Melhus 1915; Smoot et al. 1958), identification of environmental factors leading to epidemics (Beaumont 1947; Wallin 1962; Weille 1964), and culturing the pathogen for use in laboratory experiments (Crosier 1933; Dickinson and Keay 1948).

Later in the 20th century saw the introduction of one of the most effective fungicides used to control oomycetes, the phenylamide compound metalaxyl

(Frost and Dowley 1979; Urech et al. 1977). The systemic-acting phenylamide fungicides were revolutionary in their preventative and curative properties, and could effectively control US-1, the dominant global lineage at the time. However, as with many means of control against pathogens, resistance to metalaxyl was soon discovered in the field (Davidse et al. 1981; Dowley and Osullivan 1981). Biochemical studies soon determined that the route of action likely involved the metalaxyl molecule inhibiting ribosomal RNA production (Davidse et al. 1983; Fisher and Hayes 1984), and that mutations in the binding site likely contribute to the acquisition of resistance. Genetically, resistance typically segregates through sexual crosses as a single major locus (Bhat et al. 1993; Fabritius et al. 1997; Lee et al. 1999; Shattock 1988), although the presences of several minor loci modify the final phenotype to behave with incomplete dominance. In addition, the observation of multiple markers linked to resistance across different sexual crosses implies the existence of multiple major loci conferring resistance to metalaxyl globally (Lee et al. 1999).

As with the rest of the life sciences, research into the oomycetes began to enter the genomics era at the turn of the millennium. *P. infestans* was first successfully transformed in 1991 via protoplast transformation (Judelson et al. 1991), facilitating a wealth of new experimental possibilities such as gene silencing based on RNAi methods (van West et al. 1999; Whisson et al. 2005). One ability crucial to managing outbreaks is successfully identifying the lineage

causing an epidemic. Among the first successful markers to reveal molecular polymorphism between different isolates was the RFLP marker RG57 (Goodwin et al. 1992a), the glucose-6-phosphate isomerase (Gpi) allozyme marker (Goodwin et al. 1995) and the mitochondrial lineage genotypes (Griffith and Shaw 1998). Later, a set of 12 SSR microsatellite markers greatly enhanced the resolution of lineage and genetic diversity determination of *P. infestans* (Lees et al. 2006). The next generation of higher-throughput markers, AFLPs, allowed for the construction of a high density genetic map (van der Lee et al. 2004) organized into about 13 major linkage groups. This number of linkage groups was roughly in line with the predicted number of haploid chromosomes determined by direct staining and counting (Ritch and Daggett 1995), and is further confirmed by more recent studies correlating flow cytometry and direct counting of chromosomes (Bertier et al. 2013; Catal et al. 2010; Li et al. 2017). Currently, the best estimate of the haploid number of chromosomes lingers around 14, though it would not be surprising if this number changes across different lineages and populations.

Phytophthora infestans in the genomics age

Phytophthora research truly entered the genomic era with the identification of over 75,000 Expressed Sequence Tags (ESTs) (Randall et al. 2005), allowing

an estimation of the presence of 18,256 genes. This dataset facilitated development of microarrays and the first major insights into transcriptional dynamics during the *P. infestans* life cycle. In 2006, the full-genome sequences were made available for *Phytohthora sojae* and *Phytophthora ramorum* were published (Tyler et al. 2006) using BAC and Sanger sequencing approaches, representing the first oomycetes to have their genomes sequenced. Genome size estimates for *P. sojae* were 95 Mb and 19,027 predicted genes, while the *P. ramorum* genome was somewhat smaller at 65 Mb and 15,743 predicted genes. A physical map of the *P. sojae* sequence resolved 132 Mb of sequence across 79 superscaffolds (Zhang et al. 2006). Following shortly after was the genome sequence for *P. infestans* (Haas et al. 2009) using a sequencing approach similar to the *P. sojae* and *P. ramorum* genome projects. The *P. infestans* genome consisted of 229 Mb of assembled sequence, out of an estimated genome size of 240 Mb, across 4921 scaffolds with an N50 scaffold value of 1.5 Mb. Initially, 17,797 genes were identified, and of these, 9,583 have orthologs in the other two previously sequenced species. In 2012, the genome sequence for *Phytophthora capsici* was assembled via *de novo* assembly of mate pair libraries sequenced via 454 pyrosequencing, resulting in 917 scaffolds, 64 Mb of sequence and 19,805 predicted genes (Lamour et al. 2012). Additionally, targeted RAD sequencing of 60 parents and F1 progeny of a sexual cross allowed 108 of these scaffolds to be incorporated into 18 linkage groups to a genetic map size of 1,654

cM. Since 2010, many more oomycete genomes have been published, include those belonging to the obligate biotrophs *Hyaloperonospora arabidopsidis* (Baxter et al. 2010)*, Plasmopara halstedii* (Sharma et al. 2015), *Pseudoperonospora cubensis* (Burkhardt et al. 2015; Tian et al. 2011), *Peronospora tabacina* (Derevnina et al. 2015)*, Plasmopara viticola* (Dussert et al. 2016; Yin et al. 2017), *Albugo candida* (Links et al. 2011), *Albugo laibachii* (Kemen et al. 2011), the necrotroph *Pythium ultimum* (Lévesque et al. 2010), and another *Phytophthora* species, *Phytophthora lateralis (Quinn et al. 2013).* In addition, several oomycetes are having their genomes re-assembled using Pacific Biosciences long read technology, including *Phytophthora capsici* and *Bremia lactucae*.

One useful outcome of these genomic resources is reconstructing the evolutionary history of *Phytophthora*. By mtDNA and second generation Illumina sequencing of 36 isolates originating from North America, South America, and Europe, Yoshida et al (2013) predicted that *P. infestans* branched off from its nearest relative, *Phytophthora mirabilis,* around the year 750 through estimating mutation rates. Next-generation sequencing has also helped to inform on the debated origin of *P. infestans.* Historically, the Toluca Valley of central Mexico was seen as the origin of the species (Gallegly and Galindo 1958; Goodwin et al. 1992b) due to diverse genotypes and a 1:1 ratio of mating types present. However, an investigation of mtDNA haplotypes lead to the conclusion of the

Andes mountains in South American being the true origin (Gomez-Alpizar et al. 2007), coinciding with the center of origin for the domesticated potato. Later, the Mexican origin was again proposed (Goss et al. 2014), as was again the Andean origin (Martin et al. 2016), and finally most recently the Mexican origin (Wang et al. 2017).

Aside from population level studies, modern genomic tools have also allowed a more complete understanding of oomycete transcriptional dynamics via RNA sequencing. Some important questions included: what genes play a major role in promoting life stage transitions, and what genes are necessary for successful infection of hosts? A recent study addressed this by examining transcriptional differences between life phases within and between *P. infestans* and *Pythium ultimun,* as well as how their transcriptomes are remodeled in response to potato tuber infection (Ah-Fong et al. 2017). Additionally, a transcriptomics study of *P. infestans* isolates which seemed to acquire increased resistance to the fungicide metalaxyl revealed upregulation of a number of genes which included the ATP Binding Cassette (ABC) transporters (Childers et al. 2014). ABC transporters are thought to be involved in increased resistance to fungicides such as metalaxyl (Judelson and Senthil 2006), and the Childers study demonstrated the utility of transcriptomics assays to answer biological questions in the oomycetes outside of development and pathogenesis.

Scope of dissertation

The primary driving factor of my research is to provide better resources to growers, extension personnel, and academic scientists on combating the late blight disease of potato and tomato. While an important aspect of combating plate disease includes preventative measures such as resistance breeding, this remains a significant challenge for late blight and *P. infestans.* Chemical control is still the most effective preventative measure and reactionary response to late blight, and thus represents an important area of research. My dissertation primarily focuses on investigating the mechanisms of resistance of the fungicide metalaxyl, as well as developing additional genomic resources further assisting research into many aspects of *Phytophthora* biology as well as towards helping growers themselves combat the disease.

Chapter I addresses a 2014 paper by another group (Randall et al. 2014) which described the large subunit of RNA Polymerase I (RPA190) as containing a particular mutation correlated with resistance to metalaxyl. 86% of their resistant isolates collected from the UK contained this resistant allele, and they further showed that complementation transformation of a sensitive isolate with a resistant allele seemed to rescue the metalaxyl resistance phenotype. We show that in a cross generated between a resistant Mexican isolate and a sensitive Californian isolate, the previously described resistance-conferring mutation is

completely unlinked to the resistance phenotype. Additionally, a diverse collection of clonal lineages collected across North America showed no association between the resistance phenotype and the presence of the resistant allele.

Chapter II details the construction of a new reference genome to improve upon the existing *P. infestans* reference genome (Haas et al. 2009) assembled by the Broad Institute. The impetuous for this effort was due to the inability to use the Broad reference to effectively map a Quantitative Trait Locus (QTL) behind the metalaxyl resistance phenotype. We employed third generation sequencing and scaffold technology in the form of Pacific Biosciences long-read sequencing, a Dovetail Genomics long-insert mate-pair Chicago library, and a Bionano Genomics optical map to generate a new draft assembly. The draft assembly was used as an input for the construction of a fully-assembled and chromosome-scale reference sequence based on genetic maps derived from two Illumina-sequenced sexual *P. infestans* populations. Comprising of 218 Mb of assembled sequence and 15 pseudochromosomes, this is the first chromosome-scale reference genome based on third generation sequencing technology for an oomycete.

Chapter III applies the new reference genome from chapter II to explore QTL mapping, genome plasticity, and transcriptional diversity among the sexual populations. We were able to identify an 8 Mb region of high statistical

association with resistance to metalaxyl at the end of chromosome 3, but were unable to identify loci associated with three other ergonomically important traits. The use of a chromosome-scale reference genome allowed us to examine copy number and structural diversity among our progeny on the whole chromosome scale, and we observed a genomic landscape filled with triploid and tetraploid progeny, aneuploid chromosomes, and chromosome truncations. This analysis also lead us to identify a 2 Mb hemizygous deletion at the end of chromosome 3 which conferred increased resistance to metalaxyl, evidence of *P. infestans* utilizing structural variation to its advantage. Our transcriptomic data for 34 parents and progeny similarly revealed a diverse landscape in which few progeny cluster together in terms of their transcriptomic profile, likely a result of the considerable structural diversity and plasticity observed among the progeny. Overall, these results underscore the potential for rapid diversification among progeny in sexual populations of *P. infestans*. They highlights the need for close observation of contemporary clonal lineages of opposite mating types when they are detected in close geographic proximity to adequately prepare for sexual progeny that may potentially overcome established plant and chemical defenses.

References

Ah-Fong, A. M. V., Shrivastava, J., and Judelson, H. S. 2017. Lifestyle, gene gain and loss, and transcriptional remodeling cause divergence in the transcriptomes of *Phytophthora infestans* and *Pythium ultimum* during potato tuber colonization. BMC Geno 18:764.

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., 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., Win, J., Wood, J., Clifton, S. W., Rogers, J., Van den Ackerveken, G., Jones, J. D. G., McDowell, J. M., Beynon, J., and Tyler, B. M. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science 330:1549-1551.

Beaumont, A. 1947. The dependence on the weather of the dates of outbreak of potato blight epidemics. Trans of the B Mycol Soc 31:45-53.

Bertier, L., Leus, L., D'Hondt, L., de Cock, A. W., and Hofte, M. 2013. Host adaptation and speciation through hybridization and polyploidy in *Phytophthora*. PLoS One 8:e85385.

Bhat, R. G., Mcblain, B. A., and Schmitthenner, A. F. 1993. The Inheritance of Resistance to Metalaxyl and to Fluorophenylalanine in Matings of Homothallic *Phytophthora Sojae*. Mycolo Res 97:865-870.

Blackwell, E. M., and Waterhouse, G. M. 1931. Spores and spore germination in the genus *Phytophthora*. Trans of the B Mycol Soc 15:294-310.

Brurberg, M. B., Elameen, A., Le, V. H., Naerstad, R., Hermansen, A., Lehtinen, A., Hannukkala, A., Nielsen, B., Hansen, J., Andersson, B., and Yuen, J. 2011. Genetic analysis of *Phytophthora infestans* populations in the Nordic European countries reveals high genetic variability. Fun Biol 115:335-342.

Burkhardt, A., Buchanan, A., Cumbie, J. S., Savory, E. A., Chang, J. H., and Day, B. 2015. Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. Mol Plan Micro Interac 28:298-309.

Catal, M., King, L., Tumbalam, P., Wiriyajitsomboon, P., Kirk, W. W., and Adams, G. C. 2010. Heterokaryotic nuclear conditions and a heterogeneous nuclear population are observed by flow cytometry in *Phytophthora infestans*. Cytometry A 77A:769-775.

Childers, R., Danies, G., Myers, K. L., Fei, Z., Small, I. M., and Fry, W. 2014. Acquired resistance to mefenoxam in sensitive isolates of *Phytophthora infestans*. Phytopathology 105:342-349.

Crosier, W. 1933. Culture of *Phytophthora infestans*. Phytopathology 23:713-720.

Danies, G., Small, I. M., Myers, K., Childers, R., and Fry, W. E. 2013. Phenotypic characterization of recent clonal lineages of *Phytophthora infestans* in the United States. Plant Dis 97:873-881.

Danies, G., Myers, K., Mideros, M. F., Restrepo, S., Martin, F. N., Cooke, D. E., Smart, C. D., Ristaino, J. B., Seaman, A. J., Gugino, B. K., Grunwald, N. J., and Fry, W. E. 2014. An ephemeral sexual population of *Phytophthora infestans* in the Northeastern United States and Canada. PLoS One 9:e116354.

Davidse, L. C., Hofman, A. E., and Velthuis, G. C. M. 1983. Specific interference of metalaxyl with endogenous RNA polymerase activity in isolated-nuclei from *Phytophthora-megasperma f. sp medicaginis*. Exp Mycol 7:344-361.

Davidse, L. C., Looijen, D., Turkensteen, L. J., and Vanderwal, D. 1981. Occurrence of metalaxyl resistant strains of *Phytophthora infestans* in Dutch potato fields. N J of Plant Path 87:65-68.

de Bary, A. 1861. Die gegenwa¨rtig herrschende Kartoffelkrankenheit, ihre Ursache und ihre Verhu¨tung (The current prevailing potato disease, its cause and treatment). Felix, Leipzig.

de Bary, A. 1863. Recherches sur le développement de quelques champignons parasites (Research on the development of selected parasitic fungi). Ann des Sci Nat Botan 20:5-148.

de Bary, A. 1876. Researches Into the nature of the potato-fungus (*Phytophthora Infestans*). J Royal Ag Soc of E 12:239-269.

Derevnina, L., Chin-Wo-Reyes, S., Martin, F., Wood, K., Froenicke, L., Spring, O., and Michelmore, R. 2015. Genome sequence and architecture of the tobacco downy mildew pathogen *Peronospora tabacina*. Mol Plan Micro Interac 28:1198-1215.

Dickinson, S., and Keay, M. A. 1948. Growth of *Phytophthora infestans* (Mont.) de Bary on artificial media. Nature 162:32.

Dowley, L. J., and Osullivan, E. 1981. Metalaxyl resistant strains of *Phytophthora-Infestans* (Mont) de Bary in Ireland. Potato Res 24:417-421.

Drenth, A., Tas, I. C. Q., and Govers, F. 1994. DNA fingerprinting uncovers a new sexually reproducing population of *Phytophthora infestans* in the Netherlands. Euro J of Plan Path 100:97-107.

Dussert, Y., Gouzy, J., Richart-Cervera, S., Mazet, I. D., Deliere, L., Couture, C., Legrand, L., Piron, M. C., Mestre, P., and Delmotte, F. 2016. Draft genome sequence of *Plasmopara viticola*, the grapevine downy mildew pathogen. Genome Announc 4:e00987-00916.

Fabritius, A. L., Shattock, R. C., and Judelson, H. S. 1997. Genetic analysis of metalaxyl insensitivity loci in *Phytophthora infestans* using linked DNA markers. Phytopathology 87:1034-1040.

Fisher, D. J., and Hayes, A. L. 1984. Studies of mechanisms of metalaxyl fungitoxicity and resistance to metalaxyl. Crop Protection 3:177-185.

Frost, C., and Dowley, L. 1979. Metalaxyl-a new fungicide. Technical Bulletin, agriculture series, an foras taluntais, Dublin 4 3:1-6.

Fry, W. E., McGrath, M. T., Seaman, A., Zitter, T. A., McLeod, A., Danies, G., Small, I. M., Myers, K., Everts, K., Gevens, A. J., Gugino, B. K., Johnson, S. B., Judelson, H., Ristaino, J., Roberts, R., Secor, G., Seebold, K., Snover-Clift, K., Wyenandt, A., Grunwald, N. J., and Smart, C. D. 2013. The 2009 late blight pandemic in the eastern United States - causes and results. Plant Dis 97:296-306.

Gallegly, M., and Galindo, J. 1958. Mating types and oospores of *Phytophthora infestans* in nature in Mexico. Phytopathology 48:274-277.

Glendinning, D., Macdonald, J. A., and Grainger, J. 1963. Factors affecting the germination of sporangia in *Phytophthora infestans*. Trans of the B Mycol Soc 46:595-603.

Gomez-Alpizar, L., Carbone, I., and Ristaino, J. B. 2007. An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. Proc Natl Acad Sci USA 104:3306-3311.

Goodwin, S. B., Drenth, A., and Fry, W. E. 1992a. Cloning and genetic analyses of two highly polymorphic, moderately repetitive nuclear DNAs from *Phytophthora infestans*. Curr Gen 22:107-115.

Goodwin, S. B., Cohen, B. A., and Fry, W. E. 1994. Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. Proc Natl Acad Sci USA 91:11591-11595.

Goodwin, S. B., Schneider, R. E., and Fry, W. E. 1995. Use of cellulose-acetate electrophoresis for rapid identification of allozyme genotypes of *Phytophthora infestans*. Plant Dis 79:1181-1185.

Goodwin, S. B., Spielman, L. J., Matuszak, J. M., Bergeron, S. N., and Fry, W. E. 1992b. Clonal diversity and genetic differentiation of *Phytophthora infestans* populations in northern and central Mexico. Phytopathology 82:955-961.

Goss, E. M., Tabima, J. F., Cooke, D. E., Restrepo, S., Fry, W. E., Forbes, G. A., Fieland, V. J., Cardenas, M., and Grunwald, N. J. 2014. The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. Proc Natl Acad Sci USA 111:8791-8796.

Griffith, G. W., and Shaw, D. S. 1998. Polymorphisms in *Phytophthora infestans*: four mitochondrial haplotypes are detected after pcr amplification of dna from pure cultures or from host lesions. App and Envi Microbio 64:4007-4014.

Gunderson, J. H., Elwood, H., Ingold, A., Kindle, K., and Sogin, M. L. 1987. Phylogenetic relationships between chlorophytes, chrysophytes, and oomycetes. Proc Natl Acad Sci USA 84:5823-5827.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., 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., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., 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., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Visser, R. G. F., and van der Vossen, E. A. G. 2008. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. Potato Res 51:47-57.

Hu, C. H., Perez, F. G., Donahoo, R., McLeod, A., Myers, K., Ivors, K., Secor, G., Roberts, P. D., Deahl, K. L., Fry, W. E., and Ristaino, J. B. 2012. Recent genotypes of *Phytophthora infestans* in the eastern united states reveal clonal populations and reappearance of mefenoxam sensitivity. Plant Dis 96:1323-1330.

Judelson, H. S., and Senthil, G. 2006. Investigating the role of ABC transporters in multifungicide insensitivity in *Phytophthora infestans*. Mol Plan Path 7:17-29.

Judelson, H. S., Tyler, B. M., and Michelmore, R. W. 1991. Transformation of the oomycete pathogen, *Phytophthora infestans*. Mol Plan Micro Interac 4:602-607.

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., and Jones, J. D. 2011. Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. PLoS Biol 9:e1001094.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S.,

Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. Mol Plan Micro Interac 25:1350-1360.

Lee, T. Y., Mizubuti, E., and Fry, W. E. 1999. Genetics of metalaxyl resistance in *Phytophthora infestans*. Fun Genet Biol 26:118-130.

Lees, A. K., Wattier, R., Shaw, D. S., Sullican, L., and L., C. D. E. 2006. Novel microsatellite markers for the analysis of *Phytophthora infestans* populations. Plant Pathol 55:311-319.

Lévesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., Raffaele, S., Robideau, G. P., Thines, M., Win, J., Zerillo, M. M., Beakes, G. W., Boore, J. L., Busam, D., Dumas, B., Ferriera, S., Fuerstenberg, S. I., Gachon, C. M., Gaulin, E., Govers, F., Grenville-Briggs, L., Horner, N., Hostetler, J., Jiang, R. H., Johnson, J., Krajaejun, T., Lin, H., Meijer, H. J., Moore, B., Morris, P., Phuntmart, V., Puiu, D., Shetty, J., Stajich, J. E., Tripathy, S., Wawra, S., van West, P., Whitty, B. R., Coutinho, P. M., Henrissat, B., Martin, F., Thomas, P. D., Tyler, B. M., De Vries, R. P., Kamoun, S., Yandell, M., Tisserat, N., and Buell, C. R. 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. Genome Biol 11:R73.

Li, Y., Shen, H., Zhou, Q., Qian, K., van der Lee, T., and Huang, S. W. 2017. Changing ploidy as a strategy: the Irish potato famine pathogen shifts ploidy in relation to its sexuality. Mol Plan Micro Interac 30:45-52.

Li, Y., van der Lee, T. A., Evenhuis, A., van den Bosch, G. B., van Bekkum, P. J., Forch, M. G., van Gent-Pelzer, M. P., van Raaij, H. M., Jacobsen, E., Huang, S. W., Govers, F., Vleeshouwers, V. G., and Kessel, G. J. 2012. Population dynamics of *Phytophthora infestans* in the Netherlands reveals expansion and spread of dominant clonal lineages and virulence in sexual offspring. G3 2:1529-1540.

Links, M. G., Holub, E., Jiang, R. H., Sharpe, A. G., Hegedus, D., Beynon, E., Sillito, D., Clarke, W. E., Uzuhashi, S., and Borhan, M. H. 2011. De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. BMC Geno 12:503.

Marcin Nowicki, M. R. F., Marzena Nowakowska, Elzbieta U. Kozik. 2012. Potato and tomato late blight caused by *Phytophthora infestans*: an overview of pathology and resistance breeding. Plant Dis 96:4-17.

Martin, M. D., Vieira, F. G., Ho, S. Y., Wales, N., Schubert, M., Seguin-Orlando, A., Ristaino, J. B., and Gilbert, M. T. 2016. Genomic characterization of a South American *Phytophthora* hybrid mandates reassessment of the geographic origins of *Phytophthora infestans*. Mol Biol Evol 33:478-491.

Matson, M. E. H., Small, I. M., Fry, W. E., and Judelson, H. S. 2015. Metalaxyl resistance in *Phytophthora infestans*: assessing role of RPA190 gene and diversity within clonal lineages. Phytopathology 105:1594-1600.

Melhus, I. E. 1915. Germination and infection with the fungus of the late blight of potato (*Phytophthora infestans*). Agricultural Experiment Station of the University of Wisconsin, Madison.

Mizubuti, E. S., and Fry, W. E. 1998. Temperature effects on developmental stages of isolates from three clonal lineages of *Phytophthora infestans*. Phytopathology 88:837-843.

Quinn, L., O'Neill, P. A., Harrison, J., Paskiewicz, K. H., McCracken, A. R., Cooke, L. R., Grant, M. R., and Studholme, D. J. 2013. Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. FEMS Micro Lett 344:179-185.

Randall, E., Young, V., Sierotzki, H., Scalliet, G., Birch, P. R., Cooke, D. E., Csukai, M., and Whisson, S. C. 2014. Sequence diversity in the large subunit of RNA polymerase I contributes to mefenoxam insensitivity in *Phytophthora infestans*. Mol Plan Path 15:664-676.

Randall, T. A., Dwyer, R. A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., Fong, A. M., Gates, K., Roberts, S., Yatzkan, E., Gaffney, T., Law, M., Testa, A., Torto-Alalibo, T., Zhang, M., Zheng, L., Mueller, E., Windass, J., Binder, A., Birch, P. R., Gisi, U., Govers, F., Gow, N. A., Mauch, F., van West, P., Waugh, M. E., Yu, J., Boller, T., Kamoun, S., Lam, S. T., and Judelson, H. S. 2005. Large-scale gene discovery in the oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. Mol Plan Micro Interac 18:229-243.

Ritch, D. L., and Daggett, S. S. 1995. Nuclear DNA content and chromosome-number in German isolates of *Phytophthora infestans*. Mycologia 87:579-581.

Saville, A. C., Martin, M. D., and Ristaino, J. B. 2016. Historic late blight outbreaks caused by a widespread dominant lineage of *Phytophthora infestans* (mont.) de bary. PLoS One 11:e0168381.

Saville, A. G., K.; Grünwald, N. J.; Myers, K.; Fry, W. E.; Ristaino, J. B. 2014. Fungicide sensitivity of US genotypes of *Phytophthora infestans* (Mont.) de bary to six oomycete-targeted compounds. Plant Dis 99:659-666.

Seidl Johnson, A. C., Frost, K. E., Rouse, D. I., and Gevens, A. J. 2015. Effect of temperature on growth and sporulation of US-22, US-23, and US-24 clonal lineages of *Phytophthora infestans* and implications for late blight epidemiology. Phytopathology 105:449-459.

Sharma, R., Xia, X., Cano, L. M., Evangelisti, E., Kemen, E., Judelson, H., Oome, S., Sambles, C., van den Hoogen, D. J., Kitner, M., Klein, J., Meijer, H. J. G., Spring, O., Win, K., Zipper, R., Bode, H. B., Govers, F., Kamoun, S., Schornack, S., Studholme, D. J., Van den Ackerveken, G., and Thines, M. 2015. Genome analyses of the sunflower pathogen *Plasmopara halstedii* provide insights into effector evolution in downy mildews and *Phytophthora*. BMC Geno 16:741.

Shattock, R. C. 1988. Studies on the Inheritance of resistance to metalaxyl in *Phytophthora infestans*. Plant Pathol 37:4-11.

Small, T. 1938. The relation between plotato blight and tomato blight. Annals of App Bio 25:271-276.

Smoot, J. J., Gough, F. J., Lamey, H. A., Eichenmuller, J. J., and Gallegly, M. E. 1958. Production and germination of oospores of *Phytophthora infestans*. Phytopathology 48:165-171.

Therrien, C. C. R., D. L.; Davidse, L. C.; Jespers, A. B. K.; Speilman, L. J. 1989. Nuclear DNA content, mating type and metalaxyl sensitivity of eighty-three isolates of *Phytophthora infestans* from The Netherlands. Mycol Res 92:140-146.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 Genome sequencing of *Pseudoperonospora cubensis* reveals

effector proteins with a QXLR translocation motif. Mol Plan Micro Interac 24:543-553.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M. K., McDonald, W. H., Medina, M., Meijer, H. J., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261-1266.

Urech, P. A., Schwinn, F., and Staub, T. 1977. CGA 48988, a novel fungicide for the control of late blight, downy mildews and related soil-borne diseases. Proc Br Crop Prot Conf Pests & Dis 2:623–632.

van der Lee, T., Testa, A., Robold, A., van 't Klooster, J., and Govers, F. 2004. High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics 167:1643-1661.

van West, P., Kamoun, S., van ’t Klooster, J. W., and Govers, F. 1999. Internuclear gene silencing in *Phytophthora infestans*. Mol Cell 3:339-348.

Wallin, J. R. 1962. Summary of recent progress in predicting late blight epidemics in United States and Canada. Am Potato J 39:306-312.

Wang, J., Fernandez-Pavia, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodriguez-Alvarado, G., Grunwald, N. J., and Goss, E. M. 2017. High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. Mol Ecol 26:1091-1107.

Weille, G. A. d. 1964. Forecasting crop infection by the potato blight fungus; a fundamental approach to the ecology of a parasite - host relationship. K. Ned. Met. Inst. 82:1-144.

Whisson, S. C., Avrova, A. O., Van West, P., and Jones, J. T. 2005. A method for double-stranded RNA-mediated transient gene silencing in *Phytophthora infestans*. Mol Plan Path 6:153-163.

Yin, L., An, Y., Qu, J., Li, X., Zhang, Y., Dry, I., Wu, H., and Lu, J. 2017. Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism. Sci Rep 7:46553.

Yoshida, K., Schuenemann, V. J., Cano, L. M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F. N., Kamoun, S., Krause, J., Thines, M., Weigel, D., and Burbano, H. A. 2013. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. Elife 2:e00731.

Zhang, X., Scheuring, C., Tripathy, S., Xu, Z., Wu, C., Ko, A., Tian, S. K., Arredondo, F., Lee, M.-K., Santos, F. A., Jiang, R. H. Y., Zhang, H.-B., and Tyler, B. M. 2006. An integrated BAC and genome sequence physical map of *Phytophthora sojae*. Mol Plan Micro Interac 19:1302-1310.

Zhu, W., Yang, L. N., Wu, E. J., Qin, C. F., Shang, L. P., Wang, Z. H., and Zhan, J. 2015. Limited sexual reproduction and quick turnover in the population genetic structure of *Phytophthora infestans* in Fujian, China. Sci Rep 5:10094.

Chapter I

Metalaxyl resistance in Phytophthora infestans: assessing role of RPA190 gene and diversity within clonal lineages

Abstract

Prior work has shown that the inheritance of resistance to metalaxyl, an oomycete-specific fungicide, is complex and may involve multiple genes. Recent research indicated that a single nucleotide polymorphism (SNP) in the gene encoding RPA190, the largest subunit of RNA polymerase I, confers resistance to metalaxyl (or mefenoxam) in some isolates of the potato late blight pathogen *Phytophthora infestans*. Using both DNA sequencing and high resolution melt assays for distinguishing RPA190 alleles, we show here that the SNP is absent from certain resistant isolates of P. *infestans* from North America, Europe, and Mexico. The SNP is present in some members of the US-23 and US-24 clonal lineages, but these tend to be fairly sensitive to the fungicide based on artificial media and field test data. Diversity in the level of sensitivity, RPA190 genotype, and RPA190 copy number was observed in these lineages but were uncorrelated. Controlled laboratory crosses demonstrated that RPA190 did not cosegregate with metalaxyl resistance from a Mexican and British isolate. We conclude that while metalaxyl may be used to control many contemporary strains

of P. *infestans*, an assay based on RPA190 will not be sufficient to diagnose the sensitivity levels of isolates.

Introduction

The development of fungicide resistance is a challenge in many pathosystems. When the genetic basis of resistance is known, DNA-based diagnostics hold the potential of predicting whether an outbreak is controllable by the fungicide. This strategy has become common for Qo inhibitors, where most cases of resistance are caused by a single point mutation in the cytochrome *b* gene (Dufour et al. 2011; Grasso et al. 2006). The approach becomes more complex when multiple genes are involved. With sterol biosynthesis inhibitors (SBIs), for example, reduced sensitivity is known to result from changes in the structure and/or expression of at least three genes (Omrane et al. 2015; Proffer et al. 2006).

The phenylamide compound metalaxyl was introduced in the 1970s as a systemic fungicide against oomycetes and has been used widely against *Phytophthora, Pythium,* and downy mildews. Strains of these pathogens with resistance to metalaxyl (including formulations with mefenoxam, the active R-isomer) have since emerged in many locations (Davidse et al. 1981; Diriwachter et al. 1987; Taylor et al. 2002). For example, starting in the 1980s in Europe and

in North America in the 1990s, many strains of the potato late blight pathogen *Phytophthora infestans* sampled from outbreaks were resistant to metalaxyl. In the United States, the most prevalent resistant genotypes were assigned to asexual lineages called US-6, 7, 8, and 11 based on allozyme and DNA fingerprints (Goodwin et al. 1998). In 2009, populations of *P. infestans* in many parts of the United States underwent another major change as the prior genotypes were displaced by lineages named US-22, 23, and 24 (Hu et al. 2012). These have been reported to lack strong resistance to metalaxyl based on assays on artificial media. US-23 is reported to be fairly sensitive and US-24 to have intermediate sensitivity, compared to resistant lineages like US-8 and 11 (Hu et al. 2012). Since US-23 now dominates many parts of the country, particularly the East and Midwest, metalaxyl has potential value for controlling late blight.

Genetic analyses of the inheritance of metalaxyl insensitivity in *P. infestans* isolates from North America, Mexico, and Europe showed that resistance in each strain was determined by a major semi-dominant locus (Fabritius et al. 1997; Judelson and Roberts 1999; Lee et al. 1999; Shattock 1988). The data also supported the influence of minor modifying loci, as would be seen for a quantitative trait. Strains thus can exhibit a range of sensitivities depending on zygosity at the major locus and the minor genes. Similar conclusions were drawn from downy mildew (Crute and Harrison 1988).

Biochemical studies in *Phytophthora* found that metalaxyl inhibits the synthesis of ribosomal RNA, suggesting that it may target RNA polymerase I (Davidse et al. 1988). Pursuing that lead, Randall *et al.* (2014) associated resistance with a single nucleotide polymorphism, SNP T1145A, in the gene encoding the large subunit of RNA polymerase I, RPA190; T and A are the sensitive and resistant types, respectively. The T→A change causes a tyr→phe substitution, and transforming an RPA190 allele from a resistant isolate into a sensitive strain was reported to decrease the effect of metalaxyl. Randall *et al.* (2014) identified a few resistant strains that lacked the T1145A SNP, which is consistent with a study that suggested that more than one major locus might determine resistance (Judelson and Roberts 1999).

The apparent identification of a metalaxyl resistance gene and reemergence of sensitive isolates suggest that a DNA-based assay for resistance may be useful for improving the management of late blight. In this paper, we demonstrate that metalaxyl can indeed be effective against the new lineages of *P. infestans* under field conditions. We also describe a high resolution melt assay (HRM; (Vossen et al. 2009)) for scoring the T1145A SNP genotype in RPA190. However, our results indicate that the genotype at that locus is a poor predictor of the sensitivity of isolates, and that RPA190 does not cosegregate with resistance in crosses involving some resistant isolates. Our experience with the HRM assay also indicates that DNA assays for diagnosing

traits in *P. infestans* must address the likelihood that many strains are not simple diploids.

Materials and methods

Strains.

Isolates of *P. infestans* were from collections at the University of California-Riverside, USDA-ARS (Corvallis, OR; a gift of N. Grünwald), or Cornell University (a gift of Bill Fry). Assignments to clonal lineages were based on microsatellite markers, or RG57 and allozyme analysis for strains collected prior to 2011. These assays were performed as described, including the use of eleven primer sets that detect polymorphic microsatellite loci (Danies et al. 2013). Progeny were from an 8811 × E13a cross (A1 × A2) described previously (Fabritius et al. 1997), and new progeny from parents (1306 and 618) that had been crossed in a prior study (Judelson et al. 1995). In the latter, oospore germination rates were 17.8%. Whether oospore-derived strains were parental hybrids or selfs was assessed by scoring SNPs in amplicons obtained by polymerase chain reaction (PCR) using G1 (5'-GCAAATTCTTCCTGAACGTGATCC, 5'-ACACCACATCACCAGTCTT CAG) or G7 (5'-GAGAACCTGGAAGAAGCTTACG, 5'-GAGAACCTGGAAGAAGCTTACG) primer pairs, followed by Sanger sequencing.

G7 was derived from previously described marker *W4* (Judelson and Roberts 1999) by *in silico* chromosome walking; *G7* is about 600kb and 1 cM away from *W4*. About five percent of strains obtained from oospores were deemed to be non-recombinant progeny (parental genotypes or selfs) and discarded.

Metalaxyl sensitivity assays.

Strains were cultured on rye-sucrose media containing 0, 0.5, and 5 µg/ml metalaxyl (technical grade, 93% active ingredient) with 1.5% agar. Sensitivity was scored by comparing radial growth in the presence and absence of metalaxyl on 100-mm plates inoculated near one edge, using at least three biological replicates. Inoculum always came from cultures lacking metalaxyl. For isolates 80029, 8880, 93H3, 9026, 9086, 9400, 1114, 11006, and 180, sensitivity was scored based on the increase in colony radius between 3 and 9 days after inoculation. For other strains, growth rates were determined from the slopes of linear regression lines fit to data recorded every 2-4 days; measurements on plates containing the drug were typically made for up to 14 days. Both types of assays yielded similar results. Resistant, intermediate, and sensitive categories were defined as >90%, 40-90%, and <40% growth at 5.0 µg/ml metalaxyl compared to unamended media, respectively.

Field tests.

Potato seed from *cv.* Yukon Gold and transplants of tomato *cv.* Primo Red were planted in Freeville, New York on July 7, 2012 and August 2, 2012, respectively. A randomized complete block design, with six blocks, was used for the experiment. Plots for each treatment (five pathogen strains, with and without fungicide) consisted of two 6 ft rows with 12 in spacing within rows, and 3 and 5 ft spacing between potato and tomato rows, respectively. Ridomil Gold SL (mefenoxam) was applied at 3.2 oz per acre on August 21 at 10 am. At 8 pm, each plot was sprayed with 20 ml of zoospores, which were prepared by incubating sporangia at 10^3/ml for 2 hr at 4°C. Visual disease severity was determined after one and two weeks (August 27 and September 3, 2012) using a modified version of the assessment key described by James (James 1971).

A general linear mixed model (JMP® Pro, version 11.2.0, SAS Institute Inc., Cary, NC) was used to investigate the effect of fungicide treatment, host, isolate, and their interactions on disease severity. The response was $\log_{10}$ transformed disease severity (%). To enable log transformation of the data, a constant of 0.001 was added to all datapoints. The transformation was conducted to satisfy statistical assumptions for the model. Fungicide treatment, host, isolate, and their full factorial interactions were considered fixed effects. The effect of block in each experiment was treated as a random effect. Each

assessment time point (7 or 14 days after infection) was analyzed separately. Non-inoculated controls did not show symptoms of late blight and were therefore excluded from the analyses. The restricted maximum likelihood method (REML) was used for the mixed model. Least squares means were compared using a Tukey Honestly Significant Difference post-hoc test or test slices (JMP® Pro, Version 11.2.0. SAS Institute Inc., Cary, NC). *P* values from test slices were corrected for multiple comparisons using a Bonferroni correction.

RPA190 genotyping.

Alleles were amplified from genomic DNA using the G1 primers listed above, resulting in a 711 bp band, or G1B primers (5'-GGCGCAAATT CTTCCTGAAC, 5'-TTGCTGCTATCCACCAGAAG). The G1B primers bind internal to the G1 primers, and amplify a 280 bp band. Amplified DNA was purified and sequenced using the Sanger method. In the case of heterozygotes, peak heights in chromatograms were calculated using QSVanalyser (Carr et al. 2009). Similar results were obtained when G1 and G1B amplicons were sequenced, and when forward and reverse sequencing primers were used. HRM assays were performed in a Bio-Rad CFX96 instrument using primers HRMU (5'-ACTTGTCGAAGATTATGACT) and HRML (5'-CCTGGCGTTTGTAGTAG). Reactions were performed in a volume of 20 μl using the Bio-Rad Precision Melt Supermix

kit with 10 ng genomic DNA, 0.2 μM of each primer, and 40 cycles of 10 sec at 95°C, 30 sec at 55°C, and 30 sec at 72°C. Melt curves were generated using Precision Melt Analysis software.

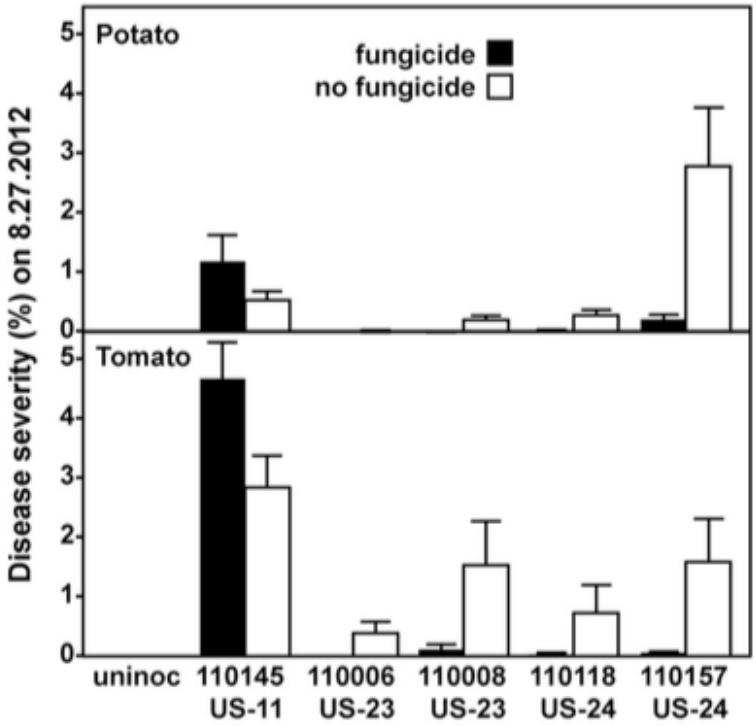
Results

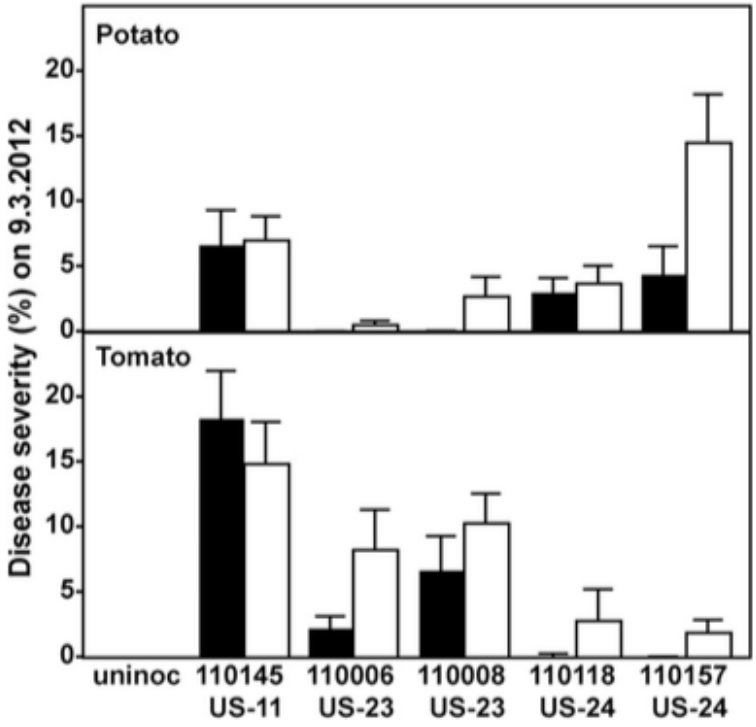
Field control of US-23 and US-24 by metalaxyl.

These lineages are reported to show high to intermediate sensitivities to metalaxyl, respectively, when assayed on artificial media (Danies et al. 2013). Since a premise of our research was that metalaxyl might be used to control these lineages, in concert with a DNA assay to predict isolate sensitivity, it was important to test how metalaxyl affected US-23 and US-24 in the field. Consequently, plots of potato and tomato with or without fungicide treatment were inoculated with two US-23 and two US-24 isolates and disease severity was scored after 7 and 14 days (8.27.2012 and 9.3.2012, respectively). In laboratory tests, these isolates had displayed sensitive and intermediate phenotypes. A metalaxyl-resistant strain from clonal lineage US-11 was used as a control.

In the field experiment (**Figure 1.1**), the effectiveness of metalaxyl depended on the isolate as indicated by a significant two-way interaction between fungicide treatment and isolate ($P<0.0001$ and $P<0.02$ for the 7 and 14-day assessments, respectively). In particular, metalaxyl was effective against

Figure 1.1. Effect of metalaxyl in field assay. Plots of potato and tomato were sprayed with Ridomil Gold SL at a rate of 3.2 oz per acre. Ten hours later, the plants were sprayed with zoospore suspensions from the indicated isolates of Phytophthora infestans. Symptoms were recorded 7 days (8.27.2012; top) and 14 days (9.3.2012; bottom) after inoculation.


Potato
fungicide
no fungicide
Tomato
Disease severity (%) on 8.27.2012
0
1
2
3
4
5
uninoc
110145 US-11
110006 US-23
110008 US-23
110118 US-24
110157 US-24


Potato
Tomato
Disease severity (%) on 9.3.2012
0
5
10
15
20
uninoc
110145 US-11
110006 US-23
110008 US-23
110118 US-24
110157 US-24

US-23 and US-24 ($P<0.05$). One week after inoculation, metalaxyl suppressed disease severity caused by all four US-23 and US-24 strains by >50%, relative to their respective untreated controls. After two weeks, the metalaxyl treatment demonstrated a continued reduction in disease severity for both isolates of US-23 ($P<0.007$) and one isolate of US-24 (US110157; $P<0.0001$) relative to the controls. In contrast, the chemical was ineffective against the US-11 strain, with no significant difference in disease severity between metalaxyl-treated and untreated plants ($P>0.94$). It is interesting to note that on artificial media this US-11 strain grows slightly faster in the presence of metalaxyl, and we also observed a trend of more growth on the fungicide-treated versus untreated plants in this field study.

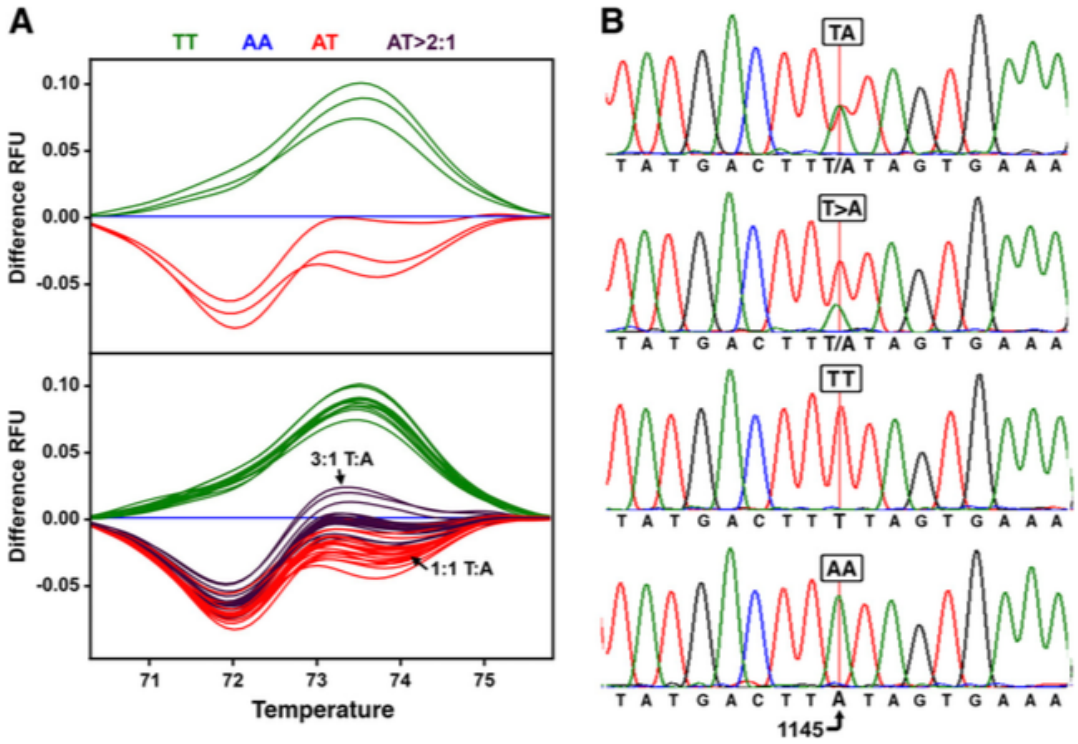
Differences in host preference were also observed for US-23 and US-24. The US-23 isolates caused more symptoms on tomato than potato ($P<0.001$) and US-24 isolates resulted in more symptoms on potato than tomato ($P<0.01$). These differences are consistent with previous reports (Danies et al. 2013). In contrast, US 11 did not differ significantly in disease severity on the two hosts ($P>0.2$). There was a non-significant three-way interaction between fungicide treatment, host, and isolate ($P=0.53$ and $P=0.44$ for the 7 and 14 day assessments, respectively).

Determining genotype by sequencing and HRM analysis.

Since we anticipated that position 1145 of RPA190 might predict metalaxyl sensitivity, two methods were developed to genotype isolates. These were Sanger sequencing of PCR products and high resolution melt (HRM) analysis. Isolates found by sequencing to be A or T homozygotes or heterozygotes were distinguishable by the HRM technique (**Figure 1.2A**, top panel). This employed primers that amplified a 63-nt region that was monomorphic in each of 22 isolates tested, not counting position 1145.

Some heterozygous isolates seemed to be triploid or tetraploid since sequencing reactions yielded A:T ratios besides 1:1 at position 1145. This was evident when peak heights in the sequencing chromatograms were compared (**Figure 1.2B**). For example, while the top chromatogram in the figure shows a 1:1 ratio of A and T, the second chromatogram presents a 1:2 ratio, which is suggestive of triploidy. Similar results were obtained regardless of the direction of the sequencing reaction, with multiple primer sets in the PCR reaction, and when isolates were passed through a zoospore (single-nucleus) purification stage. These genotypes nevertheless could still be distinguished from A or T homozygotes in the HRM assay (**Figure 1.2A**, bottom panel).

Figure 1.2. Methods for genotyping RPA190. **A**) High resolution melt assays using a 63-nt amplicon that spans position 1145. The top panel indicates melt curves for strains scored as TT (green), AA (blue), and AT (red); the latter portrays strains with 1:1 ratios between A and T alleles based on Sanger sequencing. The bottom panel indicates curves from 55 individuals, including heterozygotes (purple) suggested to be other than diploid by DNA sequencing; highlighted are isolates 120137 and 130165, which sequence analysis indicates have genotypes AT (1:1) and ATTT (3:1), respectively. **B**) Sequencing chromatograms of region containing the single nucleotide polymorphisms. These portray cases (top to bottom) in which T and A alleles are present at equal levels, the dosage of T exceeds A, a TT homozygote, and an AA homozygote.


A
TT
AA
AT
AT>2:1
Difference RFU
0.10
0.05
0.00
-0.05
3:1 T:A
1:1 T:A
71
72
73
74
75
Temperature
B
TA
T>A
TT
AA
T A T G A C T T T/A T A G T G A A A
T A T G A C T T T T A G T G A A A
T A T G A C T T A T A G T G A A A
1145

Imperfect association of T1145A SNP with resistance.

Twenty-six strains of *P. infestans* from North America and Europe with diverse levels of resistance were scored for the SNP by sequencing (**Table 1**). These included representatives of the historically prominent lineages from the United States, namely US-1, 6, 7, 8, 11, 22, 23, and 24. Most were also tested using the HRM assay, which yielded the same result as DNA sequencing.

Containing the A allele associated by Randall et al. (2014) with resistance were eight of eleven resistant isolates, four of seven intermediately resistant isolates, and none of six sensitive isolates. The exceptions included three resistant (110145, 88069, and 8880) and three intermediate isolates (618, 1114, and 11017) that were T homozygotes. Isolate 8880, for example, belongs to the US-6 lineage which is one of the most metalaxyl-resistant North American genotypes (Goodwin et al. 1996), despite being homozygous for T at position 1145. Two resistant US-11 isolates varied in genotype, with isolate 110145 being homozygous for T and isolate 9400 an AT heterozygote. It should be noted that when isolates scored as resistant were found to lack an A allele, or when isolates recorded as sensitive contained an A allele, the genotypic and phenotypic analyses were repeated.

One explanation for these findings is that resistance in some strains is determined by genes other than RPA190. It is also possible that SNPs at other

TABLE 1.1. RNA190 genotype in 22 isolates of *P. infestans.*

Isolate	Phenotype[a]	1145 alleles[b]	A:T ratio[b]	Origin and lineage[c]
80029	R	AA	2:0	Netherlands, 1980
8811	R	AT	1:1	United Kingdom, 1988
88069	R	TT	0:2	Netherlands, 1988
8880	R	TT	0:2	Canada, 1994 (US-6)
93H3	R	AT	1:1	USA, 1993 (US-7)
9086	R	AT	1:1	Canada, 1994 (US-8)
9026	R	AT	1:1	Canada, 1994 (US-8)
9400	R	AT	1:2	Canada, 1995 (US-11)
110145	R	TT	0:2	USA, 2011 (US-11)
11007	R	AT	1:1	USA, 2011 (US-8)
1114	I	TT	0:2	Netherlands, 1985
618	I	TT	0:2	Mexico, 1987
11017	I	TT	0:2	USA, 2011 (US-23)
110006	I	AT	1:2	USA, 2011 (US-23)
110008	I	AT	1:3	USA, 2011 (US-23)
110118	I	AT	1:1	USA, 2011 (US-24)
110157	I	AT	1:2	USA, 2011 (US-24)
11019	I	AT	1:2	USA, 2011 (US-24)
180	S	TT	0:2	USA (US-1)
510	S	TT	0:2	Mexico, 1983
550	S	TT	0:2	Mexico, 1983
1306	S	TT	0:2	USA, 1982
R0	S	TT	0:2	USA, pre-1980 (US-1)
E13a	S	TT	0:2	Egypt, 1984
11016	S	TT	0:2	USA, 2011 (US-22)
11006	S	TT	0:2	USA, 2011 (US-1)

[a]Resistant (R), intermediate (I), and sensitive (S) are defined as >90%, 40-90%, and <40% growth at 5.0 µg/ml metalaxyl compared to unamended media.

[b]Based on Sanger sequencing of RPA190 gene. Allele ratios determined using QSV Analyzer are shown to the nearest integer. When a homozygous state was inferred, the strain is scored as a diploid (0:2 or 2:0), although some could be of higher ploidy.

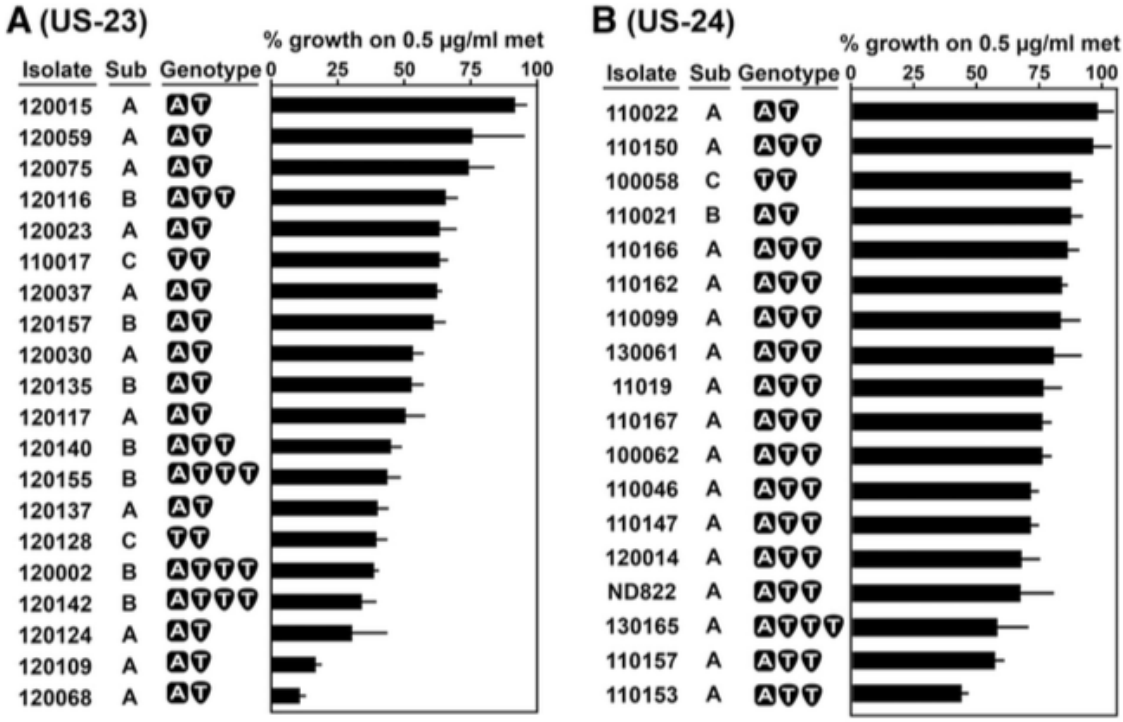
[c]Genotypes matching common "US" clonal lineages are indicated in parentheses. The date of isolation of some strains are unknown.

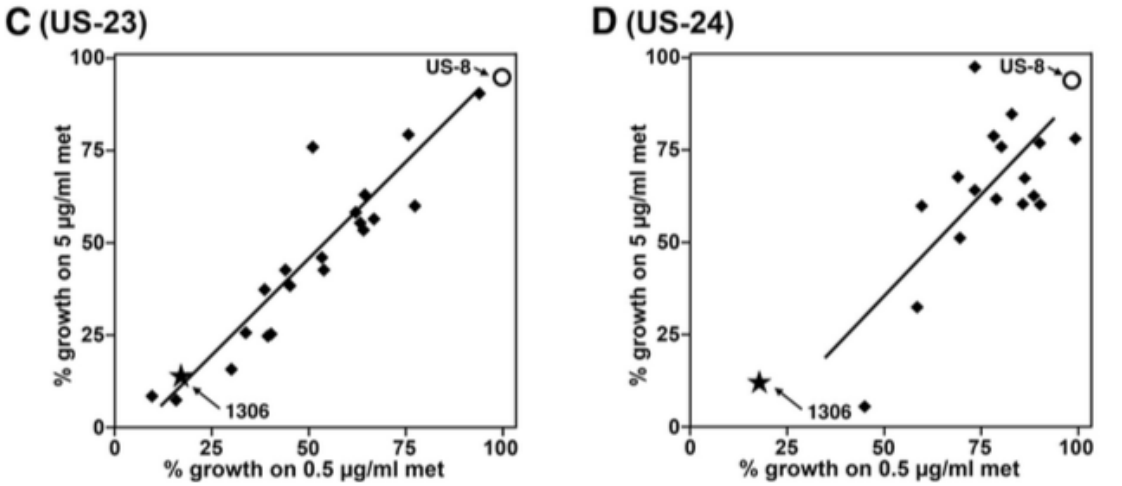
locations in RPA190 may cause resistance. We sequenced the entire 5.4 kb gene from several of the strains in **Table 1** in an attempt to identify such SNPs (Supplemental **Table 1**), but failed to detect other SNPs associated with resistance.

Variation in US-23 and US-24 is not correlated with RPA190 genotype.

Some diversity in metalaxyl sensitivity in these lineages is reported (Danies et al. 2013), and we sought to test if this could be explained by the genotype at position 1145. Twenty isolates assigned to US-23 and eighteen to US-24 by microsatellite analysis (Danies et al. 2013) were thus analyzed. Their relative growth rates in 0.5 µg/ml metalaxyl compared to unamended media ranged from 11% to 95% for US-23 and 45 to 100% for US-24 (**Figure 1.3A, B**). By comparison, strain 1306, which predates the description of metalaxyl resistance in North America, showed <20% growth. The relative growth rates of the most highly resistant strains on both 0.5 and 5.0 µg/ml metalaxyl were close to that of a US-8 isolate (**Figure 1.3C, D**). Note that the US-23 and US-24 strains used in this experiment were selected for diversity, therefore the sensitivity distribution shown in the figure does not reflect frequencies observed in random isolates.

Figure 1.3. Characteristics of US-23 and US-24 isolates. **A**) US-23. Indicated are assignments to genotype subgroups A, B, or C as defined in Table 2 (sub), genotypes at position 1145 with the number of A and T symbols reflecting ratios from QSZ Analyzer rounded to the nearest integer, and relative growth on metalaxyl at 0.5 µg/ml and 5.0 µg/ml compared with unamended media. **B**) Same as panel A, but for US-24. **C**) Relative growth of US-23 strains on metalaxyl at 0.5 µg/ml and 5.0 µg/ml compared with unamended media. Included for comparison are sensitive isolates 1306 (star) and resistant US-8 isolate 110007 (circle). **D**) Same as panel C, but for US-24


A (US-23)
% growth on 0.5 µg/ml met
Isolate Sub Genotype
120015 A AT
120059 A AT
120075 A AT
120116 B ATT
120023 A AT
110017 C TT
120037 A AT
120157 B AT
120030 A AT
120135 B AT
120117 A AT
120140 B ATT
120155 B ATTT
120137 A AT
120128 C TT
120002 B ATTT
120142 B ATTT
120124 A AT
120109 A AT
120068 A AT
0 25 50 75 100
B (US-24)
% growth on 0.5 µg/ml met
Isolate Sub Genotype
110022 A AT
110150 A ATT
100058 C TT
110021 B AT
110166 A ATT
110162 A ATT
110099 A ATT
130061 A ATT
11019 A ATT
110167 A ATT
100062 A ATT
110046 A ATT
110147 A ATT
120014 A ATT
ND822 A ATT
130165 A ATTT
110157 A ATT
110153 A ATT
0 25 50 75 100


C (US-23)
% growth on 5 µg/ml met
% growth on 0.5 µg/ml met
US-8
1306
0 25 50 75 100
D (US-24)
% growth on 5 µg/ml met
% growth on 0.5 µg/ml met
US-8
1306
0 25 50 75 100

Genotypes obtained by DNA sequencing indicated that both the US-23 and US-24 lineages contained individuals that were T homozygotes or AT heterozygotes at position 1145. The A:T ratios of heterozygotes included 1:1, 1:2, and 1:3, suggesting that many were triploid or tetraploid. This did not appear to be due to heterokaryosis, since the same results were obtained when several isolates were retested after single-zoospore purification; zoospores are typically mononucleate.

Regardless of the basis of the nonstandard allele ratios, the dosage of the A allele was not strongly linked to the level of metalaxyl sensitivity. For example, AT genotypes resided at both ends of the sensitivity spectrum in US-23, and some of the more resistant US-23 and US-24 strains were TT. We considered the possibility that different results might have been obtained if testing was done at a higher concentration of metalaxyl, but this was not the case (**Figure 1.3C, D**). A fairly good correlation was observed between relative growth at 0.5 and 5 µg/ml metalaxyl, with the correlation coefficient R equaling 0.89 and 0.71 for US-23 and US-24, respectively.

The sequence analysis also detected three distinct genotypes in US-23 and three in US-24, when SNPs throughout the whole 711-nt amplicon used to score position 1145 were considered. Variation was seen at five positions including 1145, representing two and three SNPs causing synonymous and nonsynonymous substitutions, respectively (**Table 1.2**). In US-23, 11 strains

TABLE 1.2. Subclsses within lineages based on RPA190 SNPs

		Position in gene					
Lineage	Subclass	1145	1338	1371	1670	1693	# of strains
US-23	23A	AT	CC	TT	GG	CT	11
	23B	AT	CT	TT	GT	CT	7
	23C	TT	CT	TT	GT	CC	2
US-24	24A	AT	CT	CT	GT	CC	16
	24B	AT	CC	CT	GG	CC	1
	24C	TT	CT	TT	GT	CC	1

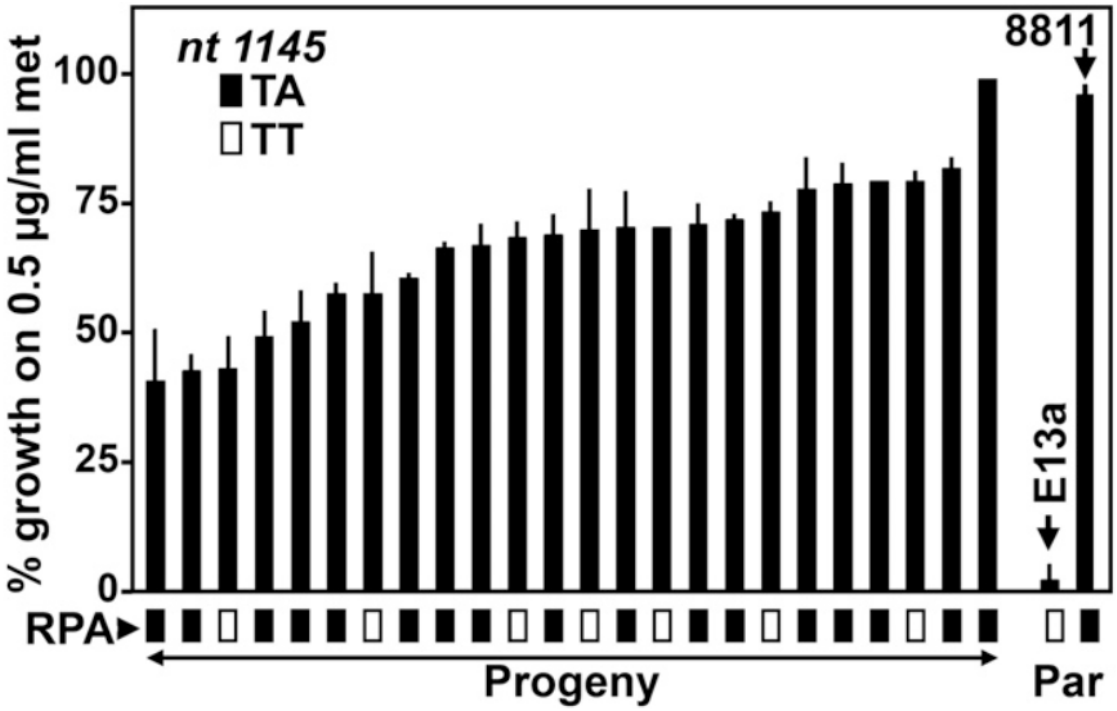
were represented by genotype 23A, followed by 7 in 23B and two in 23C. US-24 was less diverse, with most strains belonging to group 24A. The differences were consistent with gene conversion events rather than spontaneous mutations. For example, at positions 1338 and 1670 in US-23, the genotypes were either CC or CT, or GT or GG, respectively. There was not an obvious correlation between sensitivity level and genotype or ploidy, and the most and least sensitive strains within US-23 both belonged to group 23A.

RPA190 SNPs fail to cosegregate with resistance.

The failure of the T1145A SNP to associate strongly with sensitivity in the isolates in **Table 1.1**, and in the US-23 and 24 strains in **Figure 1.3**, does not exclude the prospect that other polymorphisms in the RPA190 open reading frame play roles in resistance. The gene's regulatory sequences may also contribute to resistance by affecting RPA190 expression. To test these possibilities, we assessed whether RPA190 alleles cosegregated with resistance in controlled laboratory crosses.

First, we examined a cross between isolates 8811 (highly resistant, AT at position 1145) and E13a (sensitive, TT). The phenotypes of 24 progeny from this cross ranged from intermediate to high resistance (**Figure 1.4**). The sensitivity levels of TT and AT progeny were not significantly different (P=0.91 by unpaired

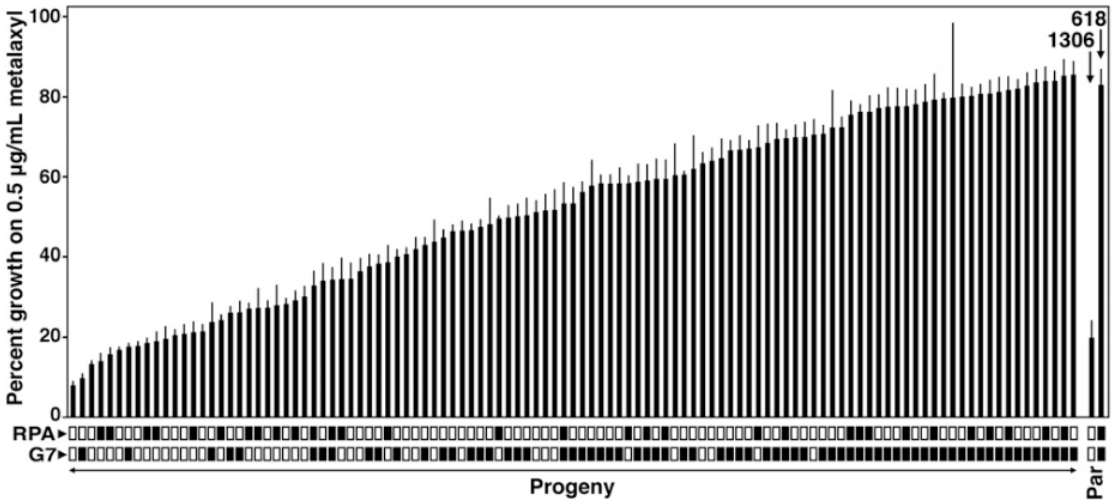
Figure 1.4. Metalaxyl resistance and 1145 genotype of progeny from 8811 × E13a cross. Shown is the relative growth of the strains on metalaxyl at 0.5 µg/ml compared with unamended media, and the genotype of position 1145 of RPA190. Closed and open boxes beneath the graph represent TA and TT genotypes, respectively.


nt 1145
TA
TT
8811
E13a
% growth on 0.5 µg/ml met
100
75
50
25
0
RPA
Progeny
Par

t-test). The most likely interpretation is that insensitivity is caused primarily by a locus unlinked to RPA190, which is likely homozygous in 8811.

A similar conclusion was drawn from a cross between intermediately resistant isolate 618 and sensitive strain 1306. As shown in **Figure 1.5**, the offspring ranged from low to high sensitivity. Both parents were homozygous for T at position 1145, but other polymorphisms were present which enabled us to assess whether RPA190 co-segregated with metalaxyl resistance. Their analysis indicated that the sensitivity of progeny inheriting the two alternate alleles of RPA190 from 618 were not significantly different (P=0.92 by unpaired t-test). An equivalent result was obtained by treating the 25 most resistant and 25 sensitive progeny as separate classes, in which case RPA190 and resistance were unlinked with a recombination fraction of 0.50. The most likely interpretation is that insensitivity is determined mainly by a locus unlinked to RPA190, which is likely heterozygous in 618. The locus determining insensitivity is however genetically linked to marker G7, which was obtained by in silico chromosome walking from marker W4, which was identified previously by screening for RAPD markers associated with resistance (Judelson et al. 1999). Treating the 25 most resistant and 25 sensitive progeny as separate classes, G7 had a recombination frequency of 0.10 (LOD=8) with the resistance determinant.

Figure 1.5. Metalaxyl resistance and genotypes of 618 × 1306 cross. Shown is the relative growth of the strains on metalaxyl at 0.5 µg/ml compared with unamended media, and the genotype of position 1145 of RPA190 and an unlinked marker, G7. Closed and open boxes beneath the graph represent TA and TT genotypes for RPA190, and CT and TT for locus G7, respectively.


Percent growth on 0.5 µg/mL metalaxyl
100
80
60
40
20
0
618
1306
RPA
G7
Progeny
Par

Discussion

The use of site-specific fungicides has expanded over past decades, since many have systemic and curative activity. Their effectiveness relative to traditional multisite inhibitors has been countered by the increased likelihood that resistance will arise. The study by Randall et al. (Randall et al. 2014) was a landmark since it revealed a genetic basis of resistance in *P. infestans* to metalaxyl, which has been an important tool against oomycetes. Our work however failed to connect that genetic change with the sensitivity level of many isolates, including lineages now prevalent in the United States. The gene encoding the RNA polymerase subunit implicated by Randall et al. also did not co-segregate with resistance in our crosses involving insensitive isolates from Mexico and the United Kingdom. While we successfully developed an assay for rapid RPA190 genotyping using the HRM technique, such tests will not be useful for informing management practices until more is learned about the basis of resistance.

That multiple genes may determine metalaxyl resistance is not surprising, since it behaves genetically as a complex trait. This was illustrated by our 618 × 1306 cross, where offspring displayed a continuous distribution of sensitivities. Genes that potentially affect resistance include those that encode the structural target of metalaxyl, inactivating enzymes, and efflux proteins. The latter include

ABC transporters, which also contribute to resistance against chemistries used against true fungi (Abou Ammar et al. 2013; Omrane et al. 2015) and were observed to be upregulated in strains of *P. infestans* adapted to growth on metalaxyl (Childers et al. 2014). Since RNA polymerase I is a multiprotein complex, alterations in more than one subunit might also influence sensitivity. The T1145A SNP changes an amino acid in the RPA190 clamp domain, through which rRNA exits the polymerase. Based on the structure of yeast RNA polymerase I holoenzyme, the protrusion domain of RPA135 resides near the clamp (Engel et al. 2013; Fernandez-Tornero et al. 2013). Consequently, changes in both RPA135 and RPA190 might influence the response to metalaxyl.

While our sequence analysis of RPA190 in US-23 and US-24 was not informative for diagnosing metalaxyl resistance, it revealed diversity in the lineages. These lineages, which are defined by an eleven-locus microsatellite genotype, are believed to reproduce asexually although changes can occur through spontaneous mutation and gene conversion. Nevertheless, we were surprised to detect three distinct genotypes based on just a short stretch of a single gene, in addition to potential variation in ploidy. The US-23 and US-24 lineages were first detected in 2009, and the pace at which their sublineages arose can be estimated knowing that the isolates studied here were collected between 2010 and 2013. There are also precedents for variation in other lineages of *P. infestans* based on virulence and molecular markers (Abu-El

Samen et al. 2003, 2003; Goodwin et al. 1995; Harbaoui et al. 2014). Variation in the level of metalaxyl sensitivity has also been reported in other lineages, such as US-8 and US-11 (Daayf et al. 2002).

Our detection of nonstandard allele ratios within US-23 and US-24 is consistent with prior information of the diversity of nuclear states in *P. infestans*. Isolates are reported to include 2n, 3n, 4n, and aneuploid types (Catal et al. 2010; Ritch et al. 1995; Tooley et al. 1987). US-24 isolates in particular were observed to contain a three-allele (100/100/111) allozyme genotype for glucose-6-phosphate isomerase, which suggests that lineage is not diploid (Hu et al. 2012). Ploidy levels are also known to vary between nuclei within a single *P. infestans* culture (Catal et al. 2010). After passing through a spore stage, this would give rise to a population of individuals with a range of nuclear contents, and this may explain the variation within US-23 and US-24.

That ploidy varies within *P. infestans* must be factored in when designing DNA assays for predicting traits. There is growing interest, both with *P. infestans* and other phytopathogens, in using DNA assays to diagnose traits such as fungicide resistance, virulence (race), and mating type (Carpezat et al. 2014; Derevnina et al. 2015; Dufour et al. 2011; Short et al. 2014). Our experience with *P. infestans* shows that consideration must be given to the phenotype that results from all possible allele dosages, since strains having 3:1, 2:1, and 1:1 ratios at a given locus may have distinct phenotypes. Quantitative rather than qualitative

DNA assays may therefore be required. Nucleic acid sequencing should be sufficient, but has a slower turn-around time than other technologies. Our HRM assay is faster than sequencing, and could distinguish AT heterozygotes with different ratios from AA and TT homozygotes within RPA190, but this approach might not succeed with other loci. The complication of diversity in ploidy is not unique to *P. infestans*. It has been described in other species of *Phytophthora* (Bertier et al. 2013), in ascomycetes and basidiomycetes (Albertin et al. 2012), in nematodes (Triantaphyllou 1991), and is suspected in *Pythium* (Spies et al. 2011).

References

Abou Ammar, G., Tryono, R., Doll, K., Karlovsky, P., Deising, H. B., and Wirsel, S. G. R. 2013. Identification of ABC transporter genes of *Fusarium graminearum* with roles in azole tolerance and/or virulence. PLoS One 8:e79042.

Abu-El Samen, F. M., Secor, G. A., and Gudmestad, N. C. 2003a. Variability in virulence among asexual progenies of *Phytophthora infestans*. Phytopathology 93:293-304.

Abu-El Samen, F. M., Secor, G. A., and Gudmestad, N. C. 2003b. Genetic variation among asexual progeny of *Phytophthora infestans* detected with RAPD and AFLP markers. Plant Pathol 52:314-325.

Albertin, W., and Marullo, P. 2012. Polyploidy in fungi: Evolution after whole-genome duplication. Proc R Soc B 279:2497-2509.

Bertier, L., Leus, L., D'hondt, L., de Cock, A. W. A. M., and Hofte, M. 2013. Host adaptation and speciation through hybridization and polyploidy in *Phytophthora*. PLoS One 8:e85385.

Carpezat, J., Bothorel, S., Daverdin, G., Balesdent, M. H., and Leflon, M. 2014. Use of high resolution melting analysis to genotype the avirulence avrLM7 gene of *Leptosphaeria maculans*, a fungal pathogen of *Brassica napus*. Ann Appl Biol 164:430-440.

Carr, I. M., Robinson, J. I., Dimitriou, R., Markham, A. F., Morgan, A. W., and Bonthron, D. T. 2009. Inferring relative proportions of DNA variants from sequencing electropherograms. Bioinformatics 25:3244-3250.

Catal, M., King, L., Tumbalam, P., Wiriyajitsomboon, P., Kirk, W. W., and Adams, G. C. 2010. Heterokaryotic nuclear conditions and a heterogeneous nuclear population are observed by flow cytometry in Phytophthora *infestans*. Cytometry A 77A:769-775.

Childers, R., Danies, G., Myers, K. L., Fei, Z., Small, I. M., and Fry, W. 2014. Acquired resistance to mefenoxam in sensitive isolates of *Phytophthora infestans*. Phytopathology 105:342-349.

Crute, I. R., and Harrison, J. M. 1988. Studies on the inheritance of resistance to metalaxyl in *Bremia lactucae* and on the stability and fitness of field isolates. Plant Pathol 37:231-250.

Daayf, F., and Platt, H. W. 2002. Variability in responses of US-8 and US-11 genotypes of potato and tomato isolates of *Phytophthora infestans* to commercial fungicides in vitro. Am J Potato Res 79:433-441.

Danies, G., Small, I. M., Myers, K., Childers, R., and Fry, W. E. 2013. Phenotypic characterization of recent clonal lineages of *Phytophthora infestans* in the United States. Plant Dis 97:873-881.

Davidse, L. C., Gerritsma, O. C. M., Ideler, J., Pie, K., and Velthuis, G. C. M. 1988. Antifungal modes of action of metalaxyl, cyprofuram, benalaxyl and oxadixyl in phenylamide-sensitive and phenylamide-resistant strains of *Phytophthora megasperma* f. sp. *medicaginis* and *Phytophthora infestans.* Crop Prot 7:347-355.

Davidse, L. C., Looijen, D., Turkensteen, L. J., and Vanderwal, D. 1981. Occurrence of metalaxyl-resistant strains of *Phytophthora infestans* in Dutch potato fields. Neth. J. Plant Pathol. 87:65-68.

Derevnina, L., and Michelmore, R. W. 2015. Wheat rusts never sleep but neither do sequencers: Will pathogenomics transform the way plant diseases are managed? Genome Biol. 16:44.

Diriwachter, G., Sozzi, D., Ney, C., and Staub, T. 1987. Cross-resistance in Phytophthora infestans and *Plasmopara viticola* against different phenylamides and unrelated fungicides. Crop Prot 6:250-255.

Dufour, M. C., Fontaine, S., Montarry, J., and Corio-Costet, M. F. 2011. Assessment of fungicide resistance and pathogen diversity in Erys*iphe necator* using quantitative real-time PCR assays. Pest Manag Sci 67:60-69.

Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D., and Cramer, P. 2013. RNA polymerase I structure and transcription regulation. Nature 502:650-655.

Fabritius, A.-L., Shattock, R. C., and Judelson, H. S. 1997. Genetic analysis of metalaxyl insensitivity loci in Phytophthora infestans using linked DNA markers. Phytopathology 87:1034-1040.

Fernandez-Tornero, C., Moreno-Morcillo, M., Rashid, U. J., Taylor, N. M. I., Ruiz, F. M., Gruene, T., Legrand, P., Steuerwald, U., and Muller, C. W. 2013. Crystal structure of the 14-subunit RNA polymerase I. Nature 502: 644-649.

Goodwin, S. B., Smart, C. D., Sandrock, R. W., Deahl, K. L., Punja, Z. K., and Fry, W. E. 1998. Genetic charge within populations of *Phytophthora infestans* in the United States and Canada during 1994 to 1996: Role of migration and recombination. Phytopathology 88:939-949.

Goodwin, S. B., Sujkowski, L. S., and Fry, W. E. 1995. Rapid evolution of pathogenicity within clonal lineages of the potato late blight disease fungus. Phytopathology 85:669-676.

Goodwin, S. B., Sujkowski, L. S., and Fry, W. E. 1996. Widespread distribution and probable origin of resistance to metalaxyl in clonal genotypes of *Phytophthora infestans* in the United States and western Canada. Phytopathology 86:793-800.

Grasso, V., Palermo, S., Sierotzki, H., Garibaldi, A., and Gisi, U. 2006. Cytochrome b gene structure and consequences tor resistance to Qo inhibitor fungicides in plant pathogens. Pest Man Sci 62:465-472.

Harbaoui, K., Hamada, W., Li, Y., Vleeshouwers, V. G. A. A., and van der Lee, T. 2014. Increased difficulties to control late blight in Tunisia are caused by a genetically diverse *Phytophthora infestans* population next to the clonal lineage NA-01. Plant Dis 98:898-908.

Hu, C. H., Perez, F. G., Donahoo, R., McLeod, A., Myers, K., Ivors, K., Secor, G., Roberts, P. D., Deahl, K. L., Fry, W. E., and Ristaino, J. B. 2012. Recent genotypes of *Phytophthora infestans* in the eastern United States reveal clonal populations and reappearance of mefenoxam sensitivity. Plant Dis 96:1323-1330.

James, W. C. 1971. An illustrated series of assessment keys for plant diseases, their preparation and usage. Can Plant Dis Surv 51:39-65.

Judelson, H. S., and Roberts, S. 1999. Multiple loci determining insensitivity to phenylamide fungicides in *Phytophthora infestans*. Phytopathology 89: 754-760.

Judelson, H. S., Spielman, L. J., and Shattock, R. C. 1995. Genetic-mapping and non-Mendelian segregation of mating-type loci in the oomycete, *Phytophthora infestans*. Genetics 141:503-512.

Lee, T. Y., Mizubuti, E., and Fry, W. E. 1999. Genetics of metalaxyl resistance in *Phytophthora infestans*. Fungal Genet Biol 26:118-130.

Omrane, S., Sghyer, H., Audeon, C., Lanen, C., Duplaix, C., Walker, A. S., and Fillinger, S. 2015. Fungicide efflux and the MgMFS1 transporter contribute to the multidrug resistance phenotype in *Zymoseptoria tritici* field isolates. Env Micro 17:2805-2823.

Proffer, T. J., Berardi, R., Ma, Z. H., Nugent, J. E., Ehret, G. R., McManus, P. S., Jones, A. L., and Sundin, G. W. 2006. Occurrence, distribution, and polymerase chain reaction-based detection of resistance to sterol demethylation inhibitor fungicides in populations of *Blumeriella jaapii* in Michigan. Phytopathology 96:709-717.

Randall, E., Young, V., Sierotzki, H., Scalliet, G., Birch, P. R. J., Cooke, D. E. L., Csukai, M., and Whisson, S. C. 2014. Sequence diversity in the large subunit of RNA polymerase I contributes to mefenoxam insensitivity in *Phytophthora infestans*. Mol Plan Pathol 15:664-676.

Ritch, D. L., and Daggett, S. S. 1995. Nuclear-DNA content and chromosome-number in German isolates of *Phytophthora infestans*. Mycologia 87:579-581.

Shattock, R. C. 1988. Studies on the inheritance of resistance to metalaxyl in *Phytophthora infestans*. Plant Pathol 37:4-11.

Short, D. P. G., Gurung, S., Maruthachalam, K., Atallah, Z. K., and Subbarao, K. V. 2014. *Verticillium dahliae* race 2-specific PCR reveals a high frequency of race 2 strains in commercial spinach seed lots and delineates race structure. Phytopathology 104:779-785.

Spies, C. F. J., Mazzola, M., Botha, W. J., Langenhoven, S. D., Mostert, L., and Mcleod, A. 2011. Molecular analyses of *Pythium irregulare* isolates from grapevines in South Africa suggest a single variable species. Fungal Biol 115:1210-1224.

Taylor, R. J., Salas, B., Secor, G. A., Rivera, V., and Gudmestad, N. C. 2002. Sensitivity of North American isolates of *Phytophthora erythroseptica* and *Pythium ultimum* to mefenoxam (metalaxyl). Plant Dis 86:797-802.

Tooley, P. W., and Therrien, C. D. 1987. Cytophotometric determination of the nuclear DNA content of 23 Mexican and 18 non-Mexican isolates of *Phytophthora infestans*. Exp Mycol 11:19-26.

Triantaphyllou, A. C. 1991. Further studies on the role of polyploidy in the evolution of *Meloidogyne*. J Nematol 23:249-253.

Vossen, R. H., Aten, E., Roos, A., and den Dunnen, J. T. 2009. High-resolution melting analysis (HRMA): More than just sequence variant screening. Hum Mutat 30:860-866.

Chapter II

Assembly of a new *Phytophthora infestans* reference genome

Abstract

Perhaps the most important tool for research in the genomics age of life sciences is the reference sequence. This glorified text file representing an organism's entire complement of DNA sequence enables investigations of the evolutionary relations within and between species, observing genome plasticity and structural variation, and linking gene sequences to phenotypic traits. The oomycete plant pathogen *Phytophthora infestans* first had a reference genome assembled by the Broad institute in 2009, yet the difficult nature of the genome resulted in a highly fragmented assembly. To facilitate the generation of a more complete reference genome, we incorporated several technologies including Pacific Biosciences long-read sequencing, Bionano optical mapping, a DoveTail long-distance mate-pair library, and a new sequence-stitching pipeline to generate an input assembly of similar contiguity to the previous Broad assembly. Genetic maps were produced through Illumina sequencing of progeny originating from two crosses, by calling SNPs/markers from alignments made to the input assembly followed by linkage analysis. Finally, the genetic coordinates of

markers in each map allowed the contigs of the input genome to be organized into chromosome-sized scaffolds. These pseudochromosomes range in size from 22.9-10.1 Mb and anchor 218.8 Mb of sequence out of a total of 247 Mb (88.5%) in the input assembly. This represents the first successful attempt at generating a chromosome-scale reference assembly for *P. infestans,* and enables far more powerful investigations into the genome of this plant pathogen.

Introduction

The oomycetes include a diverse assortment of both plant and animal pathogens that are responsible for devastating diseases (Jiang and Tyler 2012; reviewed in Kamoun et al. 2015). Though oomycetes resemble true fungi in that they often adopt a filamentous and pathogenic lifestyle, they are phylogenetically unrelated to true fungi and are more related to brown algae and diatoms. While many genera of oomycetes cause important diseases of diverse crops, the genus *Phytophthora* represents a lineage of hemibiotrophic pathogens causing especially devastating diseases. Some examples of current challenges imposed by *Phytophthora* include the epidemic of Sudden Oak Death in forests across the western United States from *Phytophthora ramorum,* as well as the explosive spread of *Phytophthora cinnamomi* through the Australian environment, threatening more than 3,000 native species (Australia 2014). However,

Phytophthora infestans, the pathogen responsible for the infamous great Irish Potato famine, is likely the most destructive, causing billions of dollars in crop damage and societal losses globally each year (Haverkort et al. 2008).

In as little as two decades, the development of the second generation of sequencing technologies, namely Illumina sequencing by synthesis, sent the cost per base of DNA sequencing plummeting (reviewed in Goodwin et al. 2016). The resulting explosion of data has enabled the examination of genome variation, structure, and plasticity on a level not previously possible. One additional effect of new technology includes a greatly altered process of generating reference genome assemblies for species lying outside of the canonical model organisms such as humans, yeast, or *Arabidopsis thaliana*. Several genome assemblies exist for oomycetes, including the hemibiotrophic species *Phytophthora sojae* and *P. ramorum* (Tyler et al. 2006), *P. infestans* (Haas et al. 2009), *P. capsici* (Lamour et al. 2012), *P. lateralis* (Quinn et al. 2013), the obligate biotrophs *Hyaloperonospora arabidopsidis* (Baxter et al. 2010), *Plasmopara halstedii* (Sharma et al. 2015), *Pseudoperonospora cubensis* (Burkhardt et al. 2015; Tian et al. 2011), *Peronospora tabacina* (Derevnina et al. 2015), *Plasmopara viticola* (Dussert et al. 2016; Yin et al. 2017), *Albugo candida* (Links et al. 2011), and *Albugo laibachii* (Kemen et al. 2011), and a handful of necrotrophs and saprotrophs such as *Pythium ultimum* (Lévesque et al. 2010), the human

pathogen *Pythium insidiosum* (Rujirawat et al. 2015), and the fish pathogen *Saprolegnia parasitica* (Jiang et al. 2013).

Many of these genomes, *P. infestans* included, were assembled with a Sanger sequencing approach. Of all the oomycetes, *P. infestans* has the largest estimated genome size at around 238 Mb and is highly repetitive with an estimated 74% of the genome showing redundancy (Haas et al. 2009). These low complexity and difficult-to-assemble regions did not allow for a telomere-to-telomere quality assembly, and the scaffold N50 value for 4,921 scaffolds was 1.57 Mb with the largest scaffold being 6.9 Mb long. Indeed, only one oomycete genome has been sequenced to the scale of pseudochromosome lengths, that of *P. capsici* (Lamour et al. 2012). In their approach, they used a combination of high depth 454 sequencing reads supplemented with low depth Sanger reads. Their *de novo* assembly of these reads was the basis for single nucleotide polymorphism (SNP) calling in a sequenced cross population, and genetic map construction revealed 18 linkage groups in which 84% of the assembled genome contigs could be anchored.

Another approach to improve genome assemblies is to generate a superior *de novo* assembly using long read third-generation sequencing technologies. Among the most popular is that of Pacific Biosciences, with the ability to consistently generate reads in the 10-15 Kb range that occasionally stretch to nearly 60 Kb (reviewed in Goodwin et al. 2016). The poor per-base

accuracy of PacBio reads can be ameliorated by redundancy through high sequencing depth, or by mapping Illumina reads, which are significantly cheaper for the same coverage, and exploiting Illumina's superior per-base quality compared to PacBio to generate consensus calls. Longer reads primarily give the advantage of potentially spanning highly redundant regions of the genome through the increased chance of having unique flanking sequences. Assemblies with much longer lengths of contiguous sequences can thus be generated.

Several technologies have emerged as further manipulations to basal assemblies resulting from *de novo* assembly of PacBio or Illumina reads. One technique is to exploit the physical interaction between DNA strands close together, either within the chromatin, or on linear high molecular weight DNA. Both Dovetail Genomics and Phase Genomics employ Hi-C technology, or chromosome conformation capture (Lieberman-Aiden et al. 2009), as some of their core services. Essentially, DNA strands in the cell belonging to the same chromosome are statistically likely to be physically close together. These strands can thus be ligated together, constructed into a sequencing library, sequenced via Illumina, and finally algorithms can parse out read pairs that tended to associate together, thus informing on their general proximity along the chromosome. Recently, the *Aedes ageypti* mosquito genome was considerably improved using Hi-C technology (Dudchenko et al. 2017). Dovetail offers another service called a Chicago library which is similar to Hi-C, but is based off of high

molecular weight (HMW) DNA reconstituted *in vitro* into chromatin. The advantage of the Chicago library is its ability to finely orient smaller assembly contigs in an intrachromosomal setting, as opposed to joining reads and contigs spanning very long distances that cannot be represented by HMW DNA. A different approach to scaffolding basal assemblies, employed by Bionano Genomics, involves optical mapping (Lam et al. 2012). This technology uses special restriction enzymes (RE) called nickases to find particular RE cut sites and label HMW DNA with a special fluorophore. The DNA strands can then be run down a special flow cell one strand at a time, and the relative distances between fluorescent signals are correlated with known cut site locations on a draft basal assembly. Thus, many smaller contigs separated by difficult repetitive regions can be placed on a large optical map based on the profile of fluorescent RE cut sites alone. Several high-quality genome assemblies have incorporated this optical mapping technology, including the 240 Mb woodland strawberry (Edger et al. 2017), and the technology has improved reference genomes for maize (Jiao et al. 2017), and the phytopathogenic fungi *Sclerotinia sclerotiorum (Derbyshire et al. 2017; Moll et al. 2017)* and *Verticillium dhaliae* (Faino et al. 2015). Several reviews have discussed how best to either incorporate both optical maps and Hi-C based technology (Moll et al. 2017), and the pros and cons of either technology based on the particular genome's properties (Jiao and Schneeberger 2017; Li et al. 2017a). One final emerging technology is

Chromium, put forth by 10x Genomics. Here, HMW DNA molecules are partitioned into tens of thousands of tiny gel beads, sheared, and subsequently labeled with unique barcodes per partition. The DNA molecules are then pooled and sequenced on a normal Illumina flowcell, but the barcodes inform which fragments were physically close to each other. Preliminary studies on human datasets appear promising, though the technology is still quite new.

We undertook efforts to use three of these technologies to generate a new and improved assembly compared to the existing Broad Institute *P. infestans* reference sequence due to its current lack of organization into pseudochromosome scaffolds. In particular, we followed the route of PacBio sequencing of isolate 1306, followed by a combination of assemblies of the PacBio reads using three *de novo* assembly tools. A custom script written by our collaborators in Stefano Lonardi's group stitched together all three *de novo* assemblies, the Bionano optical map hybrid scaffolds, and a Dovetail Chicago HiRise assembly to form a superior input assembly, hereto referred as stitched7.

Our previous work (Matson et al. 2015) included the generation of a cross population of *P. infestans* derived from parents 1306 and 618, meaning that markers derived from the progeny can be used to aid the construction of genetic maps based on the parent which was PacBio sequenced. Illumina sequencing a total of 118 individuals from two cross populations allowed us to generate genetic maps for each of the populations, anchoring 218 Mb of sequence into 15

pseudochromosomes ranging from 10.1 to 22.9 Mb. This represents the first chromosome-scale reference sequence for *P. infestans*, and opens up new possibilities for the characterization of the our sexual populations in terms of not only their SNP variations, but their chromosomal and structural variations.

Materials and methods

PacBio and Dovetail DNA prep

Genomic DNA from *P. infestans* was obtained from the isolate 1306, an A1 strain isolated from tomato in California, by Dr. Audrey Ah Fong of the Judelson laboratory. Three 100 mm plates each with 20 ml of clarified rye-sucrose broth were inoculated with 10^4 sporangia, and after 5 d were harvested by filtration and frozen at -80°C. DNA was extracted by grinding tissue to a powder under liquid nitrogen. The powder was transferred to an Oak Ridge tube to which 10 ml of extraction buffer (0.2 M Tris pH 8.5, 0.25 M NaCl, 25 mM EDTA, 0.5% SDS) was added. The tube was gently inverted until the powder was in suspension. Seven ml of phenol plus 3 ml of chloroform:isoamyl alcohol (24:1) was then added. The tube was laid on its side and gently shaken on an orbital shaker for 40 min at 50 rpm. The sample was centrifuged at 10,000 ×*g* for 30 min in a swinging bucket rotor. The supernatant was removed using a wide-

bore transfer pipet and an equal volume of chloroform:isoamyl alcohol was added. The mixture was gently mixed and spun for another 5 min. The aqueous phase was carefully transferred to another Oak Ridge tube to which 50 µl of 10 mg/ml RNAse A was added, and the tube was incubated at 37°C for 30 min. After the addition of 0.6 volumes of isopropanol, DNA was precipitated by centrifugation at 10,000 ×*g* for 20 min. The supernatant was poured off, and the pellet rinsed with 70% ethanol and briefly air-dried at room temperature. DNA was resuspended in TE buffer (10 mM Tris pH 7.5; 1 mM EDTA) overnight at 4°C. The DNA concentration was measured using the Qubit Broad Range DNA quantification method. DNA quality was assessed by running a 0.8% agarose gel in 0.5 x TBE buffer using field-inversion gel electrophoresis. The running conditions were 75 V for 8h at room temperature.

Bionano DNA isolation protocol

High molecular weight DNA for Bionano Genomics (San Diego, CA) optical map analysis was isolated from *P. infestans* protoplasts. These were prepared as described by Judelson et al. (1993). Mycelia was digested in osmoticum (0.4M mannitol, 20 mM KCl, 20 mM MES pH 5.7, 10 mM $CaCl_2$) containing 5 mg/ml cellulase (Sigma, from *Trichoderma reesi*) and 10 mg/ml β-glucanase, using gentle shaking at 50 rpm for 30 min at room temperature.

Protoplasts were filtered through Miracloth and harvested by centrifugation at 700 ×*g* for 4 min. Protoplasts were washed three times in osmoticum according to Judelson et al., (1993). A fourth wash was done in MCE buffer (1M mannitol, 0.1 M sodium citrate pH 7.0 and 60 mM EDTA). Protoplasts were resuspended in the same buffer at a final concentration of 8 x10^7protoplasts/ml.

The protoplasts were then embedded in agarose plugs, using a pre-cooled plug mold (Bio-Rad, Cat # 170-3713). Before casting, a solution of 2% low melting agarose in MCE buffer was equilibrated in a 43°C water bath. The protoplast suspension was also warmed at 43°C for 30 s. Different volumes of protoplast suspension were mixed with the low melting agarose while keeping the final concentration of agarose at 0.75%. Mixing was done gently using a 200 µl wide bore pipet tip while avoiding bubble formation. The mixture was immediately transferred to the plug mold and allowed to solidify at 4°C for at least 45 min. The plugs were then dislodged using a plug mold plunger and up to 5 plugs were then transferred to a 50 ml conical tube containing 2.5 ml of freshly made lysis buffer (200 µl Qiagen Proteinase K enzyme, 0.5 M EDTA pH 9.3, 1% (w/v) sodium lauroyl sarcosine, 25 µl β-mercaptoethanol). Plugs were kept fully submerged and the digest was carried out overnight at 50°C. At the end of the overnight incubation, the lysis buffer was drained and 2.7 ml of freshly proteinase K buffer was added and incubation was carried out for a further 2 h at 50°C. After the digest, the conical tube was equilibrated to room temperature for 5 min

and 50 µl of Qiagen RNase enzyme was added. RNase incubation was for 1 h at 37°C. The plugs were first rinsed three times in 10 ml of wash buffer (10 mM Tris pH 8.0, 50 mM EDTA pH 8.0) and washed four times in 10 ml of the same buffer by gently shaking the tube on a horizontal platform at 180 rpm. The plugs were stored in the last wash buffer for transport.

Incorporating scaffolding technology and multiple *de novo* assemblies into a final input assembly

HMW DNA from *Phytophthora infestans* isolate 1306 was sent to the UC-Davis Sequencing Core for PacBio SMRT sequencing, Bionano optical map construction at Kansas State university, and finally to Dovetail genomics (Santa Cruz, CA) for Chicago library assembly and sequencing. The PacBio output resulted in an unfiltered output of 24.8 Gb of sequence in 1.66 million reads. A collaborator from the UCR Computer Sciences department generated five *de novo* assemblies from the PacBio reads. Three originated from ABruijn v0.4 (Lin et al. 2016) and two from CANU v1.5 (Berlin et al. 2015; Koren et al. 2017), and one from FALCON (Chin et al. 2016). The ABruijn assembler was run on (1) all the PacBio reads at least 10Kb long (75x genome equivalents), (2) the PacBio reads corrected by CANU (70x genome equivalents), and (3) the PacBio reads corrected and trimmed by CANU (68x genome equivalents). The CANU assembler was run on (1) all the PacBio reads (41x genome equivalents) and (2)

all the PacBio reads at least 10 Kb long (33x genome equivalents). The FALCON assembly was carried out by the UC Davis core facility.

A custom script was used by Weihua Pan and Qihua Li in Stefano Lonardi's laboratory to detect and break chimeric contigs in the six assemblies by aligning them to the optical map. The chimera-free assemblies were merged using Novo&Stitch (Pan et al. Personal communication) using, again, the optical map. The stitched assembly was polished using Quiver (PacBio), followed by PILON (Walker et al. 2014). The PILON assembly was processed by Dovetail, using their HiRise pipeline. Finally, scaffolds were created using the optical map with the Kansas State University pipeline (Shelton et al. 2015) and assembled into a reference sequenced named "stitched6". In order to check for completeness, we first assembled Illumina RNA-seq reads into 25,881 putatively expressed transcripts using Trinity (Grabherr et al. 2011), allowing us to examine if the expected number of genes is present in the assembly. Only 96.71% of these transcripts mapped to the assembly, prompting us to identify and manually add in 34 contigs from the UC-Davis FALCON assembly, representing 21.4 Mb of missing sequence. The source of this apparent loss of sequence is unclear at the moment. This assembly is named "stitched7", and was not further modified prior to genetic mapping and pseudochromosome assembly. Next, the Benchmarking Universal Single-Copy Orthologs (BUSCO)(Simão et al. 2015) suite was used to further investigate completeness. Both the *Toxoplasma-*

trained protist lineage dataset, and the *Aspergillus* fungal lineage dataset, were used in the search for BUSCOs.

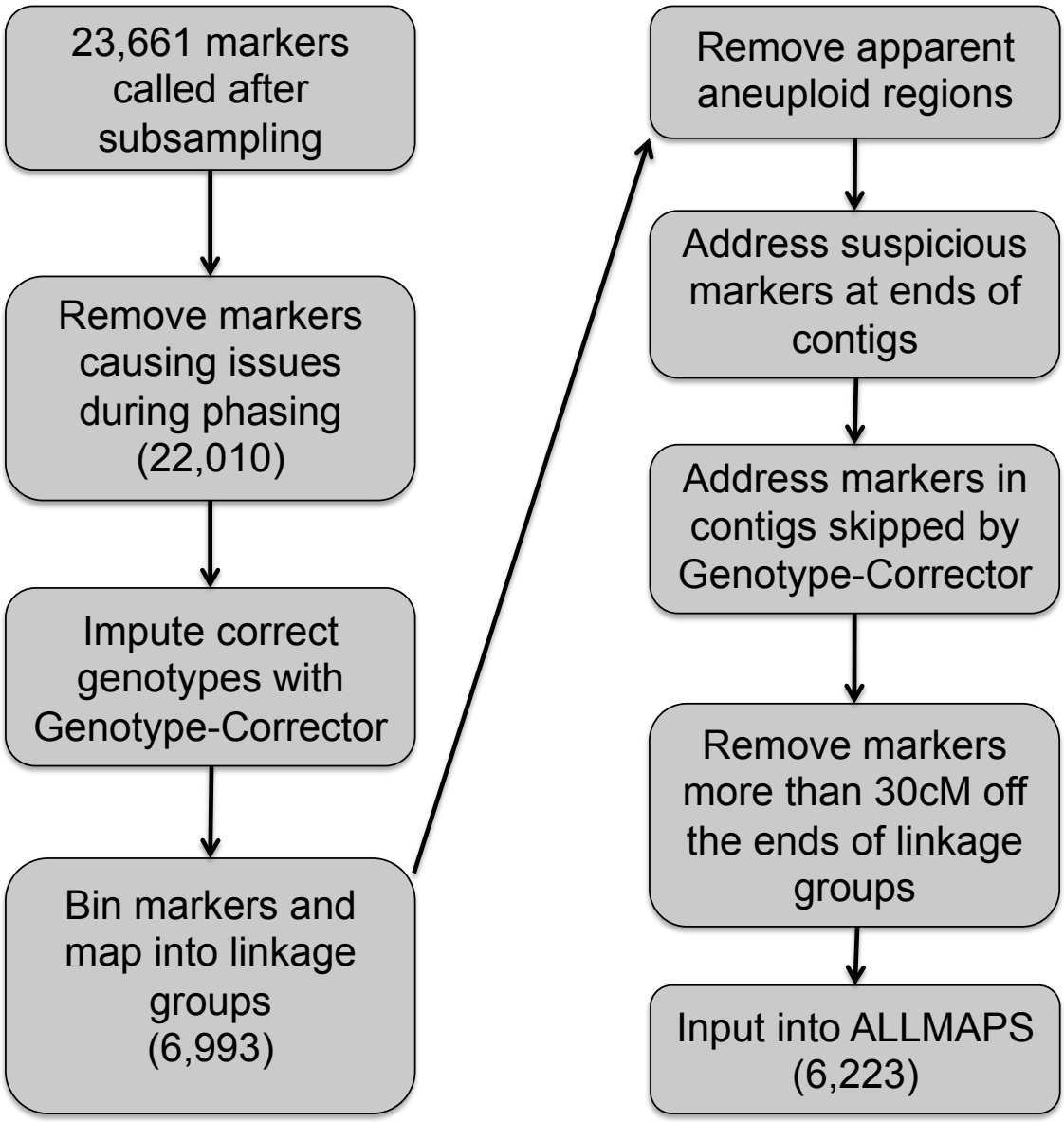
Illumina sequencing and marker selection

To achieve a chromosome-scale assembly, we used a genetic-mapping approach. Illumina sequencing was carried out on a total of 118 parents and progeny originating from two separate crosses, referred to as “crossM” and “cross20”. CrossM was the result of a cross between parents 1306 (A1 mating type) and 618 (A2 mating type), and of 82 sequenced progeny, 76 ultimately contributed to genetic map construction (Matson et al. 2015). Cross20 resulted from a cross between parents 6629 and 550, and all 32 sequenced progeny contributed to the genetic maps (Judelson 1997). Single oospore cultures were grown in liquid rye-B media and either extracted using a phenol/chloroform method (Judelson and Michelmore 1991), or with the GeneJET Plant Genomic DNA Purification Kit (Thermo Scientific, Waltham MA). All progeny were whole-genome sequenced to a depth of 15-30 ×, with ~500bp inserts, and over three rounds. For the first round, libraries for 29 crossM progeny were generated using the Illumina PCR-free library prep kit (Illumina, San Diego CA) and sequenced at 2×100 bp at the UCR genomics core. For the second round, PCR-free libraries for 28 crossM and all 34 cross20 progeny and parents were constructed and

sequenced at Cofactor genomics at 2×150 bp, and PCR-free libraries for the remaining 25 crossM progeny were constructed and sequenced at the UC-Davis sequencing core at 2×150 bp. Libraries for crossM parents 1306 and 618 were constructed and sequenced at UCR, and an additional Illumina miSeq library for 1306 was constructed and sequenced at UC-Davis (2×250 bp).

All sequencing reads were aligned to the stitched7 assembly using bwa-mem (Li and Durbin 2009) and variants called using an adaptation of the GATK best practices guidelines (McKenna et al. 2010)(**Figure 2.1**) retaining all SNPs and indels 3bp and smaller. Markers were generated to resemble a pseudo-testcross, *i.e*, Aa × AA or AA × Aa, and filtered to include markers showing Mendialian segregation (X^2 < 0.05). Due to the relatively low sequencing coverage, a high rate of miscalls was present, typically due to a low number of alternate allele reads forcing a homozygous-reference call when a heterozygous call is the true genotype. These miscalls can considerably bloat genetic map sizes due to the large number of implied recombination events, and we sought to impute correct genotypes when possible. Genotypes of each parent were first split into two vcf files consisting of SNP marker types: heterozygous in parent1 and homozygous reference in parent2, and vice versa. Each of these vcf files was subsampled with vcftools (Danecek et al. 2011) to ensure at least a 5 Kb gap between consecutive markers, resulting in 23,661 markers distributed between the four parents. The phase of each marker was determined by

Figure 2.1. Flowchart of the process used to curate markers used as input for the genetic maps. Markers originated from variants existing in a pseudo-testcross state in all four parents from both crosses, and consisted of SNPs and indels < 3bp. Curation began with a subset of 23,661 markers subsampled every 5 kb from all called variants. 1,651 markers were removed since they were found to be causing problems in proper haplotype phase determination. These 22,010 markers were used as an input for the program Genotype-corrector, and following correction and marker binning, 6,993 markers could be further curated by hand. Four rounds of genetic mapping and manual correction addressed aneuploid regions, inconsistent markers at the very ends of contigs, contigs skipped by Genotype-corrector due to low marker (threshold > 15) count, and markers hanging more than 30 cM off the end of a linkage group. The remaining 6,223 markers were inputted into the final genetic maps used in pseudochromosome assembly.


23,661 markers called after subsampling
Remove markers causing issues during phasing (22,010)
Impute correct genotypes with Genotype-Corrector
Bin markers and map into linkage groups (6,993)
Remove apparent aneuploid regions
Address suspicious markers at ends of contigs
Address markers in contigs skipped by Genotype-Corrector
Remove markers more than 30cM off the ends of linkage groups
Input into ALLMAPS (6,223)

Beagle/4.0 (Browning and Browning 2007), using a window size of 1000 and an overlap of 300. This generates another two files for a grand total of eight total genotype files corresponding to each phase, of each marker type, for each parent in either cross.

Since the markers are organized in a pseudo-testcross, the expected ratio of heterozygous:homozygous markers is 1:1, but some markers show inconsistent segregation patterns which can cause Beagle to improperly sort phased SNPs. These are usually caused by miscalls in the parents and show a 1:2:1 or all-heterozygous marker ratios. For parents 1306, 618, 6629, and 550, 758, 567, 246, and 271 markers showing unexpected segregation ratios were removed, respectively, leaving a total of 22,010 properly phased markers. The python script Genotype-corrector (https://github.com/freemao/Genotype-corrector/ wiki/Genotype-Corrector) was then used to identify miscalls and impute a more parsimonious genotype using default parameters and an F2 population type. Contigs containing fewer than 15 markers were excluded from correction. A total of 1,380,769 genotypes between both crosses were taken into consideration, and a total of 15,888 corrections were made (1.16%), the majority of which were corrected from the homozogous to the heterozygous states (85.1%). Genotype-corrector then binned consecutive identical markers, leaving a total of 6,119 markers to use as an input for genetic mapping from the stitched6

assembly. The addition of the FALCON contigs brought another 874 markers, for a total of 6,993 markers passed to JoinMap.

Genetic mapping, error correction, and pseudochromosome assembly

The program JoinMap/5.0 (Kyazma, Wageningen, Netherlands) generated four independent genetic maps, each representing one of the cross parents. The number of linkage groups ranged from 14-17 for each parent, and initial genetic map sizes ranged upwards of 5,000 cM, appearing considerably bloated due to errors not modified by Genotype-corrector. Four additional iterative rounds of error correction took place which attempted to flag non-parsimonious genotypes in the context of recombination events. Briefly, genotypes were either manually corrected if the error appeared to be a homozygous miscall and was flanked by markers residing in the phase of the majority, or omitted if the error appeared to be a heterozygous miscall in the first round. The first round primarily addressed miscalls due to apparent aneuploidy, which can force genotypes to exist as completely hetero- or homozygous across an entire contig regardless of haplotype phase. All markers showing this were eliminated from the entire region of the contig until alternate phases resumed. The second round of corrections focused on apparent miscalls at the beginning or end of contigs, and removed all miscalls if no other contig recombined within the linkage group. The third round

removed miscalls inside the smaller contigs which could not be analyzed by Genotype-corrector. Markers spanning less than 100 Kb in an entire contig were assumed to not recombine, and a majority-rules approach either omitted minority miscalls, or eliminated all calls per progeny per contig. The final round eliminated beginning or ending markers sitting more than 30 cM from the next nearest marker. After the iterations, 770 markers were removed entirely from the set of 6,993 input markers. This resulted in 3,281 individual genotypes being omitted from the cross20 progeny, and 10,196 from the crossM progeny. The final, polished genetic maps were then assembled into pseudochromosomes using ALLMAPS (Tang et al. 2015) and incorporated 6,166 total markers. Most of these markers are redundant, but removing redundancy to keep markers at somewhat regular intervals was not performed. The 1306 linkage groups were used as a guide for representing the assembly in ALLMAPS, and 100 Ns were inserted between the adjacent assembled stitched7 contigs that lacked sequence overlaps. All Illumina sequencing fastq files were subsequently remapped to the new assembly in the same manner as to stitched7. All heterozygous polymorphisms were called via GATK in the same fashion as in Stitched7. Telomere sequences were searched by using BLAST and the hypothetical telomere repeat sequence "GTTTAGGGTTTAGGGTTTAGGGTTT".

Gene annotation

Annotations were generated by Qihua Liang in the Lonardi lab at UCR. Input genes consisted of improved models of the Broad Institute's original gene annotations, and were previously generated in our laboratory (Shrivistava 2017). Protein sequences from these models were mapped to the new assembly using TBLASTX to identify their new genomic coordinates. Repeats were predicted with RepeatScout (Price et al. 2005) and assigned to common transposon families with TransposonPSI (Haas 2010). Genes mapping to the pseudochromosomes and all heterozygous polymorphisms (SNPs, small indels > 6 bp) were plotted along the assemblies with the R package karyoploteR (Gel and Serra 2017).

Results

Genome Assembly

The summary statistics for each assembly of the PacBio reads are shown in **Table 2.1**. Briefly, neither the Dovetail large-insert Chicago library, nor Bionano optical map improved the Falcon *de novo* PacBio assemblies on their own to chromosome-sized, with the Dovetail N50 doubling the Falcon-Quiver

Table 2.1. Summary statistics of the input sequences for the stitched7 assembly

	Broad assembly	Abruijn k=17 >10Kbp reads	Abruijn k=17 Canu corrected reads	Abruijn k=17 Canu trimmed reads	Canu all reads	Canu v1.5 >10Kb reads	Falcon + Quiver (UCDavis)	Falcon + Dovetail	Falcon + Bionano	Stitched6	Stitched7
PacBio: input (fold coverage)[a]	n/a	74.75	69.70	68.03	40.66	33.23	105.23	n/a	n/a	n/a	n/a
PacBio: # input reads	n/a	1,079,466	1,965,291	1,902,058	571,547	457,793	3,451,682	n/a	n/a	n/a	n/a
Assembly: contig/scaffold N50 (bp)	1,588,622	430,704	302,078	315,766	135,469	133,977	558,424	1,149,151	679,423	1,003,410	969,765
Assembly: contig/scaffold L50	38	113	160	162	464	449	87	49	76	64	71
Assembly: total assembled (bp)	228,543,505	195,768,168	177,232,870	175,149,119	292,352,599	305,686,040	215,910,203	215,939,503	219,335,213	232,339,072	246,652,131
Assembly: # assembled contigs	4,921	770	867	849	2,853	3,480	1,318	1,263	1,286	1,048	1,073
Assembly: # contigs > 1Mb	51	26	13	11	7	13	37	54	47	65	49
Assembly: longest contig (bp)	6,928,287	3,248,854	2,004,950	1,638,783	1,813,497	1,810,393	4,206,720	4,501,375	4,206,720	5,545,037	4,269,999
# of alignments between assembly and opt map[b]	241	645	550	541	679	653	605	-	-	517	627
# chimeric contigs (overhang >100Kb)[c]	119	96	30	27	27	22	68	-	-	26	37
Transcripts, % mapped (25,881)[d]	-	-	-	-	-	-	99.7905%	99.7905%	99.7905%	96.7080%	99.5958%
Transcripts, % MAPQ>=30 (25,881)[e]	-	-	-	-	-	-	76.9750%	76.9750%	76.9750%	66.0253%	61.2787%

[a]Genome wide coverage of the input reads assuming a 238 Mb genome size

[b]Successful alignments (confdience score > 20.0) between Bionano optical maps and assembly contigs

[c]Number of contigs partially aligned to an optical map, but with >100Kb of the optical map unaligned to corresponding sequence

[d]% of total Trinity transcripts mapped to each assembly

[e]% of total Trinity transcripts mapped with high-confidence to each assembly

assembly, and the Bionano N50 adding about 120Kb to the N50 value. Despite this, the optical map proved useful in resolving between 27-37 chimeric contigs originating from the assemblers. Our final stitched7 assembly consisted of 246.7 Mb of sequence spread across 1,073 scaffolds with an N50 of 0.97 Mb (**Table 2.1**), similar to the N50 of the the Falcon + Dovetail and Falcon + Bionano assemblies, but with additional sequence.

Genome completeness

99.6% of the Trinity transcripts (25,881) mapped to the stitched7 assembly using bwa-mem (Li and Durbin 2009), 61.3% of which (15,860) had MAPQ scores great than 40 (**Table 2.1**). BUSCO employs curated datasets to search for genes predicted to be not only universal to all species within a lineage, but present only in single copies. In the protist lineage training set, we identified 209 complete BUSCOs out of a total of 215 possible (97.2%), 191 (88.8%) of which are present in single copies and 18 present in at least two copies. 12 of these duplicated BUSCOs had one member on a pseudochromosomes (described below) and a second on a contig that had not mapped to a pseudochromosome. We repeated the analysis using the fungal dataset, and of 290 total BUSCOs, only 251 were observed in the genome (86.5%), 230 of which were present in single copies and 21 in two copies. For comparison, 5 protist BUSCOs are

duplicated in the Broad assembly, only one of which was shared with the Stitched7 assembly. Additionally, 17 fungal BUSCOs are duplicated in the Broad assembly, eight of which are shared with the 21 Stitched7 duplicated BUSCOs.

Genetic mapping and pseudochromosome assembly

23,661 pseudo-testcross markers from each parent, subsampled every 5 Kb, were curated through four rounds of genotype correction and marker consolidation (**Figure 2.1, 2.2, Table 2.2, 2.3**). These markers consisted of SNPs and small indel variants, of which 2,039, 2,108, 874, and 1,202 were integrated into linkage groups of parents 1306, 618, 6629, and 550, respectively (**Table 2.4**). These 6,166 markers, integrated into the genetic maps, allowed the resolution of 15 linkage groups each for parents 1306 and 618 of CrossM, and 16 and 17 linkage groups for parents 6629 and 550 of Cross20, respectively (**Table 2.5**). The greater number of linkage groups for the Cross20 parents likely originated from the fewer number of progeny contributing to the genetic maps, thus lowering the mapping resolution. Total genomic genetic sizes for parents 1306, 618, 6629, and 550 were 1,870, 1,289, 1,608, and 1,658 cM, respectively, and all four genetic maps averaged at 1606.3 cM. However, two linkage groups in parent 1306 (lg2, 251 cM and lg6, 498 cM) could not be compacted to genetic sizes comparable to the 618 linkage groups (104 and 103 cM, respectively). This

Figure 2.2. Examples of genotypes to be corrected in contig20. Haplotype phase was determined by Beagle, and the genotypes are coded as heterozygous, “X”, or homozygous, “AA”. Highlighted genotypes are in the “A” phase and heterozygous or in the “B” phase and homozygous, revealing recombination tracks and crossover events. **A**) Uncorrected genotype calls for contig20 from positions 465,936-1,653,120. ~85% of miscalls fall into the category of a false homozygous call, where the true heterozygous genotype is more likely, as evidenced by marker ‘contig20_787347’ in sample 5.36. This happens because low sequencing coverage, and occastional low alternate read depth, leads HaplotypeCaller to assume homozygosity due to lack of evidence. Genotypes called as heterozygous when the homozygous genotype is more likely were excluded form analysis, as evidenced by marker ‘contig20_576015’ in sample 5.36. Genotypes with a suspected miscall residing at the end of contigs were excluded from analysis, such as in marker ‘contig20_1653120’. Progeny showing gain or loss of heterozygosity suggest aneuploidy, as seen in progeny 7.13 and 7.19, and all such markers were excluded within the contig. Additionally, some progeny showed what appeared to be local gain or loss of heterozygosity, such as in sample M127, and suspect markers were excluded until phase seemed to return to normal. **B**) Corrected markers of contig20. These markers and their genetic positions would be passed onto ALLMAPS for final pseudochromosome assembly.

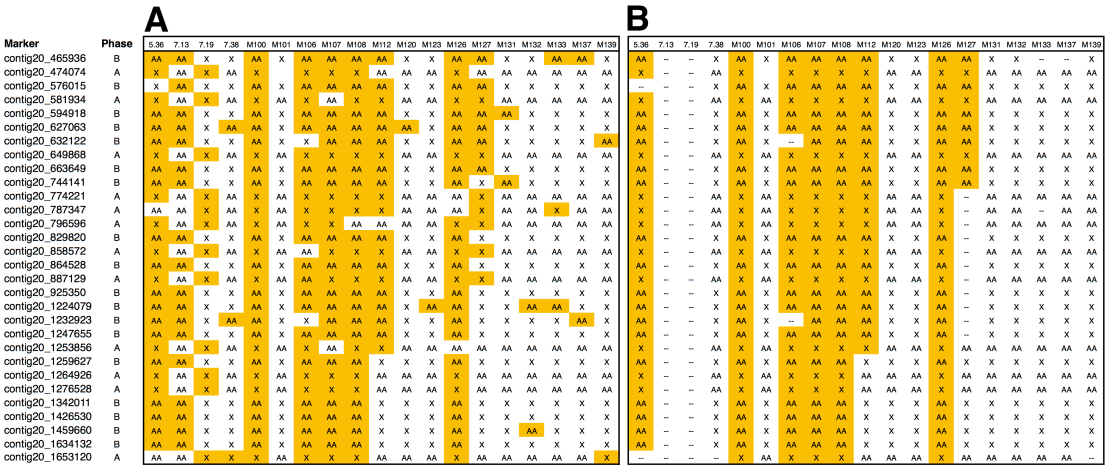

A
B
Marker
Phase
5.36 7.13 7.19 7.38 M100 M101 M106 M107 M108 M112 M120 M123 M126 M127 M131 M132 M133 M137 M139
contig20_465936 B
contig20_474074 A
contig20_576015 B
contig20_581934 A
contig20_594918 B
contig20_627063 B
contig20_632122 B
contig20_649868 A
contig20_663649 B
contig20_744141 B
contig20_774221 A
contig20_787347 A
contig20_796596 A
contig20_829820 B
contig20_858572 A
contig20_864528 B
contig20_887129 A
contig20_925350 B
contig20_1224079 B
contig20_1232923 B
contig20_1247655 B
contig20_1253856 A
contig20_1259627 B
contig20_1264926 A
contig20_1276528 A
contig20_1342011 B
contig20_1426530 B
contig20_1459660 B
contig20_1634132 B
contig20_1653120 A

Table 2.2. Summary of genetic markers and individual genotypes curated in preparation for the final genetic maps.

	1306-A[a]	1306-B[b]	618-A	618-B	6629-A	6629-B	550-A	550-B	**All markers**
Total input markers[c]	2558	4229	2887	4704	1593	2896	811	2332	**22010**
HomRef genotypes	100443	159574	109773	178306	22580	46886	13446	36831	**667839**
Het genotypes	99080	170286	115255	188604	28547	51577	14126	45455	**712930**
Total genotypes	199523	329860	225028	366910	51127	98463	27572	82286	**1380769**
HomRef -> Het corrections	1561	2902	2434	3919	641	1117	257	650	**13481**
Het -> HomRef corrections	359	618	409	586	110	169	35	121	**2407**
Total corrections	1920	3520	2843	4505	751	1286	292	771	**15888**
Binned markers[d]	822	1182	791	1224	414	802	258	626	**6119**
FALCON markers injected[e]	272	-	335	-	137	-	130	-	**874**
Total Markers input to JoinMap[f]	2276	-	2350	-	1353	-	1014	-	**6993**
									Average
% Error rate[g], HomRef -> Het	0.78	0.88	1.08	1.07	1.25	1.13	0.93	0.79	**99.03**
% Error rate[g], Het -> HomRef	0.18	0.19	0.18	0.16	0.22	0.17	0.13	0.15	**17.12**
% HomRef -> Het of total # of corrections	81.30	82.44	85.61	86.99	85.35	86.86	88.01	84.31	**85.11**
% Het -> HomRef of total # of corrections	18.70	17.56	14.39	13.01	14.65	13.14	11.99	15.69	**14.89**
% correction rate	0.96	1.07	1.26	1.23	1.47	1.31	1.06	0.94	**1.16**
Markers in the final genetic maps[h]	2039	-	2108	-	874	-	1202	-	**6223**

[a,b]A and B refer to the arbitrary haplotype phases as determined by Beagle
[c]Markers remaining following 5kb subsampling and removal of markers inducing phasing errors
[d]Adjacent markers with identical recombination profiles were consolidated with Genotype-Corrector
[e]Markers origination from the FALCON assembly added during the Stiched6 to Stitched7 transition
[f]Markers passed to JoinMap after the first round of error correction per parent
[g]% Error rate among all genotypes of the total input markers
[h]Markers passed to the final ALLMAPS step after four rounds of error correction

Table 2.3. Genetic size of parent 1306 over four rounds of genotype correction and curation.

Chromosome/ Linkage group	Round1[a]	Round2	Round3	Round4	**Final Size[b]**
chr1	593.4	299.3	114.7	114.6	**114.6**
chr2	576.8	326.3	251.3	271.3	**251.2**
chr3	259.0	98.9	70.9	103.3	**103.3**
chr4	379.2	150.3	109.8	120.9	**120.9**
chr5	563.3	71.2	50.7	77.7	**77.7**
chr6	1009.8	811.4	766.9	497.7	**497.7**
chr7	536.0	174.3	66.2	71.6	**71.6**
chr8	332.5	89.2	55.4	58.0	**58.0**
chr9	312.7	95.5	72.7	72.0	**72.0**
chr10	283.0	101.1	72.9	78.5	**78.5**
chr11	319.4	108.1	68.9	71.5	**71.5**
chr12	215.6	72.3	60.0	57.3	**57.3**
chr13	323.0	187.9	139.8	154.3	**154.3**
chr14	-	84.8	72.8	84.5	**84.5**
chr15	194.0	81.9	54.1	56.7	**56.7**
Total[c]	5897.8	2752.5	2027.1	1889.9	**1869.8**

[a]Sizes in cM

[b]Size of each linkage group when passed to ALLMAPS

[c]Total sizes of the linkage groups

Table 2.4. Summary of the markers and linkage groups contributing to the assembly of the pseudochromosomes for each parent.

Parent Linkage Groups	1306	618	6629	550
Number of linkage groups	15	15	16	17
Markers for all linkage groups	2,039	2,108	874	1,202
Markers per Mb	10.9	11.5	6.1	7.8
Number scaffolds incorporated	274	266	223	252
Scaffolds with 1 marker	87	69	91	83
Scaffolds with 2 markers	29	31	34	36
Scaffolds with 3 markers	15	18	23	22
Scaffolds with >=4 markers	143	148	75	111
Total contributed sequence	187,206,658 (75.7%)	183577805 (74.3%)	143057160 (57.9%)	153,937,509 (62.3%)

Consensus Summary	Anchored	Oriented	Unplaced
Markers in all linkage groups	6,166	4,707	27
Markers per Mb	28.2	30.6	1.0
Total scaffolds incorporated	417	181	656
Scaffolds with 1 marker	100	0	10
Scaffolds with 2 markers	44	10	4
Scaffolds with 3 markers	22	6	3
Scaffolds with >=4 markers	251	165	0
Total contributed sequence	218,832,396 (88.5%)	153,620,982 (62.1%)	28359575 (11.5%)

likely originated from unfavorable marker configurations, making error correction difficult and substantially inflating the genetic size of parent 1306.

Since DNA from parent 1306 was used to construct the PacBio sequences and thus the stitched7 input, the linkage groups from parent 1306 were used as a guide to building the chromosomes, resolving 15 pseudochromosomes in the *P. infestans* assembly. Each genetic map from the crossM parents could individually resolve 74.3 and 75.7% of the stitched7 assembly, while the cross20 parents could only each resolve 57.9-62.3% (**Table 2.5**). Altogether, though, these chromosomes anchored 218.8 Mb (88.5%) of sequence and confidently oriented 153.6 Mb (62.1%) in the direction suggested by the genetic maps. Telomere sequences were observed on seven assembled pseudochromosomes and four unassembled contigs. Only chromosomes 2, 7, and 8 contained telomere hits within 50Kb of the end of the assembly, indicating missing assembly and misplaced component contigs.

Many genetic maps and their relative physical coordinates conformed to metacentric patterns of higher recombination rates near the chromosome ends, with lower recombination rates near the center and centromere. However, some chromosomes, such as chr7, 14, and 15 only show a plateau of recombination rates at the terminal ends of the assembly. It is unclear if those that did not follow this pattern were due to a broken chromosome assembly or if the centromere resided at the end of the chromosome. Most chromosomes

Table 2.5. Genetic sizes of linkage groups for each parent.

Chromosome Number	Size[a]	1306 LGs	Size[b]	618 LGs	Size[b]	6629 LGs	Size[b]	550 LGs	Size[b]
chr1	22923084	lg1	115	lg6	138	lg3,lg11	324	lg2,lg6	180
chr2	21262383	lg2	251	lg4	104	lg2,lg5	224	lg4,lg7	236
chr3	16897382	lg3	103	lg3	67	lg14	43	lg10	151
chr4	16741615	lg4	121	lg9	78	lg10	168	lg9	176
chr5	14697857	lg5	78	lg13	67	lg1	179	lg3	113
chr6	14531184	lg6	498	lg5	103	-	-	lg12,lg14	186
chr7	13769070	lg7	72	lg1	51	lg4	75	lg11	70
chr8	13450935	lg8	58	lg10	71	lg8	49	lg17	27
chr9	13431482	lg9	72	lg12	80	lg7	92	lg15,lg16	67
chr10	13206068	lg10	79	lg2	76	lg15,lg16	128	-	-
chr11	12896474	lg11	72	lg7	137	lg13	61	lg8	119
chr12	12552869	lg12	57	lg15	75	lg12	55	-	-
chr13	11760846	lg13	154	lg8	76	lg9	123	lg1	140
chr14	10674240	lg14	85	lg11	111	lg6	87	lg11	111
chr15	10077107	lg15	57	lg14	56	-	-	lg13	82
Total sizes[c]	218872596		1870		1289		1608		1658

[a]In base pairs

[b]In centiMorgans (cM)

[c]Physical (bp) and genetic (cM) sizes for the assembled pseudochromosomes

contained nearly 1:1 coverage of all four parental linkage groups along the physical assembly (**Figure 2.3**). However, not all regions of the pseudochromosomes were equally covered by markers originating from all four genetic maps (**Figure 2.4**). In particular, some crossM linkage groups/pseudochromosomes seemed to lack marker representation from one of the cross20 genetic maps (**Table 2.5**). Examples include a lack of markers from parent 6629 mapping to chromosomes 6 and 15, while parent 550 lacks markers mapping to chromosomes 10 and 12. Variants for parents 550 and 6629 can still be identified on all 15 chromosomes, meaning the chromosomes are still present in both parents. Chromosome8/lg10 from parent 618 suggests an apparent large deletion in compared to parent 1306 due to a lack of contributing markers spanning over 8 Mb (**Table 2.5, chr8**). Curiously, parents 550 and 6629 only contribute markers on opposing ends of chromosome8, suggesting a relationship between those regions.

The failure of smaller Cross20 linkage groups to be consolidated into one, such as in chromosomes 1 and 2 where both parents 6629 and 550 contributed two linkage groups each per chromosome (**Figure 2.4, chr1,2**) highlights another potential problem with the Cross20 linkage groups. In these cases, their markers still map to the physical in a mostly 1:1 ratio, though they appear genetically unlinked by themselves. Finally, signatures of structural variation between parent 1306 and the other three parents can be seen in the form of “forked” scatter plots

Figure 2.3. ALLMAPS representation of the genetic and physical structure of pseudochromosome 14. **A**) Linkage groups from all four parents and their markers contributing to the assembly of chromosome 14. Each black horizontal line represents an individual marker, connected to the physical assembly by each colored line. Alternating shades of white and grey represent divisions between different stitch7 contigs. **B**) Scatter plots of genetic (cM) vs physical (Mbp) distance for all four linkage groups. Noise is attributed to slight order differences between markers and the genomic contigs they map to, which in turn probably originates from minor structural rearrangements between the parents or genotype errors still present, complicating a consensus assembly. The physical representation of the assembly again alternates white and grey to represent the component contigs, which are projected to the scatter plots via the vertical yellow lines. The Pearson correlation coefficient (ρ) is calculated for genetic vs physical distance each linkage group. Values closer to 1 or-1 represent better colinearity of markers, indicating a better relationship between genetic and physical distance. Based on these values, the nearly complete coverage of markers from all four parents along the length of the assembly, and the lack of obvious rearrangements, chromosome 14 seems reasonably well assembled despite the genetic vs physical distance noise.

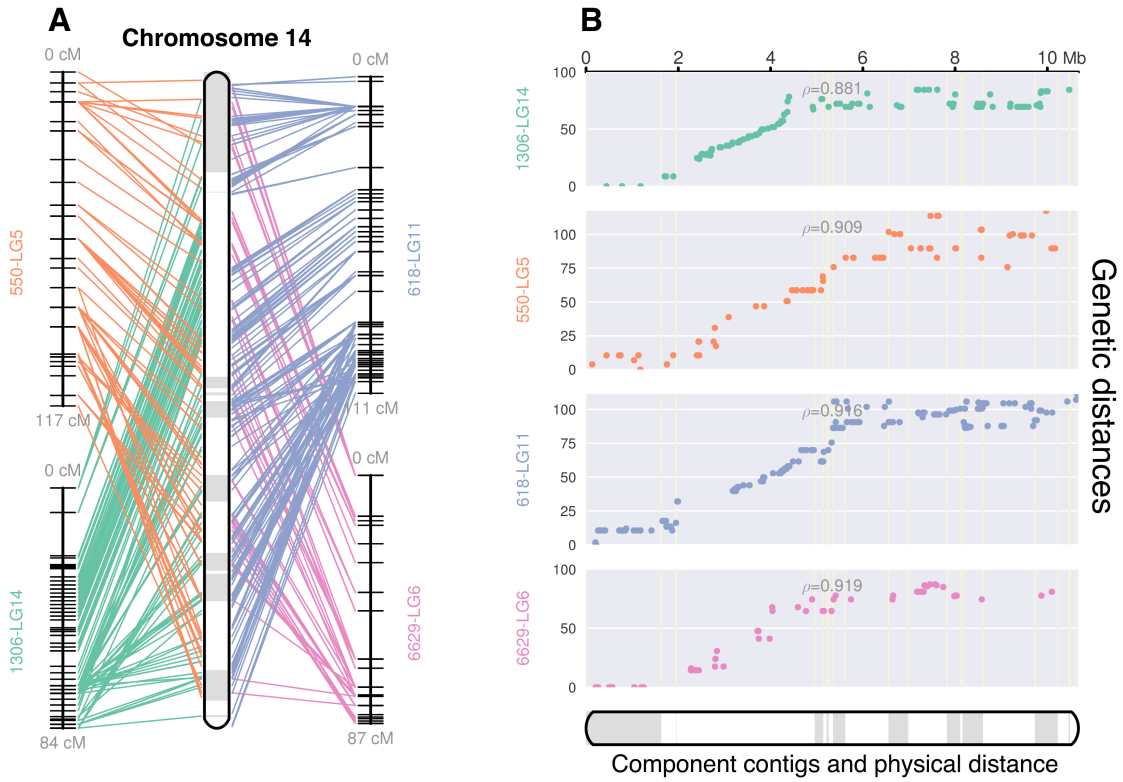

A
Chromosome 14
0 cM
0 cM
550-LG5
618-LG11
117 cM
11 cM
0 cM
0 cM
1306-LG14
6629-LG6
84 cM
87 cM
B
0
2
4
6
8
10 Mb
1306-LG14
ρ=0.881
550-LG5
ρ=0.909
618-LG11
ρ=0.916
6629-LG6
ρ=0.919
Genetic distances
Component contigs and physical distance

Figure 2.4. Fifteen pseudochromosome assemblies. Representations of each pseudochromosome include the relation between linkage group markers and their physical coordinates (left) and dot plots of the genetic distance compared to physical distance (right). Many chromosomes contain more than one linkage group from a single parent contributing to the assembly. For example, parent 550 LG2 and LG6, and parent 6629 LG3 and LG11 all contribute to chromosome 1. Some chromosomes completely lack markers from one parent, such as chromosomes 10 and 12 which lack markers from parent 550, while chromosome 15 lacks markers from parent 6629. Chromosome 8 appears to lack markers from parent 618 contributing to around 8Mb of the assembly, while the markers from parent 1306 in chromosome 6 appear highly bloated. Either case could be explained by challenges during the genotype correction process if erroneous markers could not be corrected or if markers were deleted. Finally, the “forking” pattern seen in the plot of chromosome 5 parent 6629, and elsewhere in the genome, suggests a chromosomal rearrangement compared to other three parents.

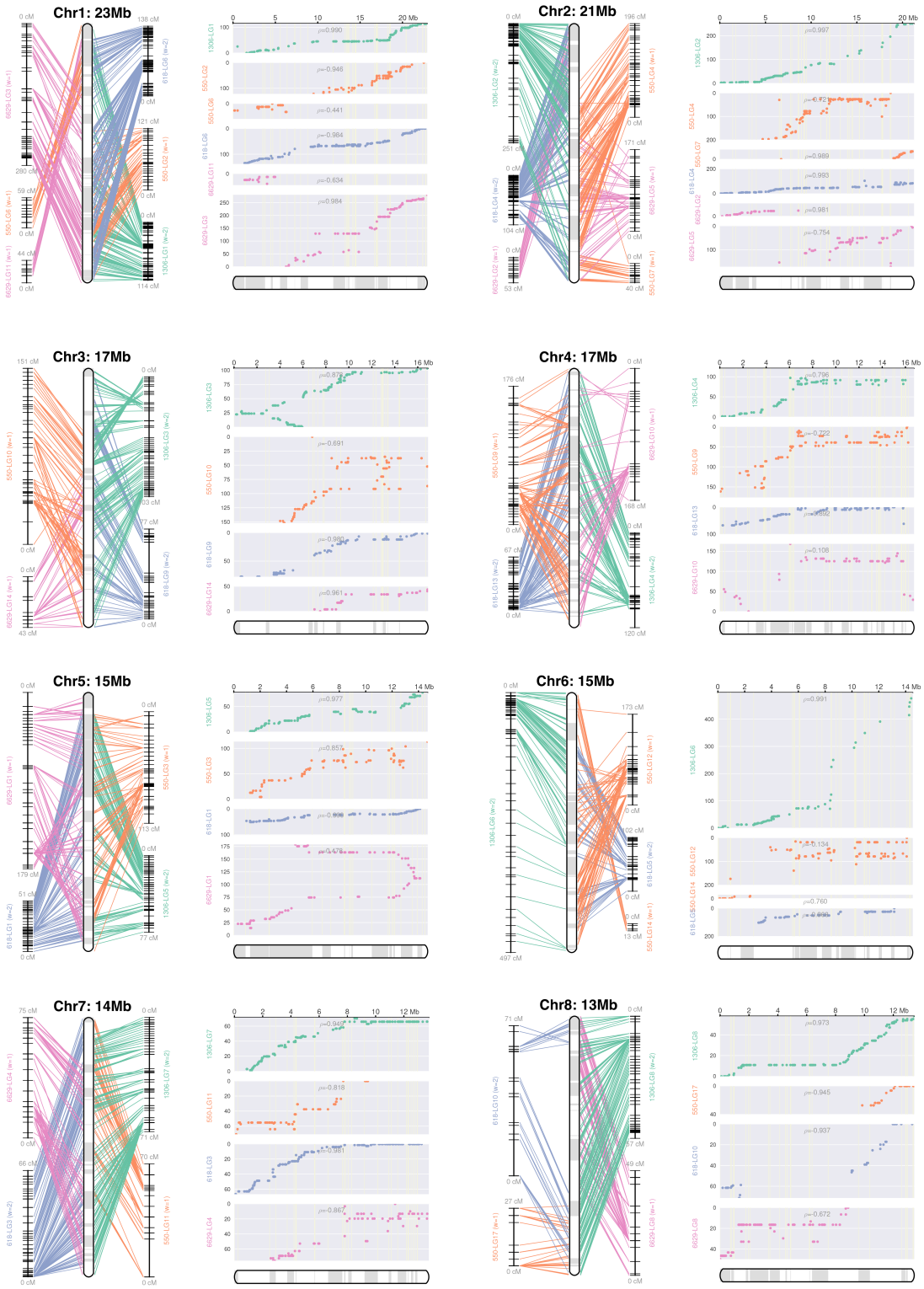
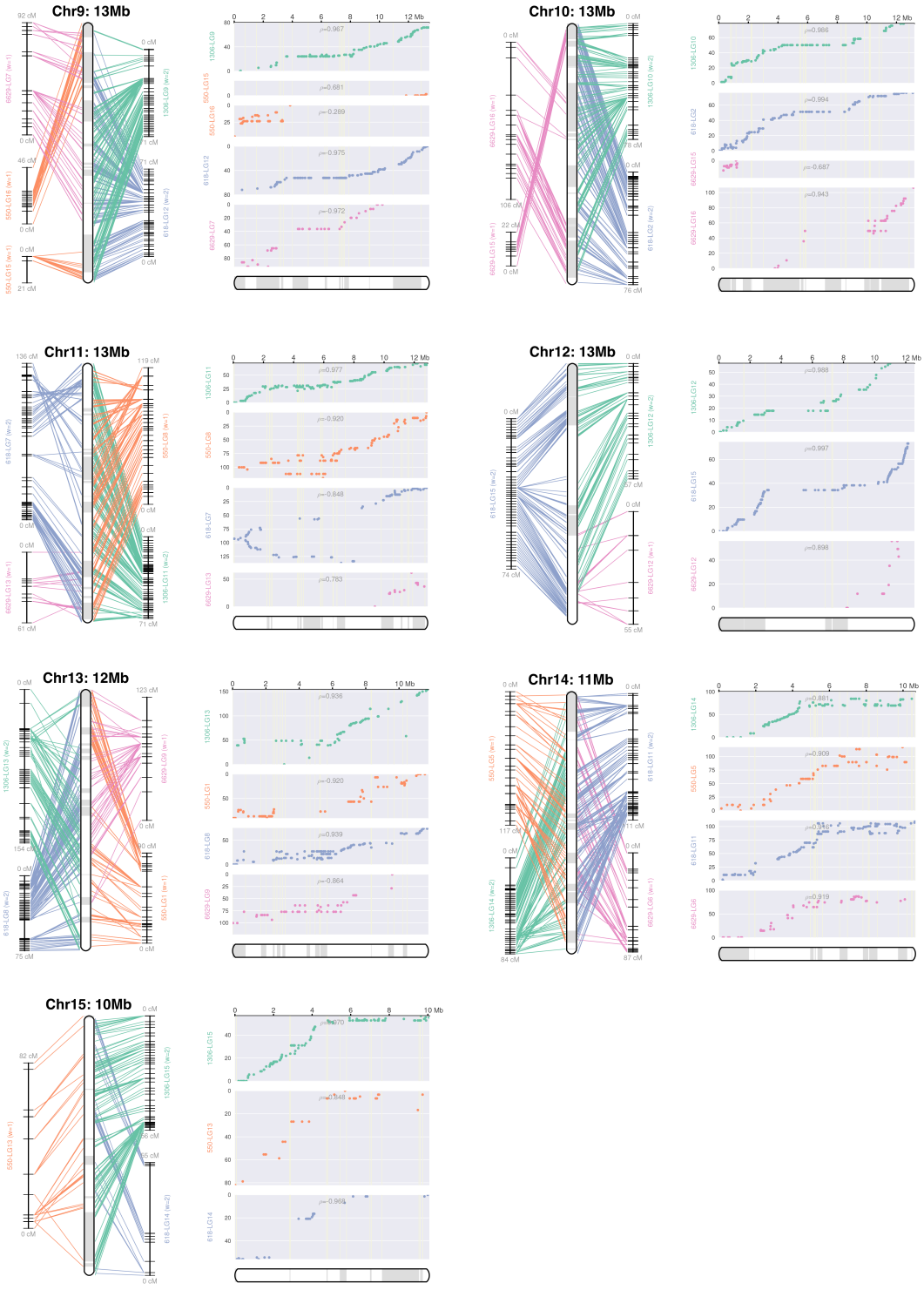

Chr1: 23Mb
Chr2: 21Mb
Chr3: 17Mb
Chr4: 17Mb
Chr5: 15Mb
Chr6: 15Mb
Chr7: 14Mb
Chr8: 13Mb

Figure 2.4 continued



when comparing genetic distance vs physical distance. Examples of this include parent 6629 lg1 in chr 5, and parent 618 lg7 in chr11.

Gene annotation

Of 18,672 unique genes (Shrivistava 2017), 14,169 uniquely mapped to the new reference assembly without duplicate hits, and 20,030 mapped in total. 7,397 of the original 18,672 genes could not be located in the new assembly and 3,149 genes were present in two at least two copies. Gene counts along the pseudochromosomes ranged from 646 for chr15, to 2,332 for chr2 (**Table 2.6**, **Figure 2.5**). Gene density, in genes per 10 Kb, ranged from 0.640 to 1.097 across the entire pseudochromosome, with an average of 0.851. The unassembled contigs contained 1026 genes at a density of 0.362 per 10Kb.

Discussion

Genetic mapping reveals 15 likely pseudochromosomes

The challenging and highly repetitive nature of the *Phytophthora infestans* genome, presents considerable challenges to the assembly of a *de novo*

Figure 2.5. Chromosome comparisons and annotations. **A**) Gene densities (green, top track) and variant densities for parent 1306 (orange, bottom, uniquely mapping reads) plotted along the assembled pseudochromosomes. Neither feature shows a consistent or uniform distribution along the chromosomes. This is consistent with the "two-speed" genomic architecture common to filamentous pathogens, but the pattern is difficult to discern on the chromosome-scale. Some chromosomes, such as 1, 3, 6, 8, 9, 10, 12, 13, and 15 show a single large relatively gene sparse region near the middle or end of the chromosome. These could represent the centromere-containing regions. **B**) The last Mb of chromosome 10 was extracted to show gene vs repeat densities. Repeated sequences are highly abundant, and the only regions to not show extensive repeats are occupied by 89 genes annotated in this interval. Genes also show local clustering, with stretches of gene deserts stretching over 100kb, consistent with the two-speed genome architecture. These patterns are difficult to observe when viewing the entire chromosome.

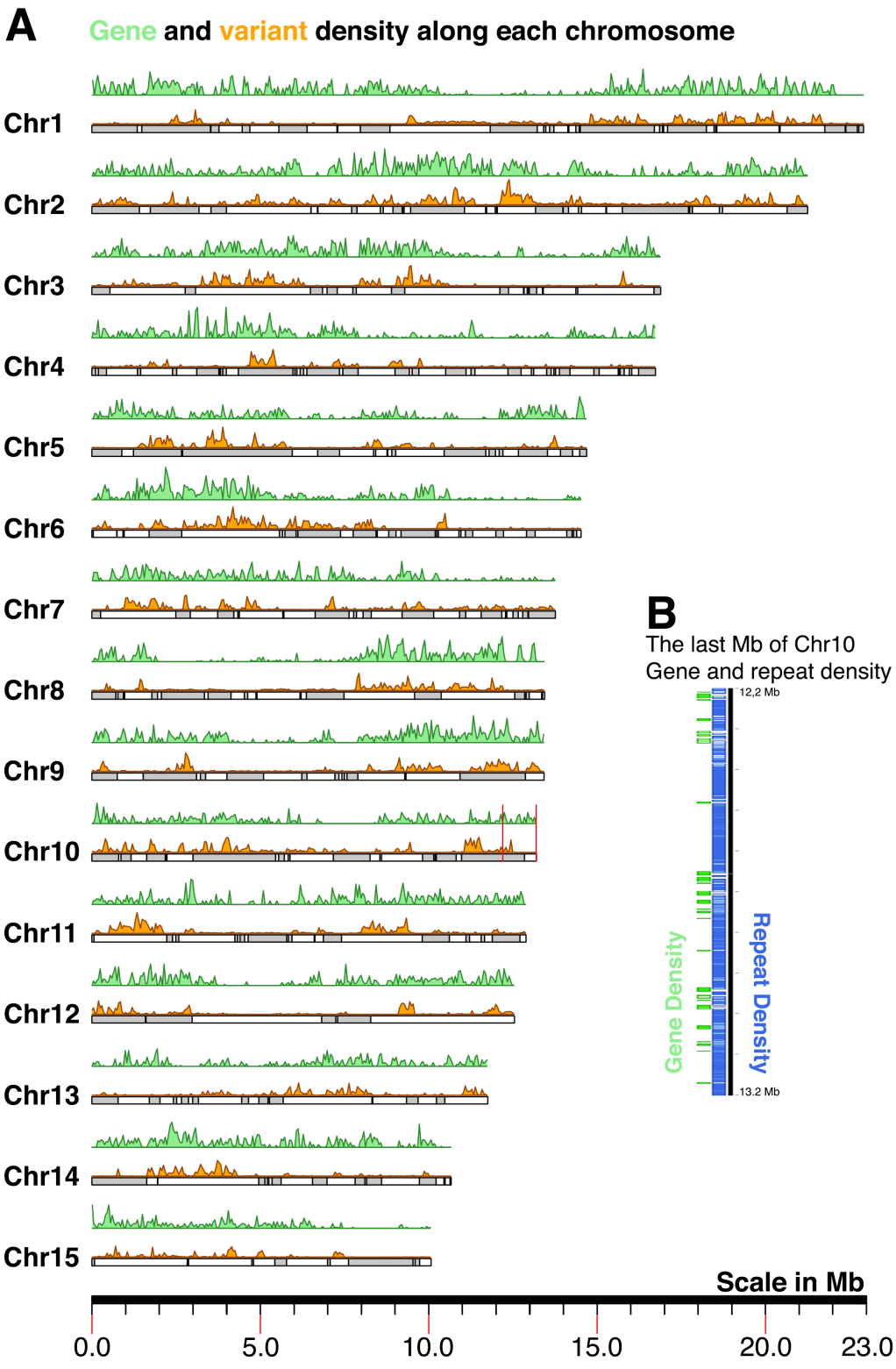

A
Gene and variant density along each chromosome
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Scale in Mb
0.0
5.0
10.0
15.0
20.0
23.0
B
The last Mb of Chr10
Gene and repeat density
12.2 Mb
13.2 Mb
Gene Density
Repeat Density

Table 2.6. Gene count and density along each pseudochromosome assembly.

Chromosome	Gene count	Chromosome size	Average gene density per 10kb
chr1	2,277	22,923,084	0.993
chr2	2,332	21,262,383	1.097
chr3	1,643	16,897,382	0.972
chr4	1,238	16,741,615	0.739
chr5	1,062	14,697,857	0.723
chr6	1,315	14,531,184	0.905
chr7	1,024	13,769,070	0.744
chr8	1,157	13,450,935	0.860
chr9	1,413	13,431,482	1.052
chr10	846	13,206,068	0.641
chr11	1,144	12,896,474	0.887
chr12	1,130	12,552,869	0.900
chr13	753	11,760,846	0.640
chr14	1,033	10,674,240	0.968
chr15	646	10,077,107	0.641
Average	1,267	**14,591,506**	**0.851**
Unassembled	1,026	28,359,575	0.362
Total	**20,039**	**247,232,171**	**-**

reference sequence. We sought to address this through using the newer 3rd generation sequencing and scaffolding technologies of Pacific Biosciences long-read SMRT sequencing, a Dovetail Genomics Chicago library, and a Bionano genomics optical map. Additionally, a pipeline developed by our collaborators to incorporate multiple *de novo* assemblies of the raw PacBio reads with the optical map allowed us to identify chimeric contigs introduced by the three *de novo* assemblers (Abrujin, Canu, Falcon)(**Table 2.1**). Finally, both our Dovetail Chicago library and the optical map aided in spanning repeat-rich regions to produce a superior assembly. However, these efforts were not enough to overcome the seemingly-insurmountable stretches of repeat-rich genome sequence in *Phytophthora infestans* which consist of mostly gypsy and copia elements (Haas et al. 2009).

To achieve a true chromosome-scale assembly, we whole-genome Illumina-sequenced a total of 114 progeny originating from two sexual crosses and constructed genetic maps representing each of the four parent isolates used to construct the sexual populations. To our knowledge, these are the first published results of whole genome resequencing of an oomycete sexual population. Our efforts to construct the genetic maps, which eventually contributed to the assembly of the chromosome-scale reference, were not without issue. Our primary challenge was a low progeny sequencing coverage of 15-35x, which introduced a genotyping error rate of around 1.1% per genotyped

variant (**Table 2.2**). Correcting these errors involved one automated strategy (Genotype-corrector), and four manual curation strategies (**Figure 2.1**), including the identification of aneuploid contigs, removal of markers from potentially misassembled regions of the contigs, manual correction of markers on contigs too small to be analyzed by Genotype-corrector, and removal of individual markers residing more than 30 cM off the end of an established linkage group (LG). These measures, however, were successful in reducing the genome size of parent 1306 from 5,898 cM to 1869 cM, more more in line with previously estimated genome sizes of *P. infestans* of 1,010-1,340 cM with 10 major and seven minor LGs (Van der Lee et al. 1997) 782 cM over 13 major and 10 minor LGs (van der Lee et al. 2004). Additionally, parents 1306, 6629, and 550 all contained individual linkage groups which appear bloated when compared to corresponding linkage groups in other parents (1306: lg2, 6, 6629: lg3+11, lg2+5, 550: lg4+7, lg12+14; **Table 2.5**), likely as a result of a failure to properly correct genotyping errors. This leaves parent 618 with the most compact genetic size at around 1,300 cM, a size closer to the previously reported genetic sizes.

Another challenge originated from the relatively small numbers of progeny for each cross, compared to model systems like *Arabidopsis thaliana* where F2 sample sizes often number in the hundreds (Guo et al. 2012). Our population numbers are similar to those used to generate previous genetic maps of *P. infestans,* however (van der Lee et al. 2004). Nevertheless, smaller progeny

sample sizes lead to less precision in placing and orienting genetic markers, and as such, we only successfully oriented 62.1% of the bases incorporated into our pseudochromosomes. This also contributed to the misplacement of many telomere sequences. Four telomere hits reside on unassembled contigs, while seven reside on assembled pseudochromosomes. Only three of these chromosome telomeres appear properly placed, near the very end of the assembly, and the rest reside up to 3 Mb within the assembly. This highlights firstly that if we assume a true chromosome count of 13-14, we would expect at a minimum 26 telomere sequences in the input assembly, considerably more than the 11 observed here. Second, more robust genetic mapping contributed by a GBS approach would likely resolve the location of telomere hits more toward the ends of the pseudochromosome assembly.

Finally, one unresolved issue is the occurrence of markers from all four parental linkage groups not mapping concurrently along the length of a physical assembly (**Figure 2.4**). This phenomenon is most frequent in the Cross20 parents, though it still occurs in parent 618 in chromosome 8. Importantly, however, sequencing reads and heterozygous variants from these parents do indeed map to the chromosomes which did not display any genetic mapping markers contributing to the pseudochromosome assembly (**Table 2.7**). A scenario explaining the situation underlying chromosome 10, for example, would involve every heterozygous position on chromosome 10 for parent 550 is also

Table 2.7. Variant numbers and variant density for each pseudochromosome and unassembled contigs, per parent.

	1306 All Het[a]	1306 Variants per Mb[b]	618 All Het	618 Variants per Mb	6629 All Het	6629 Variants per Mb	550 All Het	550 Variants per Mb
chr1	25,417	1,109	24,574	1,072	33,637	1,468	21,993	960
chr2	35,342	1,662	29,335	1,380	32,778	1,542	35,319	1,661
chr3	21,068	1,247	14,370	850	28,938	1,712	14,230	842
chr4	13,164	786	9,157	547	10,976	656	10,713	640
chr5	17,696	1,204	20,652	1,405	20,458	1,392	19,874	1,352
chr6	21,758	1,497	16,258	1,119	22,450	1,545	21,084	1,451
chr7	20,356	1,478	9,244	671	14,949	1,086	14,323	1,040
chr8	16,250	1,208	11,101	825	15,495	1,152	14,839	1,103
chr9	17,260	1,285	17,466	1,301	23,587	1,756	17,512	1,304
chr10	21,128	1,599	17,726	1,342	3,968	300	12,438	942
chr11	17,640	1,367	18,291	1,418	22,117	1,714	10,023	777
chr12	11,888	947	21,394	1,705	22,452	1,789	23,938	1,907
chr13	12,160	1,034	18,083	1,538	15,559	1,323	13,028	1,108
chr14	11,842	1,110	13,468	1,262	14,641	1,372	16,688	1,564
chr15	7,936	787	5,132	509	18,942	1,879	7,152	710
Unassembled	14,225	502	6,346	224	13,646	481	11,135	393
Genome total	**275,041**	**1,112**	**247,917**	**1,003**	**314,593**	**1,272**	**264,289**	**1,069**

[a]All heterozygous variants per chromosome per individual, as determined by UnifiedGenotyper
[b]Based on chromosome, unassembled, and genome sizes in Table 3.6

heterozygous for parent 6629, through the reverse is not the case, since 6629 markers map to chromosome 10. Since markers used for mapping were first subsampled every 5 Kb, it is possible that variants heterozygous in parent 550 and homozygous in parent 6629 might have been skipped, though this seems unlikely given the high marker density of our linkage groups and the fact that this pattern repeats itself across multiple chromosomes. Given that parents 6629 and 550 both originate from sexual population in central Mexico, it is possible that some of their chromosomes are related through relatively recent sexual reproduction events, or faced similar selection pressures for individual chromosomes. The latter scenario seems unlikely since the vast majority of heterozygous variants are not expected to have functional consequences and thus are not under any kind of selective pressure. Another scenario could involve improper segregation of the markers due to structural variation between parents 6629 and 550. The non-Mendelian segregation ratios among the progeny would cause the markers to be discarded from genetic mapping, and thus would not contribute to pseudochromosome assembly. Revisiting the genetic maps with a denser concentration of markers may resolve whether or not a biological process underlies this phenomenon, or if it is simply a statistical fluke. Nevertheless, the genetic maps of parents 6629 and 550 do not hurt the overall assembly because they were put at a lower weight when incorporated into the assembly by ALLMAPS, and still provided a limited amount of benefit.

Advantages of our chromosome-scale reference genome

While it is not the first oomycete genome organized into genetic linkage groups (Lamour et al. 2012), the *P. infestans* 238 Mb genome (compared to *P. capsici* at 56 Mb) is still the largest and first described oomycete reference sequence utilizing third generation sequencing and scaffolding technologies. As a result, several new techniques are now possible for the analysis of *P. infestans* using the reference. First, effective QTL mapping is facilitated by having much longer scaffolds than in previously available oomycete genomes. This is due to the ability to map markers that confidently flank QTLs, a feat made more difficult with shorter contig sizes, especially when compounded with small population sizes. While the process of cross generation and analysis in oomycetes presents more challenges than in other model systems like *Arabidopsis thaliana* or *Zea mays*, it will be possible to investigate the contributions of one or more genes underlying certain phenotypic traits such as fungicide resistance.

A second major advantage of using chromosome-sized scaffolds is in identifying copy number variants (CNVs) and large structural variants (SVs). While estimation of copy number through allele ratio examination is possible on a genome-wide scale, identifying either whole or partial chromosome can only be done visually through karyotyping, or through the use of a suitable reference

sequence. Both allele ratio plotting and sequencing read depth tiling can facilitate this.

Our gene annotations reveal where the *Phytophthora* centromeres lie could potentially lie. Centromere regions are often marked by low recombination rates, and low gene density, through little is known about oomycete centromere biology. Observing gene density along the chromosomes (**Figure 2.5**) gives hints of lower gene densities in the center of some chromosomes such as from 11-15 Mb in chr1and 4-6 Mb in chr 12. Closer inspection of the more gene sparse regions within such intervals might reveal centromere-specific repeats.

Finally, comparing the reference sequences of multiple oomycete species is aided by a complete reference genome. It would be interesting to see if the chromosome numbers are conserved across both distantly and closely related species, or how much syntenic regions have diverged along the chromosomes. This analysis would be most appropriate if the reference genomes being compared are all in the chromosome-level state of completion.

Future improvement potential

While the new *P. infestans* reference sequence represents a considerable upgrade over the existing genome, it still has potential for improvement. First and foremost would be to apply a Hi-C based approach from Dovetail or Phase

Genomics, instead of the Dovetail Chicago sequencing library. This is due to the difficulty in obtaining High Molecular Weight (HMW) DNA of sufficient length and quality to resolve the larger gaps of repetitive and difficult to assemble sequence, though improvements with Oxford Nanopore (Oxford, United Kingdom) technology may resolve this. Instead, an intact chromatin based approach such as Dovetail or Phase Genomics (Seattle, WA) Hi-C systems may overcome this shortcoming. Second, employing a technology such a Genotyping By Sequencing (GBS)(Elshire et al. 2011) can fix two additional shortcomings of this new *P. infestans* sequence: the number of progeny sequenced and the depth to which markers can be sequenced. Since we opted for an Illumina whole genome sequencing approach, we had to limit progeny to around 8 samples per lane to achieve a 15-25x coverage for 78 progeny to remain cost-effective. While the additional sequence will prove useful in future studies examining CNVs, the relatively low coverage limited the accuracy of our marker calls, as evidenced by the need for detailed marker curation and genotype imputations. Using a GBS approach would have allowed us to sequence considerably more progeny and observe more recombination events to more accurately incorporate genetic maps into the final contig assembly phase.

The current gene annotations also reveal shortcomings in our assembly. 3,149 of 20,039 genes (15.7%) mapped to at least two genomic loci, suggesting that redundant sequence was incorporated into the final assembly. An

explanation for this could be the divergence of haplotypes in parent 1306, leaving certain contigs with identical sequence at their ends, and diverged sequence within. These contigs would not have been collapsed into one single contig, and the non-repetitive regions within that contain genes now have two mapping positions within the assembly. 1,054 genes also reside on the unassembled contigs not placed in the genetic maps, though it is currently unclear how many of these are unique due to the redundant and repetitive nature of the unassembled sequences. One way to resolve which contigs represent this failure of haplotype consolidation would be to investigate sequencing read depth over duplicated genes. Those that show half the depth of the genome average are likely artificially duplicated. Additionally, a proper gene annotation approach using the Maker pipeline (Cantarel et al. 2008) is in progress, and currently being refined. This will incorporate multiple levels of evidence, including those from other *Phytophthora* species, and will likely result in a much more biologically relevant set of genes by identifying which genes only have spurious mapping information, if they coincide with known repetitive elements, and identifying completely novel genes not present in the previous annotations provided by the Broad institute (Haas et al. 2009).

Future efforts to improve the genome through improved genetic maps may wish to include constructing a completely new cross to take advantage of how ALLMAPS can perform better with additional genetic maps. An ideal setup could

consist of two Mexican isolates from the same geographic region as parent 618 to minimize the structural differences one might expect from more highly-diverged isolates originating outside of Mexico. On the other hand, the continued use of isolate 1306 as one of the cross parents would take advantage of the fact that our underlying assembly is based on that particular isolate through keeping the number of chromosomes contributing to the genetic maps by parent 1306 constant.

The addition of more genetic maps, high confidence genotype calls, and larger progeny numbers may help to clarify the number of major linkage groups/chromosomes. Our current number of pseudochromosome assemblies sits at 15, slightly more than the 10-13 major linkage groups of previous studies ((Carter et al. 1999; Li et al. 2017b; Van der Lee et al. 1997; van der Lee et al. 2004), and eight (Ritch and Daggett 1995) to 13-14 (Li et al. 2017b) chromosomes counted via microscopic approachs. An explanation for this may involve particularly large centromeres in a handful of the larger chromosomes, resulting in recombination profiles among the progeny that are too diverged to link together. However, the notion that recombination rates are typically suppressed within centromere regions would suggest that linkage between markers flanking centromere sequences would be maintained to a relatively high degree. Nevertheless, the existence of chromosomes showing a profile of genetic distance vs physical distance that does not follow the expected curve of a

central suppression of recombination, such as for chromosomes 7, 14, and 15 (**Table 2.4**) suggests that two of these assemblies might actually reside on a single, very large, metacentric chromosome. Whether or not these represent true acrocentric chromosomes, will depend on better genetic mapping data or another scaffolding approach, such as Hi-C.

References

Australian Commonwealth. 2014. Threat abatement plan for disease in natural ecosystems caused by *Phytophthora cinnamomi*.

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., 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., Win, J., Wood, J., Clifton, S. W., Rogers, J., Van den Ackerveken, G., Jones, J. D. G., McDowell, J. M., Beynon, J., and Tyler, B. M. 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. Science 330:1549-1551.

Berlin, K., Koren, S., Chin, C. S., Drake, J. P., Landolin, J. M., and Phillippy, A. M. 2015. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. Nat Biotech 33:623-630.

Browning, S. R., and Browning, B. L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 81:1084-1097.

Burkhardt, A., Buchanan, A., Cumbie, J. S., Savory, E. A., Chang, J. H., and Day, B. 2015. Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. Mol Plan Micro Interac 28:298-309.

Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., and Yandell, M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188-196.

Carter, D. A., Buck, K. W., Archer, S. A., Van der Lee, T., Shattock, R. C., and Shaw, D. S. 1999. The detection of nonhybrid, trisomic, and triploid offspring in sexual progeny of a mating of *Phytophthora infestans*. Fun Genet Biol 26:198-208.

Chin, C. S., Peluso, P., Sedlazeck, F. J., Nattestad, M., Concepcion, G. T., Clum, A., Dunn, C., O'Malley, R., Figueroa-Balderas, R., Morales-Cruz, A., Cramer, G. R., Delledonne, M., Luo, C., Ecker, J. R., Cantu, D., Rank, D. R., and Schatz, M. C. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. Nat Meth 13:1050-1054.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., and Genomes Project Analysis, G. 2011. The variant call format and VCFtools. Bioinformatics 27:2156-2158.

Derbyshire, M., Denton-Giles, M., Hegedus, D., Seifbarghy, S., Rollins, J., van Kan, J., Seidl, M. F., Faino, L., Mbengue, M., Navaud, O., Raffaele, S., Hammond-Kosack, K., Heard, S., and Oliver, R. 2017. The complete genome sequence of the phytopathogenic fungus *Sclerotinia sclerotiorum* reveals insights into the genome architecture of broad host range pathogens. Genome Biol Evol 9:593–618.

Derevnina, L., Chin-Wo-Reyes, S., Martin, F., Wood, K., Froenicke, L., Spring, O., and Michelmore, R. 2015. Genome sequence and architecture of the tobacco downy mildew pathogen *Peronospora tabacina*. Mol Plan Micro Interac 28:1198-1215.

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., and Aiden, E. L. 2017. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. Science 356:92-95.

Dussert, Y., Gouzy, J., Richart-Cervera, S., Mazet, I. D., Deliere, L., Couture, C., Legrand, L., Piron, M. C., Mestre, P., and Delmotte, F. 2016. Draft genome sequence of *Plasmopara viticola*, the grapevine downy mildew pathogen. Genome Announc 4:e00987-00916.

Edger, P. P., VanBuren, R., Colle, M., Poorten, T. J., Wai, C. M., Niederhuth, C. E., Alger, E. I., Ou, S., Acharya, C. B., Wang, J., Callow, P., McKain, M. R., Shi, J., Collier, C., Xiong, Z., Mower, J. P., Slovin, J. P., Hytönen, T., Jiang, N., Childs, K. L., and Knapp, S. J. 2017. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. GigaScience 7:1-7.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

Faino, L., Seidl, M. F., Datema, E., van den Berg, G. C., Janssen, A., Wittenberg, A. H., and Thomma, B. P. 2015. Single-Molecule Real-Time Sequencing combined with optical mapping yields completely finished fungal genome. MBio 6.

Gel, B., and Serra, E. 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33:3088-3090.

Goodwin, S., McPherson, J. D., and McCombie, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17:333-351.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotech 29:644-652.

Guo, Y. L., Todesco, M., Hagmann, J., Das, S., and Weigel, D. 2012. Independent FLC mutations as causes of flowering time variation in *Arabidopsis thaliana* and *Capsella rubella*. Genetics 192:729-739.

Haas, B. 2010. TransposonPSI: an application of PSI-alast to mine (retro-)transposon ORF homologies. http://transposonpsi.sourceforge.net/.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., 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., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B.,

Ren, Q., 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., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Visser, R. G. F., and van der Vossen, E. A. G. 2008. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. Potato Res 51:47-57.

Jiang, R. H., and Tyler, B. M. 2012. Mechanisms and evolution of virulence in oomycetes. Annual review of phytopathology 50:295-318.

Jiang, R. H., de Bruijn, I., Haas, B. J., Belmonte, R., Lobach, L., Christie, J., van den Ackerveken, G., Bottin, A., Bulone, V., Diaz-Moreno, S. M., Dumas, B., Fan, L., Gaulin, E., Govers, F., Grenville-Briggs, L. J., Horner, N. R., Levin, J. Z., Mammella, M., Meijer, H. J., Morris, P., Nusbaum, C., Oome, S., Phillips, A. J., van Rooyen, D., Rzeszutek, E., Saraiva, M., Secombes, C. J., Seidl, M. F., Snel, B., Stassen, J. H., Sykes, S., Tripathy, S., van den Berg, H., Vega-Arreguin, J. C., Wawra, S., Young, S. K., Zeng, Q., Dieguez-Uribeondo, J., Russ, C., Tyler, B. M., and van West, P. 2013. Distinctive expansion of potential virulence genes in the genome of the oomycete fish pathogen *Saprolegnia parasitica*. PLoS Genet 9:e1003272.

Jiao, W. B., and Schneeberger, K. 2017. The impact of third generation genomic technologies on plant genome assembly. Current opinion in plant biology 36:64-70.

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., Campbell, M. S., Stein, J. C., Wei, X., Chin, C.-S., Guill, K., Regulski, M., Kumari, S., Olson, A., Gent, J., Schneider, K. L., Wolfgruber, T. K., May, M. R., Springer, N. M., Antoniou, E., McCombie, W. R., Presting, G. G., McMullen, M., Ross-Ibarra, J., Dawe, R. K., Hastie, A., Rank, D. R., and Ware, D. 2017. Improved maize reference genome with single-molecule technologies. Nature 546:524.

Judelson, H. S. 1997. Expression and inheritance of sexual preference and selfing potential in *Phytophthora infestans*. Fun Genet Biol 21:188-197.

Judelson, H. S., and Michelmore, R. W. 1991. Transient expression of genes in the oomycete *Phytophthora infestans* using *Bremia lactucae* regulatory sequences. Curr Gen 19:453-459.

Kamoun, S., Furzer, O., Jones, J. D., Judelson, H. S., Ali, G. S., Dalio, R. J., Roy, S. G., Schena, L., Zambounis, A., Panabieres, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X. R., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B. M., Grunwald, N. J., Mukhtar, M. S., Tome, D. F., Tor, M., Van Den Ackerveken, G., McDowell, J., Daayf, F., Fry, W. E., Lindqvist-Kreuze, H., Meijer, H. J., Petre, B., Ristaino, J., Yoshida, K., Birch, P. R., and Govers, F. 2015. The Top 10 oomycete pathogens in molecular plant pathology. Mol Plan Path 16:413-434.

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., and Jones, J. D. 2011. Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. PLoS Biol 9:e1001094.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722-736.

Lam, E. T., Hastie, A., Lin, C., Ehrlich, D., Das, S. K., Austin, M. D., Deshpande, P., Cao, H., Nagarajan, N., Xiao, M., and Kwok, P.-Y. 2012. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. Nat Biotech 30:771-776.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S., Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. Mol Plan Micro Interac 25:1350-1360.

Lévesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., Raffaele, S., Robideau, G. P., Thines, M., Win, J., Zerillo, M. M., Beakes, G. W., Boore, J. L., Busam, D., Dumas, B., Ferriera, S., Fuerstenberg, S. I., Gachon, C. M., Gaulin, E., Govers, F., Grenville-Briggs, L., Horner, N., Hostetler, J., Jiang, R. H., Johnson, J., Krajaejun, T., Lin, H., Meijer, H. J.,

Moore, B., Morris, P., Phuntmart, V., Puiu, D., Shetty, J., Stajich, J. E., Tripathy, S., Wawra, S., van West, P., Whitty, B. R., Coutinho, P. M., Henrissat, B., Martin, F., Thomas, P. D., Tyler, B. M., De Vries, R. P., Kamoun, S., Yandell, M., Tisserat, N., and Buell, C. R. 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. Genome Biol 11:R73.

Li, C., Lin, F., An, D., Wang, W., and Huang, R. 2017a. Genome sequencing and assembly by long reads in plants. Genes (Basel) 9:6-6.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.

Li, Y., Shen, H., Zhou, Q., Qian, K., van der Lee, T., and Huang, S. W. 2017b. Changing ploidy as a strategy: the Irish potato famine pathogen shifts ploidy in relation to its sexuality. Mol Plan Micro Interac 30:45-52.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., and Dekker, J. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289-293.

Lin, Y., Yuan, J., Kolmogorov, M., Shen, M. W., Chaisson, M., and Pevzner, P. A. 2016. Assembly of long error-prone reads using de Bruijn graphs. Proc Natl Acad Sci USA 113:E8396-E8405.

Links, M. G., Holub, E., Jiang, R. H., Sharpe, A. G., Hegedus, D., Beynon, E., Sillito, D., Clarke, W. E., Uzuhashi, S., and Borhan, M. H. 2011. De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. BMC Geno 12:503.

Matson, M. E. H., Small, I. M., Fry, W. E., and Judelson, H. S. 2015. Metalaxyl resistance in *Phytophthora infestans*: assessing role of RPA190 gene and diversity within clonal lineages. Phytopathology 105:1594-1600.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297-1303.

Moll, K. M., Zhou, P., Ramaraj, T., Fajardo, D., Devitt, N. P., Sadowsky, M. J., Stupar, R. M., Tiffin, P., Miller, J. R., Young, N. D., Silverstein, K. A. T., and Mudge, J. 2017. Strategies for optimizing Bionano and Dovetail explored through a second reference quality assembly for the legume model, *Medicago truncatula*. BMC Geno 18:578.

Price, A. L., Jones, N. C., and Pevzner, P. A. 2005. De novo identification of repeat families in large genomes. Bioinformatics 21 Suppl 1:i351-358.

Quinn, L., O'Neill, P. A., Harrison, J., Paskiewicz, K. H., McCracken, A. R., Cooke, L. R., Grant, M. R., and Studholme, D. J. 2013. Genome-wide sequencing of *Phytophthora lateralis* reveals genetic variation among isolates from Lawson cypress (*Chamaecyparis lawsoniana*) in Northern Ireland. FEMS Micro Lett 344:179-185.

Ritch, D. L., and Daggett, S. S. 1995. Nuclear DNA content and chromosome-number in German isolates of *Phytophthora infestans*. Mycologia 87:579-581.

Rujirawat, T., Patumcharoenpol, P., Lohnoo, T., Yingyong, W., Lerksuthirat, T., Tangphatsornruang, S., Suriyaphol, P., Grenville-Briggs, L. J., Garg, G., Kittichotirat, W., and Krajaejun, T. 2015. Draft genome sequence of the pathogenic oomycete *Pythium insidiosum* strain pi-s, isolated from a patient with pythiosis. Genome Announc 3:e00574–00515.

Sharma, R., Xia, X., Cano, L. M., Evangelisti, E., Kemen, E., Judelson, H., Oome, S., Sambles, C., van den Hoogen, D. J., Kitner, M., Klein, J., Meijer, H. J. G., Spring, O., Win, K., Zipper, R., Bode, H. B., Govers, F., Kamoun, S., Schornack, S., Studholme, D. J., Van den Ackerveken, G., and Thines, M. 2015. Genome analyses of the sunflower pathogen *Plasmopara halstedii* provide insights into effector evolution in downy mildews and *Phytophthora*. BMC Geno 16:741.

Shelton, J. M., Coleman, M. C., Herndon, N., Lu, N., Lam, E. T., Anantharaman, T., Sheth, P., and Brown, S. J. 2015. Tools and pipelines for Bionano data: molecule assembly pipeline and FASTA super scaffolding tool. BMC Geno 16:734.

Shrivistava, J. 2017. Gene structure and variation in *Phytopthora infestans*: Resources for understanding and managing a global food threat to food security. PhD Thesis, University of California, Riverside.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210-3212.

Tang, H., Zhang, X., Miao, C., Zhang, J., Ming, R., Schnable, J. C., Schnable, P. S., Lyons, E., and Lu, J. 2015. ALLMAPS: robust scaffold ordering based on multiple maps. Genome Biol 16:3.

Tian, M., Win, J., Savory, E., Burkhardt, A., Held, M., Brandizzi, F., and Day, B. 2011. 454 Genome sequencing of *Pseudoperonospora cubensis* reveals effector proteins with a QXLR translocation motif. Mol Plan Micro Interac 24:543-553.

Tyler, B. M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R. H., Aerts, A., Arredondo, F. D., Baxter, L., Bensasson, D., Beynon, J. L., Chapman, J., Damasceno, C. M., Dorrance, A. E., Dou, D., Dickerman, A. W., Dubchak, I. L., Garbelotto, M., Gijzen, M., Gordon, S. G., Govers, F., Grunwald, N. J., Huang, W., Ivors, K. L., Jones, R. W., Kamoun, S., Krampis, K., Lamour, K. H., Lee, M. K., McDonald, W. H., Medina, M., Meijer, H. J., Nordberg, E. K., Maclean, D. J., Ospina-Giraldo, M. D., Morris, P. F., Phuntumart, V., Putnam, N. H., Rash, S., Rose, J. K., Sakihama, Y., Salamov, A. A., Savidor, A., Scheuring, C. F., Smith, B. M., Sobral, B. W., Terry, A., Torto-Alalibo, T. A., Win, J., Xu, Z., Zhang, H., Grigoriev, I. V., Rokhsar, D. S., and Boore, J. L. 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313:1261-1266.

Van der Lee, T., De Witte, I., Drenth, A., Alfonso, C., and Govers, F. 1997. AFLP linkage map of the oomycete *Phytophthora infestans*. Fun Genet Biol 21:278-291.

van der Lee, T., Testa, A., Robold, A., van 't Klooster, J., and Govers, F. 2004. High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics 167:1643-1661.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., and Earl, A. M. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963.

Yin, L., An, Y., Qu, J., Li, X., Zhang, Y., Dry, I., Wu, H., and Lu, J. 2017. Genome sequence of *Plasmopara viticola* and insight into the pathogenic mechanism. Sci Rep 7:46553.

Chapter III

Functional consequences of genome plasticity and transcriptional variation in *Phytophthora infestans* sexual populations

Abstract

The oomycete phytopathogen, *Phytophthora infestans,* causes the devastating late blight diseases of potato and tomato. While much is known about the relationships between the current and historic *P. infestans* global populations and the nature of its molecular plant-pathogen interactions, little is known how genomic structure can contribute to its success as a pathogen. A new reference genome sequence has been made available, however, and is shedding light on aspects of its genome biology not before possible with the previous reference sequence. Here, we used the new assembly to explore the interactions between structural variation, copy number variation, quantitative trait loci, and transcriptomics. In particular, the deletion of an arm of chromosome 3 is strongly associated with resistance to the fungicide metalaxyl. Additionally, many cases of aneuploidy and chromosome truncation can be observed within sexual populations. Overall, our data demonstrate that as with other pathogens, the *P.*

infestans genome is highly adaptable, potentially giving rise to new and more competitive lineages on short timescales.

Introduction

Diseases of plants represent a major threat to global food security. Four major lineages of life contribute to these issues, with representatives including the bacterium *Pseudomonas syringae* (Xin and He 2013), the fungus *Fusarium osysporum* (Gordon 2017), cauliflower mosaic virus (Hohn 2013), and the the oomycete *Phytophthora cinnamoni* (Australia 2014). Despite the importance of these outbreaks, few diseases in the plant world rival the destructiveness of late blight on tomato and potato, as evidenced by the Great Famine in Ireland in the 1840s. Even to this day the causal agent of late blight, the oomycete *Phytophthora infestans,* still causes billions of dollars in economic damages globally (Haverkort et al. 2008).

The oomycetes are a diverse collection of both plant and animal pathogens and are responsible for devastating diseases of both crop and natural species (Jiang and Tyler 2012; reviewed in Kamoun et al. 2015). Oomycetes adopt filamentous lifestyles, about half of all species are pathogenic, and while they have come to resemble true fungi through convergent evolution, they are phylogenetically unrelated to true fungi and reside within the

stramenopile/heterokont kingdom alongside brown algae and diatoms. Of the oomycetes, the genus *Phytophthora* represents a lineage of hemibiotrophic phytopathogens causing especially devastating diseases. Additional examples of challenges imposed by *Phytophthora* include the epidemic of Sudden Oak Death in forests across the western United States from *P. ramorum* and soybean root rot due to *Phytophthora sojae*. Species can vary from those infecting a relatively narrow host range, such as *P. infestans* and *P. sojae,* to those with a broad host range, such as *P. cinnamomi* and *P. parasitica* (reviewed in Kamoun et al. 2015).

Control can consist of preemptive measures such as avoiding transport of infected potato tubers for *P. infestans*, as well as host resistance, as seen in *P. sojae* (Jiang and Tyler 2012) and *P. capsici* (Quesada-Ocampo and Hausbeck 2010). These strategies, however, typically fall short. The thick-walled sexual oospores of *Phytophthora* species can overwinter for several years (Brurberg et al. 2011; Drenth et al. 1995), leaving fields unsuitable for economical crop cultivation. Furthermore, stable host resistance has proved elusive, especially for *P. infestans*. In particular, 9 classical resistance genes have been observed to fail before ever being introduced into commercial potato varieties (Malcolmson and Black 1966). A total of 21 R genes have been identified which have varying levels of success against *P. infestans* (Vleeshouwers et al. 2011), however, their durability remains to be seen.

This lack of success in preemptive and host-mediated control leaves chemical control as by far the most important strategy to combat the oomycetes. Several commercially available fungicides exist in the market that can control oomycetes, either through broad-spectrum modes of action, or through an oomycete-specific manner (Pscheidt and Ocamb 1999). Copper-based agents, such as the Bordeaux Mixture, have been in use since the 19th century, and still have good efficacy against both fungi and oomycetes. The resulting heavy ion pollution, however, has lead to a shift towards other means of chemical control in more recent years. Several classes of fungicides can either prevent or be used in response to epidemics: the quinone outside inhibitors (QoI), phosphonates, carboxylic acid amides, and phenylamides. Each class has costs and benefits, including mode of uptake and transport through plant tissue, as well as predisposition for the development of resistance to the compound. Integrative Pest Management (IPM) strategies (Saville et al. 2012) can prescribe spray schedules of several compounds to minimize outbreaks of several plant diseases and, compounded with local weather predications, and can give growers an informed management option when it comes to disease epidemics.

Among the various chemicals used to control oomycetes, the phenylamide compound metalaxyl has been a preferred choice in both preventative and reactionary roles for *P. infestans* control. Introduced in the late 1970s (Frost and Dowley 1979; Urech et al. 1977), “metalaxyl” itself is a racemic mixture of the

inactive L stereoisomer, and the active R stereoisomer, which is also marketed as Mefenoxam. The compound moves systemically through the plant xylem, and is active only against the oomycetes and not true fungi (Cohen and Coffey 1986). Biochemical studies indicate that it interferes with the transcription of ribosomal RNA, suggesting a method of action against the RNA Polymerase I complex in some way (Davidse et al. 1983; Fisher and Hayes 1984). Despite initial promise, resistance to metalaxyl arose soon after its introduction (Dowley and Osullivan 1981; Reuveni et al. 1980), causing major crop losses in Ireland. The trait of resistance typically acts incompletely-dominantly when observed segregating through sexual population of *P. infestans* (Shattock 1988), indicating a single major locus conferring resistance. Globally, several of these major loci have arisen independently in *P. infestans* (Fabritius et al. 1997; Judelson and Roberts 1999; Lee et al. 1999), as well as in other oomycete species (Bhat et al. 1993; Bower and Coffey 1985; Layton and Kuhn 1988).

In 2014, a mutation was identified in the large subunit of RNA Polymerase I (RPA190) which was strongly associated with resistance to metalaxyl in UK isolates (Randall et al. 2014). This polymorphism, SNP T1145A (AA F382Y), was transformed into a sensitive isolate of *P. infestans* and a slight phenotype rescue was observed, suggesting that this SNP is one of the major loci determining metalaxyl resistance. Furthermore, a recent study suggested multiple pathways of resistance developing through additional polymorphisms in

RPA190 (Chen et al. 2018). Both of these studies, as well as earlier biochemical studies (Davidse et al. 1988; Davidse et al. 1983), strongly implicate RPA190 as being a major locus conferring metalaxyl resistance. However, recent work (chapter 1)(Matson et al. 2015) demonstrated that resistance in an isolate of Mexican origin contained a completely separate major locus of resistance. The RPA190 gene was found to segregate in a manner completely unlinked to resistance through a sexual population generated from the resistant Mexican isolate. While this does not negate the role that RPA190 might play in isolates of European and Chinese origin, it does support earlier studies demonstrating the presence of several major loci conferring resistance to metalaxyl in a broader collection of isolates.

While major loci control the majority of the final phenotype of resistance, several other factors contribute. Minor loci modify the phenotype into being somewhat quantitative (Fabritius et al. 1997; Judelson and Roberts 1999; Lee et al. 1999), and it is thought that genes underlying this variation include transporters or cytochrome P450 genes (Childers et al. 2014; Judelson and Senthil 2006). The differential expression of these genes likely contributes to enhanced or decreased fungicide sensitivity more than sequence variants. These dosage variations could also be induced by differences in copy number. For the major loci, isolates homozygous for resistant alleles have been demonstrated to confer increased resistance over those heterozygous for

resistant alleles (Fabritius et al. 1997; Lee et al. 1999). Presumably, genome instability could confer increased resistance through the addition of an additional allele or the loss of a sensitive allele, exposing the resistant allele.

As a pathogen, genome instability can greatly increase the effectiveness of *P. infestans* isolates and their success in establishing epidemics by constantly facilitating adaptations to new environmental or host challenges (Barchenger et al. 2017; Berman 2016; Judelson and Yang 1998; Li et al. 2017; Patel 2016; Raffaele et al. 2010; Sheltzer and Amon 2011; van der Lee et al. 2001; Whittaker 1991). The sudden oak death pathogen, *Phytophthora ramorum*, undergoes genome instability and acquires aneuploidies when passed through it's host, showing that aneuploidy is a potential stress-induced phenonomenon in the oomycetes (Kasuga et al. 2016; Kasuga et al. 2012). In fungi, a notable example of pathogen genomic plasticity is that of the lineage specific chromosomes in the *Fusarium osysporum* species complex which confer pathogenicity when present (Ma et al. 2010). Varying copy number of the EnCYP51 gene was found to affect pathogenicity in the powdery mildew fungus *Erysiphe necator* (Jones et al. 2014). That study found that that a mutation in EnCYP51 conferring increased resistance to sterol demethylase inhibitor fungicides were also associated with an increase in copy number of the gene itself, further compounding resistance in a dose dependent manner. Finally, major human pathogenic fungi such as *Cryptococcus neoformans* and *Candida*

albicans have been shown to undergo considerable structural variation. Titan cells, from *C. neoformans*, are large polyploid variations of normal cells, showing enhanced resistance to antifungal drugs and often produce aneuploid daughter cells (Gerstein et al. 2015). In *C. albicans,* extra chromosome copies were shown to be readily acquired in response to antifungal drugs (Selmecki et al. 2008; Selmecki et al. 2009). One caveat of these studies and examples is that genome instability was observed when subjecting asexual isolates to various stresses, and little data of instability following sexual reproduction is available.

P. infestans has long been known to show deviation from its native diploid state (Bertier et al. 2013; Carter et al. 1999; Hamed and Gisi 2013; van der Lee et al. 2004). This variation includes whole genome ploidy shifts (Carter et al. 1999; Li et al. 2017; Yoshida et al. 2013), aneuploidy (van der Lee et al. 2004), and combinations of variation through heterokaryon mixtures of nuclei in the mycelia (Catal et al. 2010; Judelson and Yang 1998; Layton and Kuhn 1988). While aneuploidy typically induces severe developmental problems in animals, such as Down syndrome in humans caused by three copies of chromosome 21, and less severe consequences in plants (Guo et al. 1996; Henry et al. 2015; Tan et al. 2015), pathogens can take advantage of sudden shifts in gene dosage to try new combinations of effectors (Dong et al. 2016) or to facilitate sequence divergence within regions of duplication (Li et al. 2017). Achieving the optimal gene dosages is important for plant and animals with defined and long-lasting

generations (Birchler et al. 2016; Birchler and Veitia 2012; Sheltzer and Amon 2011), but filamentous pathogens typically produce billions of individual spores, and thus countless dosage configurations can be tested until a superior variant can outcompete others.

In this study, we analyze whole genome sequencing data from a sexual cross of *P. infestans* to better understand resistance to the fungicide metalaxyl on a genome wide scale. Our previous work identified a major locus of metalaxyl resistance segregating in a sexual population generated between a resistant Mexican isolate, and a sensitive Californian isolate. Using our new chromosome-scale reference assembly, we identify a single broad interval at the end of chromosome 3 highly linked to resistance, encompassing 97 genes and four candidates theorized to be functionally involved in metalaxyl resistance. Furthermore, this interval coincides with a hemizygous deletion observed in a subset of isolates displaying spontaneous gains in metalaxyl resistance. We show that the sensitive allele originating from parent 1306 was deleted, with the remaining resistant allele presumably conferring high levels of resistance.

In addition to examining resistance to metalaxyl segregating in our population, we also sought to examine three other phenotypic traits: pathogenicity on potato tubers, relative sporulation proficiency on artificial media, and overall colony growth rate on artificial media. We hypothesized that any linked intervals might contain genes important to each phenotype such as

effectors allowing for tuber colonization, genes critical for sporulation such as Cdc14 (Fong and Judelson 2003), or loss-of-function mutations in genes impacting growth. However, no significant intervals were observed to be linked to either of these three additional traits.

Our whole genome-sequenced sexual population also allowed us to track Copy Number Variants (CNVs) throughout the progeny, and revealed considerable variation. We focused not only on CrossM, made between parents 1306 and 618, but also 34 members of a second cross, Cross20. Variants in both crosses ranged from deletions or duplications of whole chromosomes to chromosomal arm truncations or rearrangements. Around 31% and 13% of all possible chromosomes in CrossM and Cross 20, respectively, were found in a copy number state greater than or equal to three, indicating that the cross progeny had highly plastic genomes.

Finally, to better elucidate the role of minor loci in the metalaxyl resistance phenotype, we performed RNA-seq on 34 parents and progeny of the cross in which metalaxyl is segregating. A total of 309 differentially expressed genes (DEGs) were identified with regards to growth in the presence or absence of metalaxyl, with sensitive progeny showing much more change in mRNA levels than resistant progeny. 120 DEGs were identified between progeny showing high or low resistance to metalaxyl, though the presence or absence of metalaxyl had much less of an effect. Common Gene Ontology (GO) terms included those

pertaining to RNA polymerase subunit-associated genes, as well as kinases and transporters.

Overall, our diverse sexual populations appear to have undergone considerable structural rearrangement, contributing to their diverse phenotypes. The cause their diverse genome structures is unclear, as few studies have investigated large oomycete sexual populations in a whole genome manner, and the study of epigenetic diversity in oomycete species is similarly limited. Nevertheless, a plausible hypothesis to explain the sudden genomic diversity of the sexual progeny could lie in transposable element (TE) activity. However, the confirmation of this hypothesis is out of the scope of this study.

Materials and Methods

Phenotypic analysis

Progeny originating from the cross described in Matson et al (2015; Chapter 1) were selected for Illumina sequencing. This cross was made between metalaxyl-resistant parent 618 originating from Mexico, and metalaxyl-sensitive parent 1306. Briefly, resistance to metalaxyl was assessed through growing single-oospore and outcrossed progeny on plain Rye-A agar, or media amended with metalaxyl at 0.5 µg/ml or 5.0 µg/ml. Each trial consisted of three 1.5% rye-A agar plates, per growth condition, growing at 18°C in the dark. Three measurements per plate were taken every 2-3 days for up to 14 days, and a

linear regression of the colony radius per measurement day was plotted for each plate to determine the colony growth rate in cm per day. The cutoff for determining resistance was defined as having a radial growth rate of at least 40% on the 0.5 µg/ml metalaxyl amended media compared to the growth rate on the unamended control. An emphasis was placed on selecting progeny at either extreme of the distribution of sensitivities, but progeny with intermediate sensitivities were selected as well.

Four traits other than metalaxyl resistance were measured for the sequenced progeny. First, the overall radial growth rates in cm/day were compiled using measurements from the non-amended media during the metalaxyl resistance trials.

Second, relative sporulation was measured. This involved first subculturing a rye agar plug containing mycelia from the growing edge of an already established culture, onto a new rye agar plate. Each trial consisted of up to 15 progeny, each growing on two duplicate plates, but always contained both parents 1306 and 618. After seven days of growth, four 1 cm diameter circles of agar plus mycelia were punched out from the region immediately adjacent to the inoculation plug. All eight plugs, four from each replicate and with a total surface area of 6.286 cm^2, were placed in a 50 ml Falcon tube containing 10 ml room temperature water, and vortexed at high speed for 30 seconds. Without filtering, samples of the suspended mycelia plus sporangia mixture were placed on a

haemocytometer, and sporangia concentration was determined. The total number of sporangia was extrapolated for the 10 ml tube of eight sample plugs, and an estimate of sporangia per cm^2 was generated. Since sporangia production can vary from trial to trial, and all progeny cannot be tested at once, a relative measure of sporangia production was calculated by comparing the sporangia concentration of each sample to that of parent 1306 for each trial. Data for 70 progeny were obtained in total, and each progeny was included in at least two separate trials. The final phenotype measurement per progeny is an average of the relative sporulation concentrations for each trial they took part in.

Third, the degree of colonization of potato tubers was measured. Russet potatoes were bought from a local supermarket, which were assumed to be relatively genetically uniform. Tubers were sliced to 0.5 cm thickness, sterilized in 1.5% bleach for 10 minutes, washed in sterile water, and placed on raised metal racks in closed containers lined with damp paper towels to maintain humidity. Sporangia were extracted by growing subcultures on rye agar plates as described above, but instead of extracting agar plugs, 10ml of chilled and sterile water was added to the plate. A sterile bent glass rod was used to rub the mycelia mat to dislodge the sporangia, and the water plus sporangia suspension was extracted and transferred to a 50 ml Falcon tube. The tube was placed in a dark chamber, maintained on ice at around 10C, for an hour. Sporangia, regardless if they were empty or not following the release of zoospores, were

counted on a haemocytometer and transferred to a new tube with an adjusted concentration of approximately 20 sporangia per µl. A single 10µl droplet of this suspension (200 sporangia plus zoospores) was subsequently placed in the middle of each sliced potato tuber in the humid container. Each parent/progeny consisted of at least three individual tubers per trial. Scores of infection success after 10 days of incubation at 18C in the dark were as follows: **0**, absolutely no necrosis or visible mycelia growth on the tuber, **1**, minor necrosis at the site of inoculation, but no hyphae, **2**, small necrotic lesion with a 1 cm maximum diameter and hyphae confined to lesion borders, **3**, spreading necrotic lesion is <50% tuber size with hyphae contained or uncontained, **4**, necrotic lesion is >50% tuber size with hyphae contained or uncontained, **4.5**, lesion and mycelia dominate entire tuber with extensive necrosis, **5**, mycelia dominates entire tuber, but necrosis has yet to set in (this phenotype was virtually exclusive to parent 1306). In total, phenotypes for 64 parents and progeny could be obtained, with each progeny having at least two bioreplicates. The final phenotype score was an average of all measurements from all trial.

Fourth, mating type was assessed for each progeny. Parent 1306 is of mating type A1, while parent 618 is mating type A2. Thin strips of agar were extracted each progeny and placed parallel to either parent 1306 or parent 618 on a new rye agar plate. After 7 days of growth, small 1-mm^2 samples of the mycelia interaction zone were taken and subsequently squashed between a

microscope slide and cover glass. A progeny was determined to be A1 if oospores were observed when paired with the A2 parent, 618, and progeny were determined to be A2 if oospores were observed when paired with the A1 parent, 1306. Some progeny showed self-fertility since they showed oospores with both A1 and A2 counterparts. Progeny were scored as A1-self if more oospores were observed when paired with the A2 parent than the A1 parent, and vice versa for progeny scored as A2-self. If oospores were observed in roughly equal numbers when paired with either parent, the individual was scored as fully self-fertile.

Single zoospore isolation

To obtain single zoospore-derived subcultures from progeny showing spontaneous acquisition of enhanced metalaxyl resistance, small samples of mycelia were extracted from the growing edge of both fast growing and slow growing sectors on the same agar plate. Sporangia were induced to release zoospores as described above. The sporangia plus zoospore mixture was then filtered through a sterile 10 µm mesh to retain only the zoospores, a 100 µl sample of the resuspended zoospore suspension was spread across a 0.8% water agar plate, and plates were incubated in the dark for 3 days. Individual germinating zoospores were physically extracted with a sharp sterile probe under a microscope and transferred to a new rye agar plate. Individuals derived from

the original progeny, which successfully established a colony, were thus renamed as “N” for normal, or “F” for fast. The metalaxyl resistance phenotype was re-assed for each pair of N and F individuals to confirm the enhanced resistance phenotype was retained through the single-zoospore process.

Sequencing

In total, 82 progeny plus parents from crossM were whole genome sequenced using Illumina technology to a depth of 15-30x, with ~500bp inserts, and over three rounds. For the first round, libraries for 29 crossM progeny were generated using the Illumina PCR-free library prep kit (Illumina, San Deigo CA) at 2x100 bp at the UCR genomics core. For the second round, PCR-free libraries for 28 additional crossM and all 34 cross20 progeny and parents were constructed and sequenced at Cofactor genomics (St. Louis, MO) at 2x150 bp, and PCR-free libraries for the remaining 25 crossM progeny were constructed and sequenced at the UC-Davis sequencing core at 2x150 bp. Libraries for crossM parents 1306 and 618 were constructed and sequenced at UCR, and an additional 2x300 bp Illumina miSeq library for 1306 was constructed and sequenced at UC-Davis at 100x coverage.

A second sexual population of 32 progeny plus two parents was also Illumina-sequenced. This cross, “Cross20”, was generated previously (Fabritius

and Judelson 1997) between A1 parent 6629 and A2 parent 550, and DNA was retrieved from long term storage at -80C. PCR-free 2x150 bp Illumina libraries were generated at CoFactor Genomics (St. Louis, MO) and sequenced to 20-30x coverage for the progeny, and 40-50x coverage for the two parents.

Alignment, variant calling, and QTL analysis

Sickle (github.com/najoshi/sickle) was used to trim the raw sequencing files using default settings. Trimmed sequencing files were aligned to the new *P. infestans* reference (**Chapter 2**) sequence using bwa-mem (Li and Durbin 2009) with default settings, and variants were called using GATK HaplotypeCaller with default settings (Van der Auwera et al. 2013). Briefly, progeny were called in GVCF mode as a cohort, filtered according to the suggested hard filtering parameters, and indels longer than 6 bp were excluded. Variants contributing to the QTL mapping markers were extracted from the cohort VCF file in a pseudo-testcross manner. For example, markers were either Aa × aa or aa × Aa, representing markers heterozygous in one parent only, respectively. If the total number of heterozygous genotypes for any marker fell below 30% or above 70% of the total calls, the marker was excluded. Since the new reference sequence is derived from parent 1306, it was expected that no homozygous alternative calls would therefore be found, and alternate (618) allele frequencies below 0.18 and

above 0.32 were excluded. The resulting high-quality mapping markers totaled 28,771 for parent 1306, and 53,345 for parent 618.

QTL analysis was performed using a custom script provided by Dr. Shizhong Xu (Xu 2013) and logarithm of odds (LOD) scores plotted with the R tool karyoploteR (Gel and Serra 2017). For the detailed analysis of chromosome 3, markers derived from parent 618 were separated into haplotype phases with Beagle/4.0 (Browning and Browning 2007), miscalled genotypes corrected with the python script Genotype-corrector (https://github.com/freemao/Genotype-corrector), and the haplotypes designated chr3A and chr3B. Each dataset was reintroduced to the QTL mapping script. Since metalaxyl likely affects the function of the RNA Pol I complex, we searched for the subunit locations described by Randall et al (2014) in the new reference genome to see if any resided in the interval linked to resistance.

A subset of non-redundant chr3B markers was used to incorporate the resistance locus into a genetic map using JoinMap/5.0 (Kyazma, Wageningen, Netherlands). This resistance marker was constructed by designating the 25 most resistant progeny as heterozygous, the 25 most sensitive progeny as homozygous, and leaving the intermediately resistant genotypes blank.

Copy number evaluation

Estimates of progeny karyotypes were made by integrating sequencing read depth, allele read ratios of all heterozygous variants, and ratios of variants expected to be heterozygous (discussed below). Normalized read depth on all parents and progeny was done using a script written by the Comai lab at UC-Davis (http://comailab.genomecenter.ucdavis.edu/index.php/Bin-by-sam) (Henry et al. 2015). In addition, analysis was run on four isolates originating from an earlier iteration of the 618×1306 (CrossM)(Fabritius et al. 1997), and 35 additional progeny originating from the more recent iteration of the CrossM (Matson et al. 2015).

To more precisely infer karyotypes, two datasets of allele read ratios were considered for each group of progeny (**Table 3.2**). The first dataset consisted of all heterozygous and biallelic polymorphisms, and was generated by extracting genotypes if they were determined as heterozygous by GATK UnifiedGenotyper (which ended up calling much fewer false-positives than HaplotypeCaller), had a total read depth >8, and an indel size no greater than 6bp. The second dataset were generated by examining homozygous polymorphisms existing between the two cross parents. A SNP/indel may be considered for this analysis if it were AA in 618, but TT in 1306, for example. Considering a CrossM progeny population size of 82, polymorphisms were excluded if the number of heterozygous

genotypes called fell below 70, leaving 38,039 total SNP/indels. The same analysis was applied to the 32 Cross20 progeny, although two sets of variants, one for each parent, could be generated instead of a single dataset for CrossM. This is due to parent 1306 from CrossM being the basis for the reference sequence, and thus no homozygous alternate calls are expected, while each parent of Cross20 is expected to show homozygous alternate genotypes. Variants heterozygous in at least 28 of 32 Cross20 progeny were included in the two datasets, resulting in 38,034 homozygous alternate variants in parent 550 and 40,017 in parent 6629.

Allele read ratios for all datasets were generated by dividing the alternate read depth into the total read depth per variant. For all heterozygous variants originating from genomic regions of copy number 2, allele ratios are expected to follow a single mode, normally distributed around 0.50. Copy number 3 variants follow a bimodal distribution around 0.33 and 0.66, while copy number 4 variants follow a trimodal distribution around 0.25, 0.50, and 0.75, or a unimodal distribution around 0.5. The R packages fitdistrplus (Delignette-Muller and Dutang 2015) and mixtools (Benaglia et al. 2009) modeled the underlying normal distributions within the variant ratio calls, and ggplot2 (Gómez-Rubio 2017) plotted the results. Most individuals show a slight skew toward the 0.0 end of the spectrum, and modeling with the expectation of multiple normal distributions often gives better overall fits than a single distribution. Despite this, read depths

and the likely sexual origins of three out of four parents (Fabritius and Judelson 1997; Fabritius et al. 1997; Judelson and Roberts 1999; Wang et al. 2017) strongly implicate a base genome state of 2, though the multimodal models were retained nonetheless for each individual.

In contrast to the set of all heterozygous variants, the sets of variants expected to be heterozygous should only show a normal distribution around a single mode regardless of ploidy. Variation in copy number will instead shift the modal peak ratio. A region with a copy number of 2, containing one reference allele and one alternate, allele will still show a mode around 0.50. However, variants from a copy number 3 region, with one reference and two alternate alleles, will show a single distribution of ratio around 0.66, while variants with two reference and one alternate alleles will show a single distribution around 0.33. Regions of copy number equaling 4 will show single distributions around 0.25, 0.50, or 0.75, depending on the allele layout. Regions of variants showing extensive loss of heterozygosity (LOH), either through chromosome nondisjunction, truncation, duplication, or copy-neutral LOH will not show normal distributions, and will instead shift to highly skewed ratios past 0.75 or 0.25, or pile up almost exclusively at ratios of 0.0 or 1.0 when only the reference or alternate allele is present, respectively.

Karyotypes were estimated by considering these three pieces of evidence: read depth from the Comai script, all heterozygous variants genome wide, and

the variants expected to be heterozygous. In cases of contradiction between the three criteria, read depth was used as the final determinant of copy number. Karyotypes were plotted in terms of copy number from 1-4. Cases of chromosome truncation, either through duplication or deletion of at least 500 kb of sequence, were designated 1+, 2+, or 3+. For instance, a chromosome present in mostly two copies, but with a major deletion, is designated as 1+, while a chromosome present in three copies, but with a duplication, is designated as 3+. Several cases of complex ploidy were observed (e.g. **Figure 3.7a:** 7.19**, b:** 20-44) and in such cases, the majority of copy number calls determined the overall ploidy.

Further confirmation of copy number and karyotype was achieved through qPCR and High Resolution Melt (HRM) analysis of a region at the end of chr3. Regions were selected with the goal of amplifying homozygous SNPs between the two parents, within genes, and with amplicons only containing a single polymorphism. The IDT PrimerQuest software was used to design primers using default settings (Tm=62°C) and with an amplicon length of between 80-120 bp. Following primer efficiency testing using a standard curve of five serial dilutions, four candidate primers were selected for analysis (**Table 3.4**). qPCR reactions were carried out according to the protocol recommended by the Bio-Rad (Hercules, CA) Precision Melt Supermix (#172-5112), on a Bio-Rad CFX Connect

PCR machine, and using the Bio-Rad CFX Manager and Precision Melt Analysis software (**Figure 3.6**).

RNA-seq

32 progeny and both parents from CrossM were selected to undergo transcriptomic analysis in relation to their underlying predisposition to metalaxyl resistance, as well as their responses to metalaxyl (**Table 3.1**). Small tufts of mycelia were extracted from a culture growing on rye agar and placed into six plates containing 50ml of clarified liquid rye media. Samples were kept in the dark at 18C for six days, after which metalaxyl was added at a final concentration of 0.5 µg/ml to three of the plates. All six plates were shaken at 90 RPM for one minute to ensure adequate dispersal of the fungicide and returned to the dark for 24 more hours.

Tissue isolation consisted of consolidating and vacuum filtering the three metalaxyl treated samples and the three control samples, and subsequently flash freezing in liquid nitrogen. 100 mg of tissue was used to extract total RNA using a Sigma-Aldrich (St. Louis, MO) Spectrum Plant Total RNA Kit (STRN205-1KT). RNA quality was assed through checking on a 1% agarose gel, as well as an Agilent (Santa Clara, CA) 2100 Bioanalyzer at the UCR Genomics Core Facility, and samples with a RNA Integrity Number (RIN) of greater than 7.0 were sent for

library construction. Each of the 32 progeny submissions was a single bioreplicate consisting of with and without metalaxyl conditions, while the each parent had two bioreplicates, resulting in a grand total of 72 libraries to be sequenced.

Libraries were prepared and Illumina-sequenced at Cofactor Genomics (St. Louis, MO) to an average of 33.6 million 1x75-nt reads. Since the newer version of the *P. infestans* chromosome scale reference assembly was not available at the time, reads were aligned to the Broad Institute T30-4 reference assembly using the SystemPipeR (TW and Girke 2016) workflow. The gene annotations used were the result of improved annotations conducted by the Judelson Lab (Shrivastava, 2017), and counts per million (CPM) values for each gene were determined with edgeR (Robinson et al. 2010). The phenotype classes used for comparison were divided into High resistance if the % growth rate on 0.5 µg/ml media was greater than 70% of unamended media, Low resistance if the % growth rate was lower than 40%, and Medium for all else in between on rye-B agar. Differentially expressed genes (DEGs) were sought for several comparisons: 1) all resistant individuals vs. all sensitive individuals, 2) all individuals with metalaxyl vs. all individuals without metalaxyl, 3) sensitive individuals with metalaxyl vs. sensitive without metalaxyl, 4), sensitive with metalaxyl vs. resistant with metalaxyl, 5), resistant with metalaxyl vs. resistant without metalaxyl, and 6), resistant without metalaxyl vs. sensitive without

metalaxyl (**Table 3.7**). edgeR was initially used to estimate DEGs though a generalized linear model, but problems with the output forced us to turn to a one way ANOVA model through the Partek Genomics Suite (St. Louis, MO) in which edgeR counts per million CPM values were used as an input. Genes showing raw CPM values of < 1.0 for at least half of the 72 samples were excluded from analysis, leaving 12,791 out of 18,693 total genes.

Results

QTL analysis of phenotypes

The four main phenotypes of the 82 Illumina-sequenced CrossM progeny which were are shown in **Figure/Table 3.1**. Metalaxyl resistance (**Figure 3.1a**) shows a linear distribution of responses when considering the entire population (Matson et al. 2015), but a bias can be seen in the data presented since more resistant and sensitive progeny were selected for sequencing compared to intermediately resistant progeny. Radial growth rates in the absence of metalaxyl are only somewhat distributed broadly (**Figure 3.1b**), and both relative sporulation percentages and pathogenicity on potato tubers show considerable diversity among the progeny (**Figure 3.1c,d**). Metalaxyl and pathogenicity typically showed phenotypes among the progeny intermediate of either parent,

Figure 3.1. Graphs of the four main phenotypes measured for the CrossM progeny. Only the tuber infection phenotype is not measured on 1.5% rye agar media. Parent 618's phenotypes are always shown in red, while parent 1306 is blue. **A**) Metalaxyl resistance, measured as the % radial growth rate in cm/day on media amended with 0.5 µg/ml metalaxyl compared to radial growth rate on media unamended with metalaxyl. Error bars are Standard Error of the Mean (SEM). **B**) Radial growth in cm/day in the absence of metalaxyl with error bars representing SEM. **C**) Sporulation proficiency of progeny, measured after seven days as the number of spores in 1.0 cm^2 compared to parent 1306, per experimental batch, with standard deviation. **D**) Infection scores with SEM of potato tuber colonization.

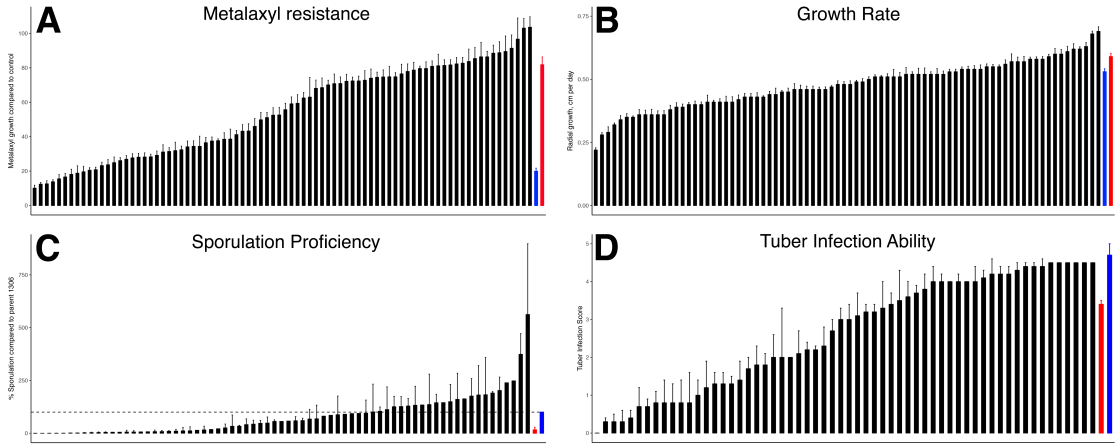

A
Metalaxyl resistance
Metalaxyl growth compared to control
B
Growth Rate
Radial growth, cm per day
C
Sporulation Proficiency
% Sporulation compared to parent 1306
D
Tuber Infection Ability
Tuber Infection Score

Table 3.1. Phenotypic summary of all sequenced CrossM progeny.

Individual	Mex Phenotype[a]	Growth Phenotype[b]	Sporulation Phenotype[c]	Pathogenicity Phenotype[d]	Approx Ploidy[e]	Mating Type[f]	RNA-Seq[g]
1306	19.9 (± 1.8)	0.53 (± 0.012)	100.0 (± 0.0)	4.7 (± 0.3)	2	A1	Yes
618	81.8 (± 4.5)	0.59 (± 0.013)	16.0 (± 12.4)	3.4 (± 0.3)	2	A2	Yes
5.36	12.6 (± 1.8)	0.39 (± 0.012)	135.8 (± 144.1)	0.7 (± 0.2)	3	NA	Yes
7.13	86.4 (± 8.4)	0.36 (± 0.015)	32.8 (± 53.6)	0.3 (± 0.3)	3	NA	Yes
7.19	12.4 (± 1.0)	0.44 (± 0.012)	144.5 (± 40.1)	1.4 (± 0.5)	4	NA	Yes
7.38	88.8 (± 6.4)	0.39 (± 0.017)	4.4 (± 2.2)	NA	3	NA	Yes
M1	10.1 (± 1.7)	0.35 (± 0.014)	112 (± 107.4)	1.3 (± 0.2)	3	A1-self	no
M2	29.2 (± 2.5)	0.46 (± 0.023)	239.0 (± NA)	NA	2	A1	no
M3	18.1 (± 3.0)	0.44 (± 0.025)	182.3 (± 176.2)	2.2 (± 0.2)	2	A1	no
M4	20.8 (± 1.5)	0.35 (± 0.005)	202.6 (± 63.3)	2.0 (± 0.0)	2	A1	Yes
M5	74.5 (± 3)	0.43 (± 0.005)	176.0 (± 83.0)	1.7 (± 0.3)	3	A1-self	no
M7	76.5 (± 5.6)	0.46 (± 0.013)	125.6 (± 49.3)	2 (± 0.6)	2	A1-self	yes
M8	41.2 (± 2.4)	0.32 (± 0.007)	7.8 (± 5.7)	0.7 (± 0.5)	3	A1	Yes
M11	91.4 (± 7.7)	0.51 (± 0.029)	25.3 (± 18.3)	2.1 (± 0.6)	2	A1	yes
M12	74.8 (± 2.4)	0.52 (± 0.013)	0.0 (± 0.0)	NA	2	A1	no
M13	26.9 (± 2.1)	0.63 (± 0.016)	7.5 (± 7.5)	1.8 (± 0.5)	2	A1-self	no
M14	37.5 (± 2.3)	0.53 (± 0.010)	2.7 (± 2.3)	0.4 (± 0.2)	2	A1	no
M15	52.6 (± 4.4)	0.62 (± 0.021)	128.9 (± 35.7)	3.1 (± 0.6)	2	A1	Yes
M18	31.9 (± 4.9)	0.22 (± 0.009)	9.2 (± 3)	0.0 (± 0.0)	2	A1	Yes
M20	20.5 (± 1.5)	0.55 (± 0.011)	4.8 (± NA)	1.3 (± 0.3)	2	A1	no
M21	31.4 (± 2.2)	0.45 (± 0.006)	NA	NA	3	A2-self	no
M24	28.2 (± 2.4)	0.46 (± 0.017)	5.6 (± 8.6)	0.8 (± 0.3)	3	A1	no
M25	46.0 (± 4.6)	0.42 (± 0.018)	0.1 (± 0.1)	0.8 (± 0.5)	4	A1	Yes
M26	43.2 (± 3.8)	0.36 (± 0.021)	48.1 (± 28.7)	0.8 (± 0.6)	4	A1	Yes
M30	70.8 (± 5.6)	0.57 (± 0.031)	0.0 (± NA)	0.3 (± 0.2)	2	A1	no
M32	16.6 (± 2.1)	0.46 (± 0.013)	126.5 (± 47.2)	0.8 (± 0.8)	2	A1	no
M35	68.5 (± 5.6)	0.50 (± 0.010)	NA	4.5 (± 0.0)	2	A1	no
M36	19.6 (± 3.2)	0.43 (± 0.014)	133.9 (± NA)	2.7 (± 0.3)	3	A1	no
M39	52.5 (± 4.3)	0.38 (± 0.017)	94.5 (± 62.6)	1.8 (± 0.3)	2	A2-self	Yes
M40	24.9 (± 3.3)	0.41 (± 0.027)	59.9 (± 11.2)	1.3 (± 0.3)	3	A1	no
M42	72.4 (± 2.2)	0.55 (± 0.009)	0.0 (± 0.0)	NA	2	A1	no
M43	31.1 (± 4.2)	0.41 (± 0.021)	14.0 (± NA)	NA	2	A1	no
M44	34.4 (± 3.2)	0.28 (± 0.009)	5.0 (± NA)	NA	2	A1	no
M45	43.3 (± 4.2)	0.36 (± 0.016)	7.0 (± NA)	NA	2	A1-self	Yes
M46	79.6 (± 3.9)	0.52 (± 0.024)	57.8 (± NA)	4.3 (± 0.2)	4	Self	no
M47	82.7 (± 3.4)	0.54 (± 0.014)	NA	NA	2	A1	no
M53	38.6 (± 5.8)	0.36 (± 0.020)	11.4 (± 19.3)	0.8 (± 0.6)	2	A1	Yes
M59	26.1 (± 1.7)	0.47 (± 0.005)	NA	NA	2	A1-self	no
M60	15.5 (± 2.5)	0.34 (± 0.016)	2.0 (± 1.7)	1.2 (± 0.7)	2	A1	Yes
M62	28.3 (± 1.5)	0.40 (± 0.008)	93.0 (± NA)	NA	2	A1	no
M63	88.5 (± 4.7)	0.52 (± 0.010)	190.5 (± 7.0)	4.4 (± 0.2)	2	A1-self	Yes
M65	70.9 (± 3.8)	0.69 (± 0.018)	248.1 (± NA)	4.5 (± 0.0)	2	A1	Yes
M66	85.4 (± 6.4)	0.57 (± 0.018)	5.8 (± 6.3)	4.5 (± 0.0)	2	A1	Yes
M70	70.1 (± 2.7)	0.48 (± 0.011)	15.7 (± 17.8)	4.2 (± 0.2)	2	A2	Yes
M71	80.9 (± 3.1)	0.55 (± 0.007)	NA	4.0 (± 0.0)	2	A1	no
M72	23.7 (± 3.0)	0.51 (± 0.005)	17.6 (± 0.4)	4.0 (± 0.0)	2	A2	no
M75	74.7 (± 4.1)	0.51 (± 0.012)	NA	3.7 (± 0.2)	3	A1	no
M76	81.7 (± 3.0)	0.53 (± 0.010)	373.8 (± 97.8)	4.0 (± 0.2)	2	A2	Yes
M77	23.2 (± 2.0)	0.41 (± 0.007)	132.6 (± 89.8)	4.4 (± 0.1)	2	A1	no
M82	81.2 (± 6.7)	0.57 (± 0.021)	163.0 (± NA)	NA	2	A2	no
M83	34.4 (± 5.9)	0.48 (± 0.014)	8.01 (± NA)	NA	2	A2	no
M84	81.4 (± 3.3)	0.46 (± 0.009)	93.8 (± NA)	3.8 (± 0.4)	2	A2	Yes
M86	72.2 (± 4.0)	0.54 (± 0.017)	6.4 (± NA)	4.2 (± 0.4)	2	A1	Yes
M88	89.5 (± 9.0)	0.41 (± 0.013)	21.1 (± NA)	3.2 (± 0.2)	2	A1-self	Yes
M89	63 (± 11.4)	0.61 (± 0.021)	1.3 (± 0.1)	2.2 (± 0.1)	2	A1-self	no
M92	72.4 (± 2.8)	0.58 (± 0.009)	90.0 (± NA)	NA	2	A1	no
M93	79.6 (± 1.4)	0.60 (± 0.018)	85.0 (± NA)	NA	2	A1	no
M95	59.4 (± 5.0)	0.41 (± 0.023)	NA	NA	2	A1	no
M99	36.5 (± 3.1)	0.43 (± 0.017)	145.0 (± NA)	NA	2	A1	no
M100	62.5 (± 3.2)	0.54 (± 0.012)	58.9 (± 19.7)	4.5 (± 0.0)	2	A1	Yes
M101	59.1 (± 4.0)	0.51 (± 0.008)	160.4 (± 123.9)	4.0 (± 0.2)	2	A1	Yes
M106	51.1 (± 3.0)	0.58 (± 0.008)	NA	NA	2	A1-self	no
M107	78.7 (± 4.5)	0.59 (± 0.009)	NA	NA	2	A1	no
M108	74.8 (± 6)	0.52 (± 0.027)	9.9 (± 8.3)	NA	2	A1	Yes
M112	55.7 (± 3.7)	0.49 (± 0.013)	NA	4.2 (± 0.2)	2	A1	no
M120	86.4 (± 3.2)	0.46 (± 0.010)	NA	3.3 (± 0.7)	3	A1	no
M123	83.7 (± 7.2)	0.52 (± 0.023)	NA	NA	2	A1	no
M126	34.1 (± 3.4)	0.54 (± 0.009)	47.1 (± 14.2)	4.5 (± 0.0)	2	A1-self	no
M127	13.8 (± 1.2)	0.43 (± 0.014)	43.1 (± 18.6)	4.1 (± 0.2)	2	Self	no
M131	37.7 (± 1.1)	0.49 (± 0.006)	33.4 (± 4.1)	2 (± 1.3)	2	A1	no
M132	68.1 (± 4.8)	0.60 (± 0.022)	11.2 (± 15.8)	4.5 (± 0.0)	2	A1	Yes
M133	28.1 (± 2.2)	0.58 (± 0.010)	149.4 (± 82.0)	4.4 (± 0.1)	2	A1	Yes
M135	27.7 (± 2.4)	0.62 (± 0.009)	NA	NA	2	A1	no
M136	82.3 (± 3.7)	0.52 (± 0.010)	0.2 (± 0.0)	0.3 (± 0.1)	2	A1-self	no
M137	18.8 (± 4.3)	0.51 (± 0.017)	104.0 (± 20.3)	3.5 (± 0.8)	2	A1	Yes
M139	38.5 (± 3.8)	0.29 (± 0.022)	181.2 (± 139.3)	4.0 (± 0.4)	2	A1	Yes
M64N	72.8 (± 5.5)	0.40 (± 0.010)	68.4 (± 64.8)	3.4 (± 0.1)	2	A1	no
M64F	103.1 (± 5.6)	0.48 (± 0.014)	562.2 (± 335.0)	3.6 (± 0.4)	2	NA	no
M78N	49.8 (± 4.3)	0.52 (± 0.025)	87.6 (± 88.4)	3.0 (± 0.4)	2	A2	no
M78F	96.7 (± 12.2)	0.36 (± 0.018)	100.9 (± 131.7)	1.0 (± 0.4)	2	NA	no
M91N	32.4 (± 2.1)	0.68 (± 0.011)	67.3 (± 45.8)	4.0 (± 0.4)	2	A1	Yes
M91F	74.0 (± 5.3)	0.45 (± 0.015)	55.9 (± 7.5)	2.3 (± 0.5)	2	NA	no
M123N	77.9 (± 4.4)	0.56 (± 0.017)	3.4 (± 2.8)	3.2 (± 0.2)	2	A1	no
M123F	103.6 (± 6.1)	0.40 (± 0.014)	1.1 (± 0.4)	3.0 (± 0.3)	2	NA	no

[a]Percent growth rate on 0.5 µg/ml metalaxyl amended media divided by rate on non amended media (Std Error)
[b]Radial growth rate in cm per day (Sted Error)
[c]Percent of number of sporangia counted compared to parent 1306 (Std Dev)
[d]Potato tuber colonization score (Std Error)
[e]Approximate ploidy
[f]Mating type
[g]Was the individual included in the RNA-seq data?

Table 3.2. Variant numbers of selected individuals per chromosome

	1306 All Het	618 All Het[a]	M64N All Het	M64F All Het	M101 All Het	618 Hom Alt[bd] 1306 Hom Ref[cd]	6629 All Het	550 All Het	20-16 All Het	20-23 All Het	6629 Hom Alt 550 Hom Ref	550 Hom Alt 6629 Hom Ref
chr1	25417	24574	17651	17300	16388	4051	33637	21993	18004	20113	2734	2928
chr2	35342	29335	14019	13769	16501	3135	32778	35319	22442	22095	1778	2383
chr3	21068	14370	12765	11009	12272	3194	28938	14230	15273	17300	2069	2603
chr4	13164	9157	3672	3615	5590	251	10976	10713	6537	8518	1655	746
chr5	17696	20652	9098	8829	8246	2265	20458	19874	12805	12917	1646	2105
chr6	21758	16258	11775	11636	12307	2746	22450	21084	13305	16416	1927	2057
chr7	20356	9244	497	9619	9484	2144	14949	14323	11226	2033	1	21
chr8	16250	11101	9694	9589	8243	2440	15495	14839	9663	5830	2269	918
chr9	17260	17466	9447	9163	12471	2095	23587	17512	14438	9673	2258	3240
chr10	21128	17726	8481	8227	10854	1137	3968	12438	14438	20128	11483	9370
chr11	17640	18291	6576	6560	10161	946	22117	10023	18971	12406	3491	3153
chr12	11888	21394	10868	10519	11219	2884	22452	23938	14083	18701	4865	4705
chr13	12160	18083	11811	11836	5733	1157	15559	13028	15550	9486	2361	2740
chr14	11842	13468	7943	7629	8218	3497	14641	16688	9665	12074	1479	1065
chr15	7936	5132	8926	8844	9986	6098	18942	7152	8780	1344	1	0
Total	**270905**	**246251**	**143223**	**148144**	**157673**	**38040**	**300947**	**253154**	**205180**	**189034**	**40017**	**38034**

[a]All heterozygous variants per chromosome per individual, as determined by UnifiedGenotyper
[b]Homozygous-alternate called variants
[c]Homozygous-reference called variants
[d]These variants are expected to be heterozygous for all progeny of the cross

while both parents grew better than many progeny and many progeny sporulated considerably better than parent 1306. The correlation between sporulation and growth rate was calculated in Microsoft Excel (Redmond, WA), and showed little to no relation with a Pearson's correlation coefficient of 0.038.

A total of 55,345 and 28,771 pseudo-testcross markers were generated for parents 618 and 1306, respectively. QTL analysis applied to metalaxyl resistance, using the mapping markers derived from parent 618, revealed a large interval of linkage on the end of chromosome 3 was observed, peaking at a LOD of 11.3 in the raw (**Figure 3.2a)** and cleaned up markers (**Figure 3.2c,d,e**). The interval appears to peak starting around 15.75 Mb and continues until the very end of the chromosome assembly, though a potential misassembly from 16.725-16.8 Mb is revealed by a subset of unlinked markers. A total of 130 genes lie within the region of 15.75 Mb to the end, 97 of which have annotations other than "conserved hypothetical protein" (**Table 3.5**). A somewhat minor peak of linkage can be observed on chromosome 13 corresponding to a maximum LOD of 4.6 and several high-LOD SNPs appear to be scattered around the genome, through these could easily be attributed to noise. None of the genes encoding the 12 subunits comprising RNA Polymerase I are observed within the major observed interval (**Table 3.3**). When the mapping markers originating from parent 1306 were applied, no significant peaks of linkage were observed, consistent with the segregating trait originating from the resistant parent 618. Furthermore,

Figure 3.2. Genome-wide analysis of markers linked to metalaxyl resistance for all 82 sequenced progeny. **A**) Markers linked to the haplotypes of parent 618. Only chromosome 3 has a significant peak of genetic linkage, while chromosome 13 has a much more modest peak. **B**) Markers linked to the haplotypes of parent 1306 which do not show significant linkage to metalaxyl resistance. **C-E**) Detail of both haplotypes of chromosome 3 parent 618 after having miscalled markers cleaned up. **C**) Both cleaned haplotypes of parent 618. **D**) Chromosome 3A haplotype, showing a drop off of association after ~15 Mb. **E**) Chromosome 3B haplotype, showing an increase in linkage after position 15 Mb.

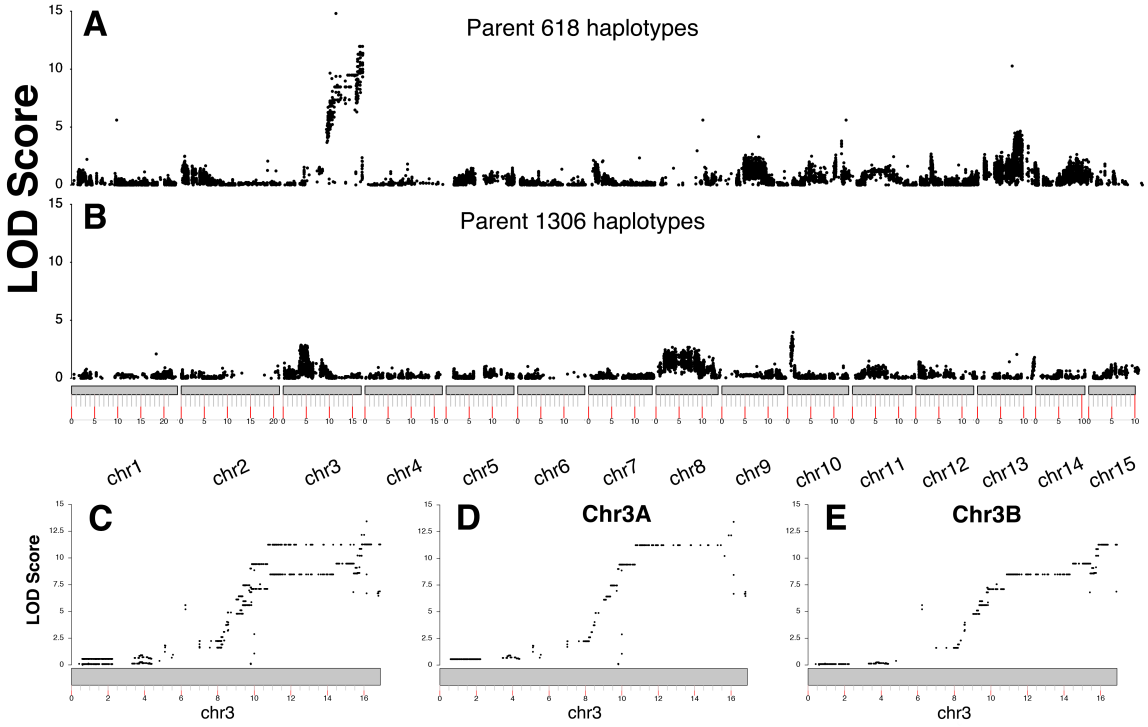

LOD Score
A
Parent 618 haplotypes
B
Parent 1306 haplotypes
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
C
D
Chr3A
E
Chr3B
LOD Score
chr3
chr3
chr3

Table 3.3, Mapping locations of all RNA Polymerase I subunits

Common Name	PITG Number	Chrom	% Identity to T30-4[a]	Alignment Length	Mismatches to T30-4	Chromosome Start	Notes
RPA190	PITG_03855	chr2	99.513	5540	27	8767859	
RPA135	PITG_02040	chr9	99.565	3681	12	11529754	
RPA49	PITG_11365	chr15	100	1486	0	2973073	
RPAC40	PITG_18777	chr11	99.915	1179	1	12747139	*
RPA43	PITG_14613	chr11	99.313	1019	6	2938333	
RPABC14.5	PITG_09712	chr11	99.647	849	1	8435953	
RPABC27	PITG_010445	chr4	99.587	726	3	6946527	
RPABC27	PITG_010445	chr4	99.449	726	4	5856950	
RPABC23	PITG_12877	chr11	98.088	523	10	1731857	
RPAC19	PITG_05854	chr7	99.609	512	2	2538792	**
RPAC19	PITG_05854	chr7	99.609	512	2	3066856	
RPA12.2	PITG_06706	chr1	98.652	445	6	16247662	
RPABC10a	predicted	chr1	100	349	0	7133316	
RPABC10a	predicted	chr1	98.567	349	0	1246944	
RPABC10b	predicted	chr11	100	244	0	743451	
RRN3	PITG_15566	chr15	99.298	2420	8	5931352	

[a]% Genomic DNA identity

* Originally predicted as three paralogs in T30-4

** Three partial hits on chr7 immediately adjacent to full length hits

Table 3.4, Summary of qPCR and HRM markers.

	LF1	LF2	CNV_chr3
Gene	PITG_10087	PITG_01520	PITG_16115
Location	chr3: 6,088,070	chr3: 11,071,360	chr3: 16,183,778
Size	98 bp amplicon	134 bp amplicon	89 bp amplicon
Type[a]	C/A, type II	C/T, type I	C/T, type I
1306 Genotype	AA	TT	CC
618 Genotype	CC	CC	TT
Forward Seq	ACGCCAAGTCAAGCAAAGA	ATCGCATCGTGGGTAGTATTG	AGCTGACTTCAGATGAGCAAG
Reverse Seq	CTCTTCGTCTTCGGGCTTTC	CGAACTCCCATTCTCCCTTAAA	TGTGGGTTCTACTGGGTTAGA

[a]Class of SNP for HRM. Type I is the easiest to genotype, while type IV is the most difficult

Table 3.5, List and annotation of all genes in the chromosome 3 deletion interval defined by fast-growing isolate M64F

Name	Annotations	Start
PITG_16783	Hect Domain Ubiquitin-Protein Ligase	14,532,246
PITG_16788	Conserved Hypothetical Protein	14,579,417
PITG_15987	Conserved Hypothetical Protein	14,581,474
PITG_16789	Conserved Hypothetical Protein	14,582,780
PITG_16791	Conserved Hypothetical Protein	14,616,058
PITG_16792	Nucleolar Gtp-Binding Protein 2	14,618,343
PITG_16794	Chitinase; Di-N-Acetylchitobiase	14,752,964
PITG_16795	Dynein Assembly Factor 3, Axonemal	14,754,723
PITG_16797	Tubulin-Specific Chaperone D	14,831,437
PITG_16798	Phospholipase D, Putative	14,835,948
PITG_16799	Inositol Polyphosphate 4-Phosphatase Type II	14,838,714
PITG_16800	Programmed Cell Death Protein 2-Like	14,842,487
PITG_16801	Protein Hira	14,863,268
PITG_16803	Mkiaa1688 Protein	14,876,721
PITG_16804	Conserved Hypothetical Protein	14,877,999
PITG_16806	4'-Phosphopantetheinyl Transferase	15,052,690
PITG_16807	3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase	15,053,423
PITG_16808	Mitochondrial ABC Transporter 3	15,209,270
PITG_16809	Oxysterol-Binding Protein-Related	15,212,730
PITG_16810	Conserved Hypothetical Protein	15,215,464
PITG_22468	Conserved Hypothetical Protein	15,216,999
PITG_16812	Anaphase-Promoting Complex 10	15,235,637
PITG_16813	Serine/Threonine-Protein Kinase Rio1	15,236,899
PITG_16814	Conserved Hypothetical Protein	15,238,432
PITG_15980	Glucan 1,3-Beta-Glucosidase Putative	15,305,517
PITG_15981	Interphotoreceptor Retinoid-Binding Protein	15,327,215
PITG_15982	Histone-Arginine Methyltransferase Carmer-Related	15,328,962
PITG_15985	Conserved Hypothetical Protein	15,337,807
PITG_21619	Homoserine O-Acetyltransferase	15,415,187
PITG_15993	Homoserine O-Acetyltransferase	15,415,730
PITG_15994	Conserved Hypothetical Protein	15,416,623
PITG_12314	Voltage-Gated Ion Channel (Vic) Superfamily, Putative	15,416,799
PITG_15996	Conserved Hypothetical Protein	15,439,764
PITG_15998	Phospholipase A-2-Activating Protein	15,461,344
PITG_15999	Conserved Hypothetical Protein	15,464,528
PITG_16000	Conserved Hypothetical Protein	15,466,399
PITG_16001	Regulator Of Chromosome Condensation	15,467,187
PITG_16002	Leucine-Rich Repeat-Containing Protein	15,471,556
PITG_16003	Chaperonin; T-Complex Protein 1 Subunit Gamma	15,476,248
PITG_16004	Conserved Hypothetical Protein	15,478,437
PITG_16005	Methylmalonate-Semialdehyde Dehydrogenase [Acylating], Mitochondrial	15,480,940
PITG_16006	Gle-1-Related	15,483,130
PITG_16007	Conserved Hypothetical Protein	15,489,406
PITG_16008	40S Ribosomal Protein S27	15,511,952
PITG_16009	Importin Subunit Alpha	15,512,890
PITG_16010	E3 Ubiquitin-Protein Ligase Rfwd3	15,516,273
PITG_16012	Sr-Related Ctd Associated Factor 6	15,530,037
PITG_16013	Regulator Of Chromosome Condensation	15,531,295
PITG_16014	Conserved Hypothetical Protein	15,537,578
PITG_16015	Ribonuclease T2 ;Ribonuclease T2	15,582,791
PITG_01495	Ribonuclease T2 ;Ribonuclease T2	15,583,079
PITG_16016	Cytochrome C Oxidase Assembly Protein Cox15	15,584,270
PITG_16017	Deoxyribose-Phosphate Aldolase	15,586,088
PITG_16018	Deoxyribose-Phosphate Aldolase	15,589,975
PITG_16020	Isopenicillin N Epimerase	15,613,447
PITG_16021	Sh3 Domain Ysc84-Like Protein 1	15,615,305
PITG_16022	Transmembrane Protein 116	15,616,701
PITG_16023	Transcription Initiation Factor TFIID	15,617,917
PITG_16024	Riboflavin Synthase	15,620,812
PITG_16026	Regucalcin	15,645,123
PITG_21037	Conserved Hypothetical Protein	15,645,805
PITG_16031	Fyve-Finger-Containing Rab5 Effector Protein	15,662,941
PITG_16033	Conserved Hypothetical Protein	15,693,266
PITG_16034	Conserved Hypothetical Protein	15,695,167
PITG_23065	General Transcription Factor Iih Subunit 2	15,732,846
PITG_23066	Probable Pip5K	15,734,322
PITG_16036	Wd-40 Repeat Protein	15,737,674
PITG_16037	Conserved Hypothetical Protein	15,746,153
PITG_16038	Conserved Hypothetical Protein	15,749,646
PITG_16039	Dynactin 1-Related Microtubule-Binding	15,753,509
PITG_16040	Dynactin 1-Related Microtubule-Binding	15,758,006
PITG_16041	Conserved Hypothetical Protein	15,758,995
PITG_16042	Conserved Hypothetical Protein	15,761,608
PITG_16043	Conserved Hypothetical Protein	15,764,010
PITG_16044	Conserved Hypothetical Protein	15,768,906
PITG_16047	Glyceraldehyde-3-Phosphate Dehydrogenase	15,789,739
PITG_19495	Glyceraldehyde-3-Phosphate Dehydrogenase	15,793,469
PITG_16048	Triosephosphate Isomerase	15,794,782
PITG_16052	Inward Rectifier Potassium Channel	15,850,665
PITG_16053	Conserved Hypothetical Protein	15,854,638
PITG_16054	60S Ribosomal Export Protein Nmd3, Putative	15,855,747
PITG_15975	Putative Gpi-Anchored Hypothetical Protein	15,856,849
PITG_16055	Dipeptidyl Peptidase 2	15,857,703
PITG_16056	Cyclin-Dependent Kinase-Like 2	15,859,515
PITG_16057	Enolase 1, Chloroplastic	15,861,830
PITG_16058	Conserved Hypothetical Protein	15,863,839
PITG_16059	Conserved Hypothetical Protein	15,875,830
PITG_16060	Set And Mynd Domain Containing	15,876,615
PITG_16061	Na/K-Transporting ATPase Subunit Alpha, Putative	15,877,699
PITG_16062	Glycosyltransferase	15,882,234
PITG_16063	Conserved Hypothetical Protein	15,884,004
PITG_16064	Conserved Hypothetical Protein	15,885,856
PITG_16065	Rnf5	15,886,876
PITG_16066	Conserved Hypothetical Protein	15,887,629
PITG_16067	Prolyl Hydroxylase; 2-Oxoglutarate; Fe-Dependent Oxygenase Domain	15,888,243
PITG_16068	Conserved Hypothetical Protein	15,889,582
PITG_16069	ATP-Dependent Clp Protease Domain Protein	15,890,646
PITG_16070	Conserved Hypothetical Protein	15,894,048
PITG_16071	Protein Farnesyltransferase Subunit Beta	15,895,085
PITG_16072	Conserved Hypothetical Protein	15,906,764
PITG_16073	Zn-Binding Alcohol Dehydrogenase Domain Protein 2	15,907,802
PITG_16074	ATP-Dependent Clp Protease Domain	15,909,364
PITG_16069	ATP-Dependent Clp Protease Domain	15,909,365
PITG_16075	Conserved Hypothetical Protein	15,912,327
PITG_16076	2-Hydroxyacid Dehydrogenase-Related	15,915,664
PITG_16077	2-Hydroxyacid Dehydrogenase-Related	15,917,115
PITG_16078	PPR-Containing Protein At1G03560, Mitochondrial; pre	15,920,529
PITG_16079	Conserved Hypothetical Protein	15,924,678
PITG_16080	Conserved Hypothetical Protein	15,928,788
PITG_16083	2-Hydroxyacid Dehydrogenase-Related	15,945,274
PITG_16084	Conserved Hypothetical Protein	15,946,922
PITG_16088	Asparagine Synthetase, Glutamine-Hydrolyzing	16,028,565
PITG_16089	Oxidored-Nitro Domain Protein	16,030,592
PITG_16090	Nedd4-Like E3 Ubiquitin-Protein Ligase Wwp1	16,032,141
PITG_16091	Conserved Hypothetical Protein	16,039,134
PITG_16092	Mate Efflux Family Protein Dtx1	16,040,550
PITG_16093	Zn finger, E3 Ubiquitin-Protein Ligase Rnf115	16,044,318
PITG_16094	Conserved Hypothetical Protein	16,045,939
PITG_16095	Nck-Associated Protein 1 Homolog	16,046,864
PITG_16096	Conserved Hypothetical Protein	16,050,794
PITG_16097	Sodium/Hydrogen Exchanger	16,053,583
PITG_16098	2-Ketoacid Reductase	16,055,401
PITG_16105	E3 Ubiquitin-Protein Ligase Praja	16,064,265
PITG_16104	Cryptochrome-2	16,067,979
PITG_16103	Major Facilitator Superfamily Domain Protein 6	16,070,367
PITG_16102	Homocysteine S-Methyltransferase Ybgg	16,072,093
PITG_16101	Ngep-Related, Anoctamin-8	16,073,218
PITG_16100	Cryptochrome-Like Protein	16,078,134
PITG_16107	Rrp5 Homolog; Programmed Cell Death Protein 11	16,148,402
PITG_16108	Cation Channel Sperm-Associated Protein 3	16,154,794
PITG_16111	Conserved Hypothetical Protein	16,174,670
PITG_16112	Intraflagellar Transport Protein 57 Homolog	16,176,312
PITG_16113	Methylthioribulose-1-Phosphate Dehydratase	16,177,758
PITG_16114	Transcriptional Activator Myb	16,179,912
PITG_16115	Conserved Hypothetical Protein	16,183,433
PITG_16116	Dna-Directed Rna Polymerase III Subunit, Putative	16,184,635
PITG_16117	Swi/Snf-Related Chromatin Binding Protein Dna Binding Protein	16,186,492
PITG_16118	Conserved Hypothetical Protein	16,189,062
PITG_16120	Conserved Hypothetical Protein	16,285,348
PITG_16121	Nucleoporin Nup43	16,287,916
PITG_16125	Conserved Hypothetical Protein	16,309,481
PITG_16126	Na-Dependent Glutamate/Aspartate Transporter 2	16,317,808
PITG_16127	Fun14 Domain Protein 1	16,359,363
PITG_16128	Conserved Hypothetical Protein	16,360,980
PITG_16131	Pym, Partner Of Y14 And Mago	16,370,720
PITG_16132	Phosphoinositide Phosphatase Sac6-Related	16,371,670
PITG_16133	MFS Domain Protein 6; Macrophage Mhc Class I Receptor 2	16,379,156
PITG_16134	Conserved Hypothetical Protein	16,381,816
PITG_16137	Ubiquitin Fusion Degradaton Protein 1	16,412,221
PITG_16138	Exosome Complex Component Rrp42	16,414,012
PITG_01879	Conserved Hypothetical Protein	16,457,636
PITG_01878	Tbc1 Domain Family Member 23	16,460,186
PITG_01875	Secreted RxLR	16,470,584
PITG_20972	Secreted RxLR	16,470,814
PITG_09120	Conserved Hypothetical Protein	16,484,716
PITG_01873	Phosphatidylinositol-4-Phosphate 5-Kinase Related	16,503,787
PITG_01872	Conserved Hypothetical Protein	16,507,448
PITG_01871	Serine-Threonine Protein Kinase	16,510,692
PITG_01870	Protein Nlrc3; Nucleotide-Binding Oligomerization Domain Protein 3	16,513,645
PITG_01869	Trehalose-6-Phosphate Synthase; Alpha	16,515,436
PITG_01868	Tetratricopeptide Repeat Protein 26	16,518,830
PITG_01867	Splicing Factor U2Af 65 Kda Subunit	16,520,662
PITG_21959	Conserved Hypothetical Protein	16,523,660
PITG_01866	F-Box/Lrr-Repeat Protein 4; Atfbl4	16,525,514
PITG_01865	Regulator Of Chromosome Condensation	16,527,806
PITG_01864	Conserved Hypothetical Protein	16,530,683
PITG_01862	Glucose-6-Phosphate Isomerase	16,532,116
PITG_01860	Transmembrane Protein, Putative	16,538,803
PITG_01859	Zswim2, E3 Ubiquitin-Protein Ligase	16,542,294
PITG_01858	Derlin-3	16,543,958
PITG_01857	Ring Finger And Swim Domain Protein 2	16,545,414
PITG_01856	Cyclic Nucleotide-Dependent Protein Kinase	16,547,076
PITG_01855	ER Mannosyl-Oligosaccharide 1,2-Alpha-Mannosidase	16,548,364
PITG_01853	Protein Stx16-Npepl1-Related	16,552,749
PITG_01852	Dynein Alpha Chain, Flagellar Outer Arm, Dhc Alpha	16,553,933
PITG_01851	Homogentisate 1,2-Dioxygenase	16,566,664
PITG_01850	Probable 125 Kda Kinesin-Related Protein	16,568,816
PITG_01849	Dual Specificity Protein Phosphatase	16,587,897
PITG_01848	Dual Specificity Protein Phosphatase	16,590,072
PITG_01847	Ring Finger And Swim Domain Protein 2	16,592,633
PITG_01846	Actin-Related Protein 2/3 Complex Subunit 2	16,594,294
PITG_01845	Ring Finger And Swim Domain Protein 2	16,596,978
PITG_01844	Ring Finger And Swim Domain Protein 2	16,597,960
PITG_01842	Ring Finger And Swim Domain Protein 2	16,609,692
PITG_01841	Ring Finger And Swim Domain Protein 2	16,613,159
PITG_01840	Ring Finger And Swim Domain Protein 2	16,660,852
PITG_01839	4-Coumarate--Coa Ligase 3	16,662,276
PITG_01838	Ring Finger And Swim Domain Protein 2	16,665,071
PITG_01837	Chaperonin; T-Complex Protein 1 Subunit Gamma	16,666,385
PITG_20759	Dihydroxy-Acid Dehydratase, Mitochondrial	16,684,169
PITG_22684	Peptidyl-Prolyl Cis-Trans Isomerase-Related	16,687,552
PITG_01833	60S Ribosomal Protein L13A	16,689,305
PITG_01832	Scy1-Related S/T Protein Kinase-Like Protein 2	16,690,831
PITG_15593	Ring Finger And Swim Domain Protein 2	16,711,164
PITG_12648	Ring Finger And Swim Domain Protein 2	16,711,189
PITG_17063	Secreted RxLR	16,769,172
PITG_16137	Ubiquitin Fusion Degradaton Protein 1	16,815,953
PITG_16138	Exosome Complex Component Rrp42	16,817,748

significant QTL peaks (LOD>5.0) were not observed for either parent in regards to the sporulation, pathogenicity, or growth rate phenotypes (**Figure 3.3**).

During the course of the metalaxyl trials, several progeny were observed to develop sectors of mycelia growing considerably faster than the rest of the culture (**Figure 3.4a**), through this phenomenon was not observed for other phenotypes, or for cultures growing on normal rye-A media. When samples of each sector were single-zoospore purified and re-phenotyped, the sizeable increase in resistance to the fungicide was maintained (**Figure 3.4b**, **Table 3.1**). Three of the four isolated progeny pairs had resistant metalaxyl response phenotypes for their normal sectors, while the sensitive individual, M91N, had a borderline intermediately resistant phenotype. Interestingly, other phenotypes were similarly affected to varying degrees, most notably being an 8.2-fold increase in sporulation proficiency in M64F over M64N, and a two score decrease in pathogenicity for M78F over M78N (**Table 3.1**). When the haplotypes of parent 618 chromosome 3 were resolved, one haplotype showed a decrease in linkage and the end of chromosome 3 (chr3A, **Figure 3.2d**) while the other showed an increase (chr3B, **Figure 3.2e**), potentially informing on which haplotype the major locus of resistance resides in parent 618. 24 unique markers originating from haplotype 3B were combined with a metalaxyl resistance marker to generate a genetic map in JoinMap (not shown). The linkage group spanned 83.8 cM, but the metalaxyl resistance marker resided

Figure 3.3. QTL analysis of **A**) Radial growth rate on rye agar, **B**) Sporulation proficiency, and **C**) Pathogenicity on potato tubers. No genomic intervals from either parent show a strong (LOD>5.0) association with any phenotype.

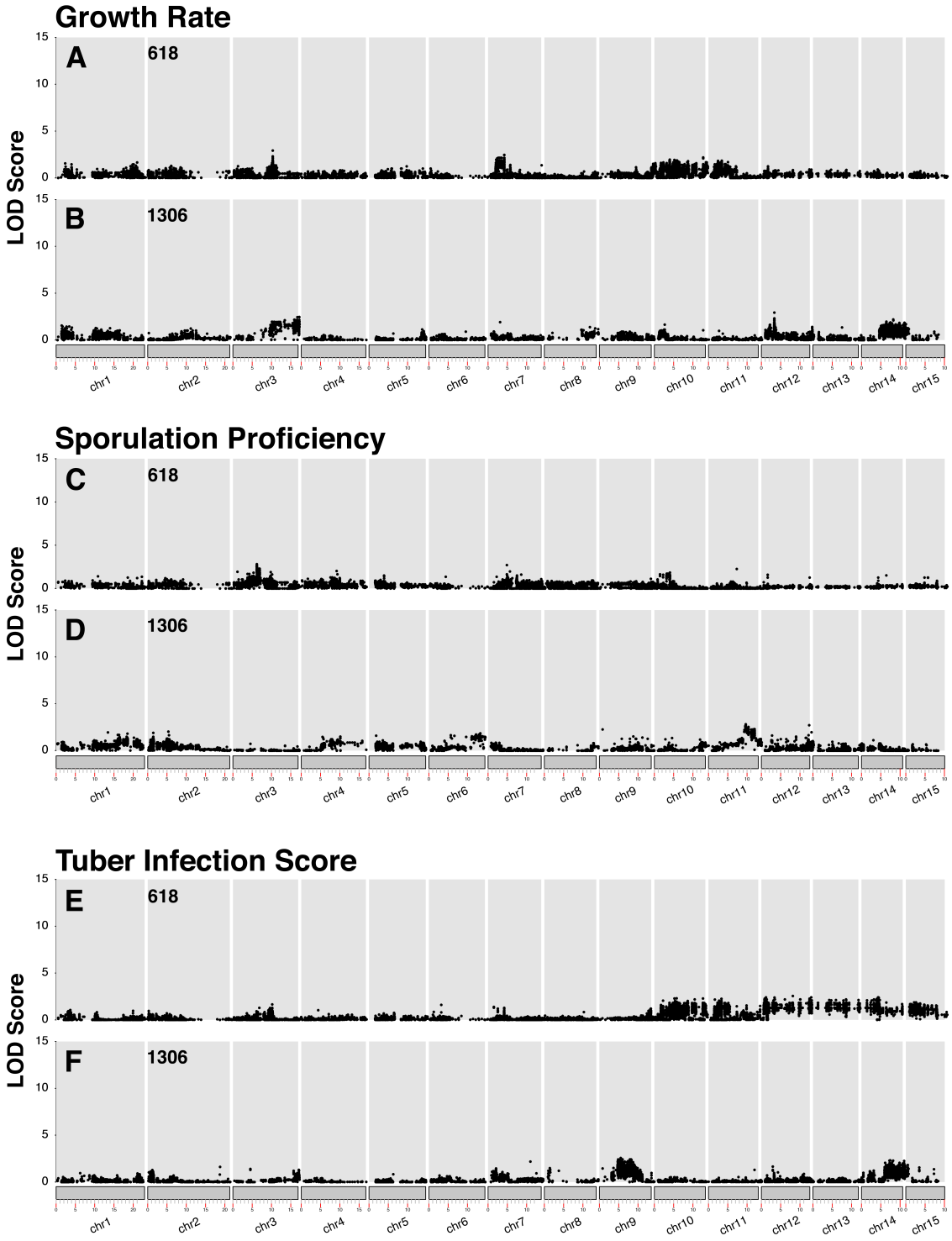
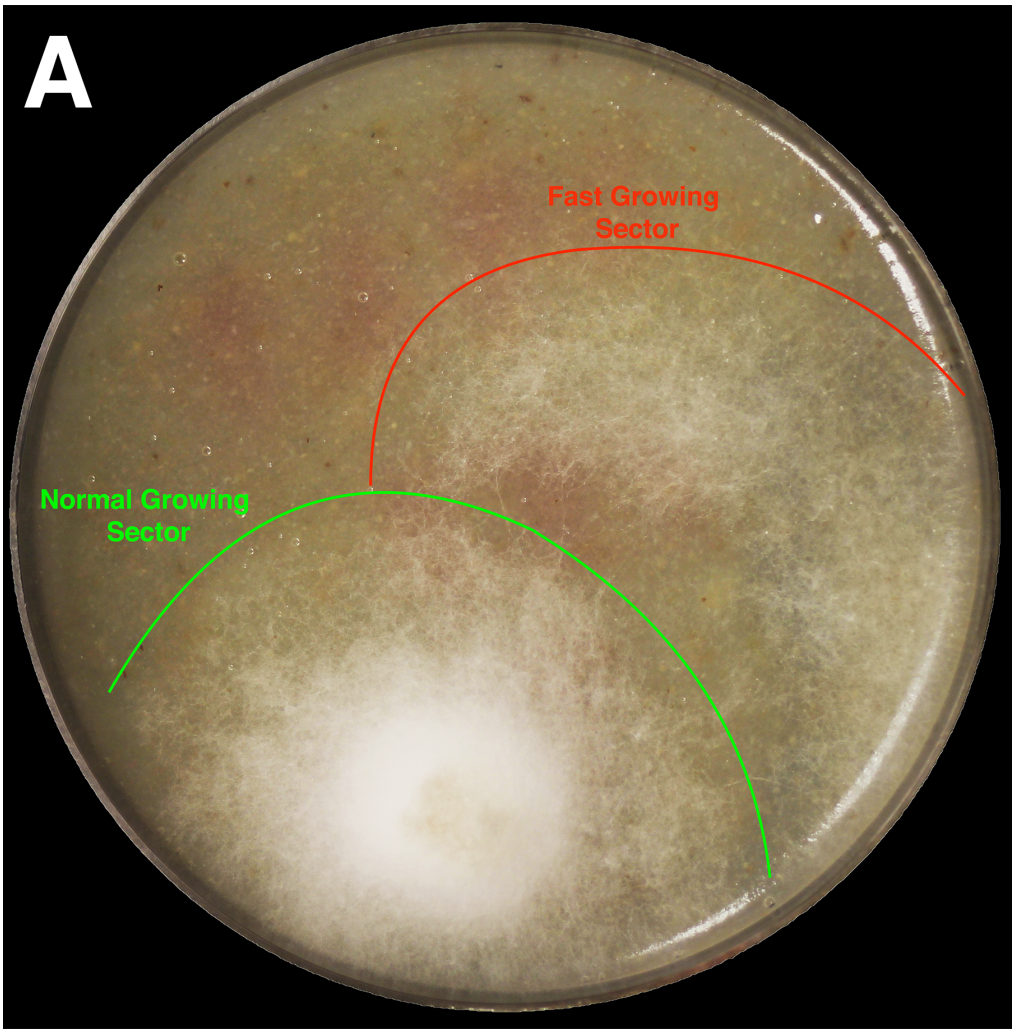
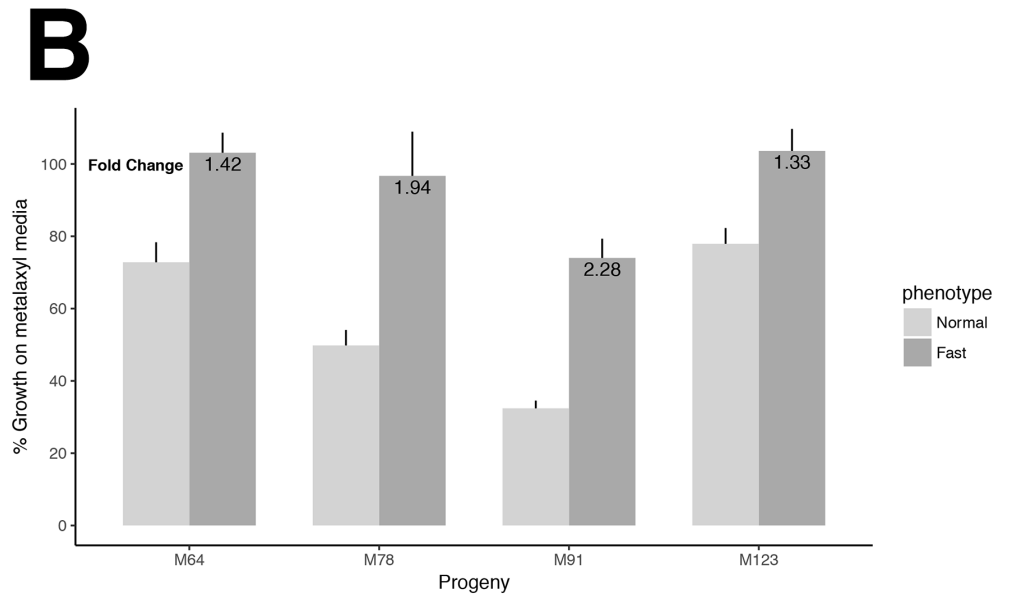

Growth Rate
A
618
B
1306
Sporulation Proficiency
C
618
D
1306
Tuber Infection Score
E
618
F
1306
LOD Score
0
5
10
15
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15

Figure 3.4. Confirmation of increased metalaxyl resistance in single-zoospore derivatives of progeny spontaneously growing faster on media amended with metalaxyl. **A**) Example of a sector (in red) growing faster than the normal growing sector (green) on a rye agar plate with 0.5 µg/ml metalaxyl. **B**) Pairs of fast and normal-growing sectors for each individual (x axis) following single-zoospore purification. All four pairs showed an increased in metalaxyl resistance (y axis) for the fast sectors compared to the normal sectors. Interestingly, the most metalaxyl-resistant "normal" isolate, M123N, showed the lowest fold-change of increased resistance following single-zoospore purification, while the most metalaxyl sensitive normal isolate, M91N, showed the largest fold-change of resistance. Error bars are SEM.


A
Fast Growing
Sector
Normal Growing
Sector

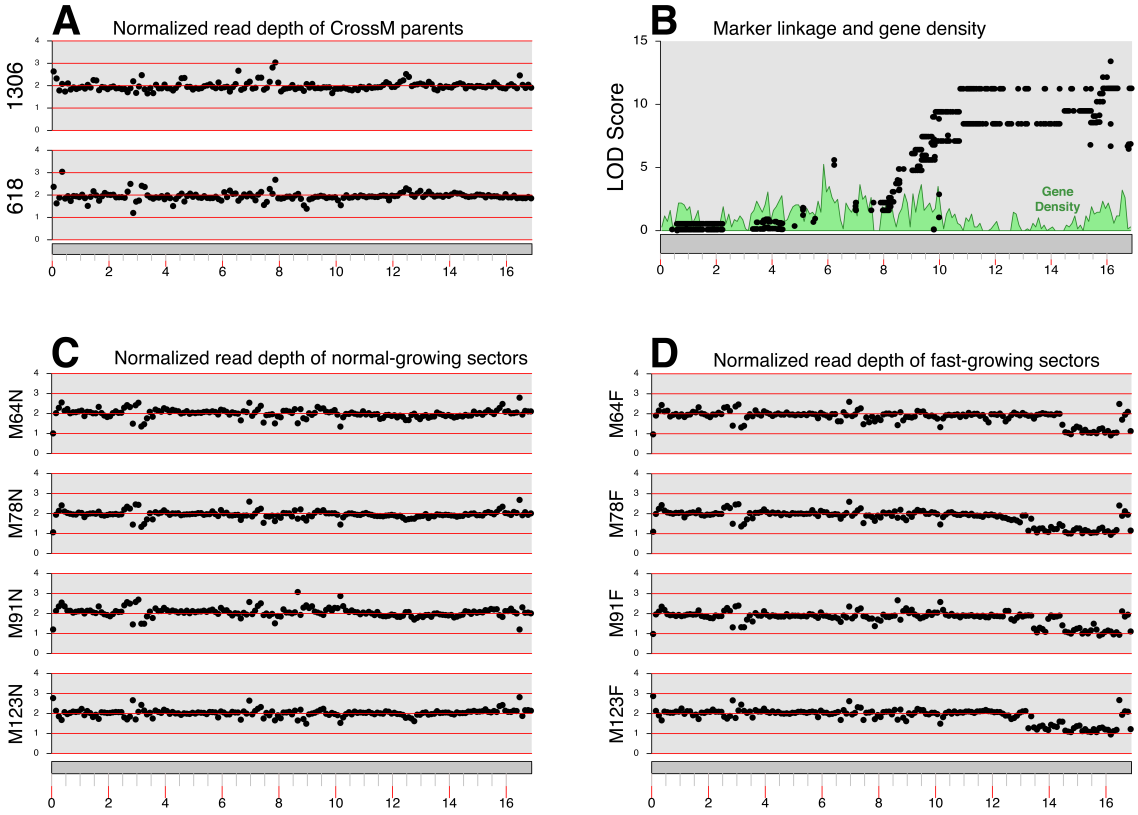

B
Fold Change
1.42
1.94
2.28
1.33
100
80
60
40
20
0
% Growth on metalaxyl media
M64
M78
M91
M123
Progeny
phenotype
Normal
Fast

17.7 cM away from the nearest chr3B marker, and no flanking markers were identified around the resistance locus.

Structural variation and metalaxyl resistance

To investigate the origin of this increased resistance, we examined the structural variation between the normal and fast growing sectors. Using the Comai Lab's script to measure normalized read depth, a hemizygous deletion at the end of chromosome 3, absent in the normal growing individuals, was observed in the fast-growing individuals (**Figure 3.5d**). This deletion coincides with the peak of genetic linkage with resistance (**Figure 3.5b**). Individual M64F contains the smallest deleted region, starting around position 14.4 Mb, resulting in a 2.4 Mb deletion and providing a minimal border for the genomic interval containing the functional element contributing to the increased resistance phenotype. The other three fast-growing individuals all showed a break around 13.4 Mb, resulting in a 3.4 Mb deletion. This deletion was not observed in other metalaxyl resistant progeny, nor was it present in the cross parents. The deletion breakpoint varied among the four individuals containing it, suggesting a biological origin as opposed to an assembly artifact. However, several regions of apparent deletion are inconsistent with the most parsimonious explanation of a single deletion event of a chromosomal arm. Two 500 kb regions of sequence appear

Figure 3.5. The sectors growing fast on metalaxyl media share a deletion at the end of chromosome 3. **A**) Normalized read depth along chromosome 3 (in Mb, x axis) shows that neither parent 1306 or 618 show unusual read depth patterns along chromosome 3 and maintain a copy number of two (y axis). **B**) Statistical association of all markers heterozygous in parent 618 and homozygous in parent 1306 to metalaxyl resistance. Relative gene density is plotted concurrently, showing the deleted regions have functional consequences. **C**) Read depths of the sectors growing normally on metalaxyl media mostly resemble the parents' read depths, and do not show any major perturbations. **D**) Read depths of the fast growing sectors all show a major deletion at the end of chromosome 3, with varying breakpoints. The breakpoints occur in a relatively gene-sparse, and thus repeat-rich regions. This suggests that they could be influenced by TE integration or recombination between repeated sequences.

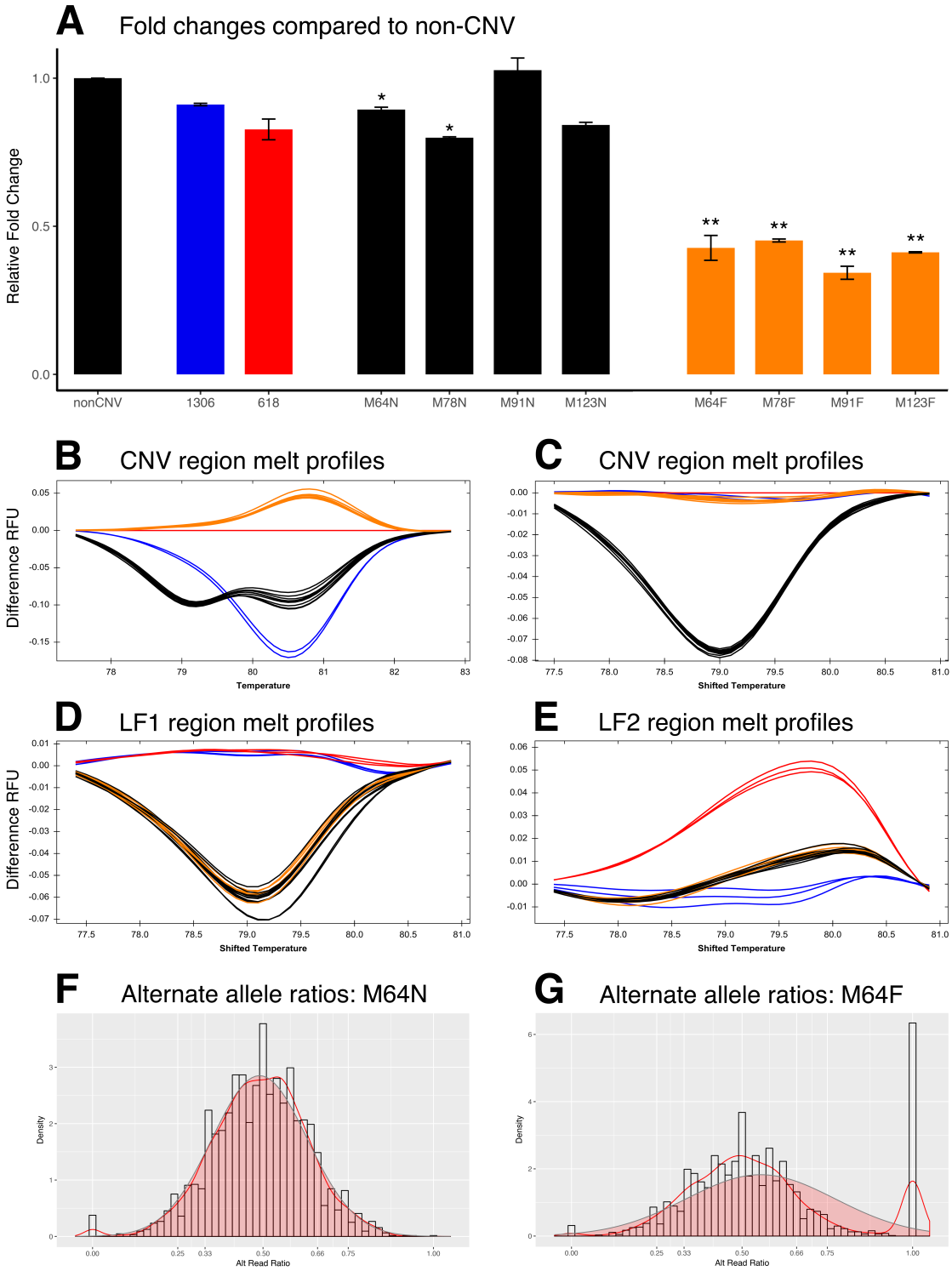

A
Normalized read depth of CrossM parents
1306
618
B
Marker linkage and gene density
LOD Score
Gene Density
C
Normalized read depth of normal-growing sectors
M64N
M78N
M91N
M123N
D
Normalized read depth of fast-growing sectors
M64F
M78F
M91F
M123F

to alternate copy number in M91F, highlighting a potential assembly artifact in the reference genome, and each fast-growing individual has a 400 kb region of two copies close to the end of the chromosome arm flanked by a single 100kb region of one copy at the very end. That chromosome 3 lacks BLAST hits against telomere sequences may also be an indication that the very end of the scaffold is not optimally assembled.

Further confirmation of this deletion was achieved through qPCR and High Resolution Melt (HRM) analysis. Three primer sets amplified SNPs from two regions outside of the deletion, and one region within the deletion (**Table 3.4**). All SNPs were homozygous for different bases in the parents, meaning they are expected to be heterozygous in all progeny. Copy number fold-changes of the deleted region of the parents and normal-growing sectors remained around 1.0, with the exception of M64N and M78N which were weakly significantly different than the non-CNV region (**Figure 3.6a**). In contrast, all four fast-growing progeny had fold changes around 0.5 for the CNV region, consistent with a loss of a single copy.

When the melt profiles were determined, there was a clear difference between the normal and fast growing individuals (black vs orange, **Figure 3.6b**), consistent with a loss of heterozygosity in the fast growing progeny. Switching to a temperature-shifted view clearly demonstrated clustering among the melt profiles of the individuals with two alleles (black), and individuals with only a

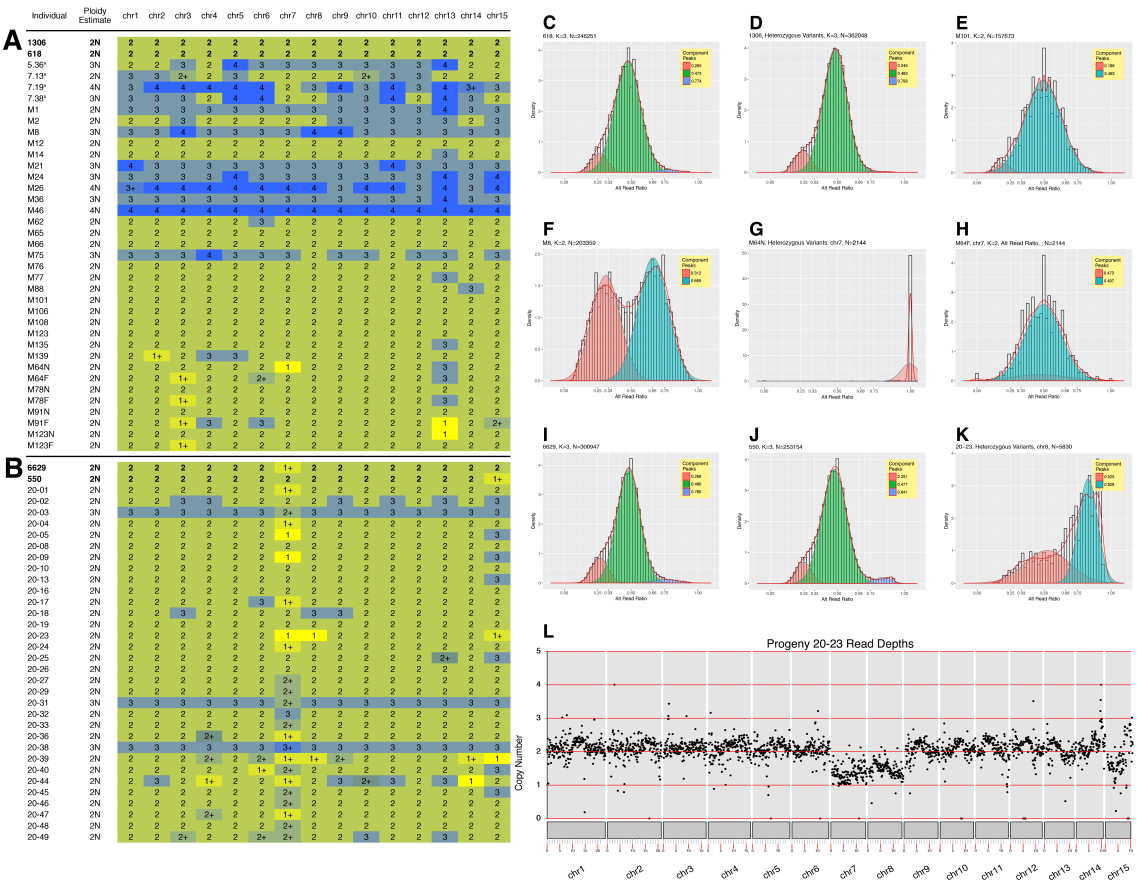
Figure 3.6. qPCR and High Resolution Melt (HRM) analysis of the deleted region at the end of chromosome 3. **A**) ΔCq values of the deletion region primers (CNV: chr3, 16.2 Mb) compared to the non-deletion region primers (LF1: chr3, 6.1 and LF2: 11.1 Mb)(**Table 3.4**). Parent 618 is shown in red, parent 1306 in blue, normal sectors black, and fast growing sectors tan. *$p < 0.05$, ** $p < 0.005$ Both parents, M91N, and M123N show no statistical different between deleted and non-deleted regions of chromosome 3, while M64N ($p = 0.045$) and M78N ($p = 0.024$) show barely significant differences between deleted and non deleted regions. All fast growing deletion regions show high levels of fold change between deletion and non deletion regions. **B, C**) HRM melt profiles of the normal and fast growing sectors in the deletion region clearly do not cluster in either temperature difference, or shifted temperature difference models. **D, E**) Outside of the deletion region, normal and fast sector isolates are expected to both be heterozygous, and they both cluster together and not with the homozygous parents. **F, G**) Ratios of alternate calls divided by total calls for 3,194 variants expected to be heterozygous on chromosome 3. Normal distributions (red fill, grey line), kernel density (red line), and histogram bins (bars) were plotted. **F**) Read ratios originating from M64N show the expected normal distribution of ratios around 0.50 with only minimal loss of heterozygosity (LOH) at 0.0. **G**) Read ratios of variants expected to be heterozygous in M64F show zero considerable LOH piled up at 1.0 due to the loss reference alleles.


A
Fold changes compared to non-CNV
Relative Fold Change
1.0
0.5
0.0
nonCNV
1306
618
M64N
M78N
M91N
M123N
M64F
M78F
M91F
M123F
B
CNV region melt profiles
Differennce RFU
Temperature
C
CNV region melt profiles
Shifted Temperature
D
LF1 region melt profiles
Differennce RFU
Shifted Temperature
E
LF2 region melt profiles
Shifted Temperature
F
Alternate allele ratios: M64N
Density
Alt Read Ratio
G
Alternate allele ratios: M64F
Density
Alt Read Ratio

single allele present (**Figure 3.6c**). The other two non-deletion amplicons show clustering among both the normal and fast growing individuals, and clearly distinct from the homozygous parents (**Figure 3.6d,e**). Withing the smallest deletion, observed in M64F from 14.4 Mb to the end, a total of 199 genes are annotated.

Finally, the allele read ratios of 3,194 variants on chromosome 3 expected to be heterozygous in all progeny were plotted. Each normal growing individual showed the expected normal distribution of alternate allele ratios centered around 0.50 (**Figure 3.6f**). However, when the allele ratios were plotted for the fast growing individuals, a large pileup of genotypes can be seen at 1.00 (**Figure 3.7g**). This loss of heterozygosity (LOH) implicates a loss of the reference allele since the total read depth consists solely of alternate basecalls. The individual variants contributing to this LOH all reside at the end of chromosome 3 along the deleted interval (not shown). Additionally, the resolved haplotypes of chromosome 3 for parent 618 show that haplotype chr3A seems to lack markers contributing to linkage to metalaxyl resistance at the end of chromosome 3 (**Figure 3.2d**), while the chr3B haplotype shows markers with increasing linkage over the same interval (**Figure 3.2e**). This is because haplotype chr3A shares the same allele with parent 1306 at the major locus and thus the chr3A variants are called as “reference”, and do not contribute to QTL linkage. On the contrary,

Figure 3.7. Progeny from CrossM and Cross20 show a diversity of karyotypes. **A**) Inferred copy numbers. of individual chromosomes of 35 progeny and two parents from CrossM. Chromosomes showing "+" have truncations or translocations. For example, the fast growing progeny on metalaxyl all have chromosome 3 listed as 1+ since they all have one complete and one partial copy of chromosome 3. .**B**) Chromosome copy numbers for Cross20. **C,D,E,F**) Distribution of allele read ratios from parents 618 and 1306, and progeny M101 and M8, respectively. 618, 1306, and M101 are not expected to be tetraploid or triploid, but modeling the distributions underlying noise gives a better fitting model of the main (green) distribution centered around 0.5. M8 displays a clearly triploid distribution. **G**) M64N lacks the 1306 copy of chromosome 7, causing variants expected to be heterozygous to pile up around 1.0. **H**) M64F still has its 1306 copy, and shows the expected distribution around 0.5 for the same set of variants. **I,J**) Parents 6629 and 550 show the expected distribution around 0.5 for the main (green) model. **K**) Progeny 20-23 is likely missing a copy of chromosome 7 and 8 (**L, Table 3.2**), causing chromosome 8 to show a highly skewed profile, possibly due to a heterokaryon mixture.


A
B
C
D
E
F
G
H
I
J
K
L
Progeny 20-23 Read Depths
Copy Number
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15

the alleles from haplotype chr3B do contain the resistant/alternate variant call, and thus contribute to QTL linkage.

Genome wide structural variation in the progeny

A subset of 35 progeny from CrossM and all 32 progeny from Cross20 were examined for copy number and structural variation at the genome-wide level. Read depth and variant ratios methods were used to reveal the presence or absence of whole chromosomes, chromosome arm duplications or deletions, the number of copies originating from each parent, and both deletion-derived and copy-neutral loss of heterozygosity (LOH)(**Figure 3.7**). The number of genome-wide heterozygous variants for all cross parents ranged from 246,251-300,947, while variant numbers in the progeny ranged from 143,223-205,180 (**Table 3.2**). The subset of 38,040 variants expected to be heterozygous in CrossM allowed for the prediction of how many copies of each parent's alleles are present due to the mode of the ratio distribution, while two datasets of 40,017 and 38,034 variants homozygous alternate/recessive in parents 550 and 6629, respectively, informed on the Cross20 progeny (**Table 3.2**). While most progeny appear to be diploid, many higher-ploidy individuals are seen in both crosses, as well as individuals showing a complex mixture of chromosome copy numbers (**Figure 3.7**). A few progeny have chromosomes that either show conflicting signals or

have intermediate and ambiguous signals, presumably do the a mixture of nuclear types in a heterokaryon state at the time of DNA isolation. In such cases, a best estimate is made through using read depth as a determining data point.

Three notable trends are observed. First, trisomy and tetrasomies are common. In CrossM, only 11 of the 37 parents and progeny lack any non-disomic chromosomes, and 8 contain a single non-disomic chromosome. 175 of 555 (32.2%) total chromosomes from parents and progeny were at, or higher than, a predicted copy number of 3, and only 8 chromosomes (1.4%) were below copy number 2. Only four of 34 Cross20 parents and progeny contained no non-disomic chromosomes, through if one disregards the deletion originating from parent 6629 on chromosome 7, four additional parents and progeny fall into that category. Cross chromosomes were at, or higher than, a predicted copy number of 3 in 87 cases (17.1%) while copy numbers lower than 2 were observed in 21 cases (4.1%).

Second, tetrasomy is often identified within an individual only through read depths higher than chromosomes showing a classic trisomy profile. For example, chromosome 1 on progeny 7.19 shows a triploid allele ratio, while chromosome 2, at a higher read depth, shows a diploid-like ratio. This is presumably due to recent autotetraploidy leaving the alleles in a diploid-like state. Third, some chromosomes are more prone to perturbations than others in each cross. Chromosomes 3 tended to show instability for CrossM in that more

progeny show monosomic or trisomic karyotypes for chromosome 13 than other chromosomes. Chromosomes 7 and 15 show instability for Cross20. Additionally, parent 6629 has a ~3 Mb deletion at the beginning of chromosome 7 and parent 550 has a ~4 Mb deletion at the end of chromosome 15, both of which segregate in the progeny.

RNA-seq

34 CrossM progeny with varying levels of metalaxyl resistance, plus the two parents were considered for transcriptomics analysis, though only 26 progeny with high (15) or low (11) resistance phenotypes were considered for analysis and the other eight intermediately resistant progeny were not considered. Differentially Expressed Genes (DEGs) between individuals with differing metalaxyl sensitivities and between progeny growing in metalaxyl amended and non-amended media were sought. Since all individuals were grown in liquid rye-A media and not on potato tubers or rye agar, the biological context for examining the three other phenotypes made analysis inappropriate. Following per-gene normalization of edgeR CPM values, we observed considerable diversity of expression levels between the progeny. While the presence or absence of metalaxyl affected some genes' expression patterns, genetic diversity of the progeny was contributed more to differential expression

than metalaxyl treatment as evidenced by the preference of individuals to cluster together based on genotype rather than metalaxyl presence versus absence (**Figure 3.8a**). Nevertheless, genome-wide, more genes were significantly differentially expressed between individuals treated with metalaxyl versus no metalaxyl (308) than between resistant and sensitive progeny (120)(**Table 3.6**). When considering all progeny, 28 and 19 genes were highly significant, showing a significant FDR (Stepup, <0.05) and a log2 fold change larger than +/- 1.0, when comparing resistant vs. sensitive progeny and when metalaxyl is present vs. absent, respectively (**Table 3.6, Figure 3.8b,d**).

Only three and four DEGs were identified when comparing resistant vs sensitive progeny between individuals treated, and not treated, with metalaxyl, respectively (**Table 3.6**). The only gene with an annotated function in both cases was PITG_14400, a MYB-domain-containg H2A deubiquitinase MYSM1-like gene downregulated in resistant progeny. Conversely, many more DEGs were identified between sensitive progeny with and without metalaxyl (342), while much fewer were identified in resistant progeny (28). All three comparisons of metalaxyl present versus absent growth conditions (**Table 3.6**) showed considerably more DEGs being upregulated in the presence of metalaxyl compared to downregulated DEGs, though this is more balanced when comparing resistant versus sensitive progeny.

Figure 3.8. Diversity of CrossM transcriptome profiles. Heat maps show per-gene normalized counts-per-million (CPM) values from edgeR, clustered on both axes, and scaled to a maximum CPM value of 3.0. **A**) Heat map of five individuals and their runs of metalaxyl present (Mex) and metalaxyl absent (NoMex). Sensitive individuals (5.36, parent 1306) are more affected by metalaxyl than resistant individuals (M76, parent 618), but metalaxyl treatments do not affect transcriptome diversity more than underlying genotype. **B**) Volcano plot of all genes demonstrating how genes are preferentially upregulated in response to metalaxyl. Data includes the log2 fold change in response to the presence of metalaxyl, the p-value of differential expression, and corrected stepup FDR value. **C**) Few progeny cluster together overall, and metalaxyl resistance shows no correlation with transcriptome profile. Progeny with high metalaxyl resistance show % growth rate on metalaxyl media > 70% compared to growth rate on unamended media, Mid resistance is less than 70% and greater thn 40% relative growth rate, and low resistance is < 40% relative growth rate. The occurrence of gene clusters highly upregulated in only a single progeny also demonstrates the diversity of transcriptomes among the progeny. **D**) Volcano plot demonstrating that fewer genes are upregulated in resistant or sensitive progeny compared to progeny treated with metalaxyl, and that the split is more evenly distributed.

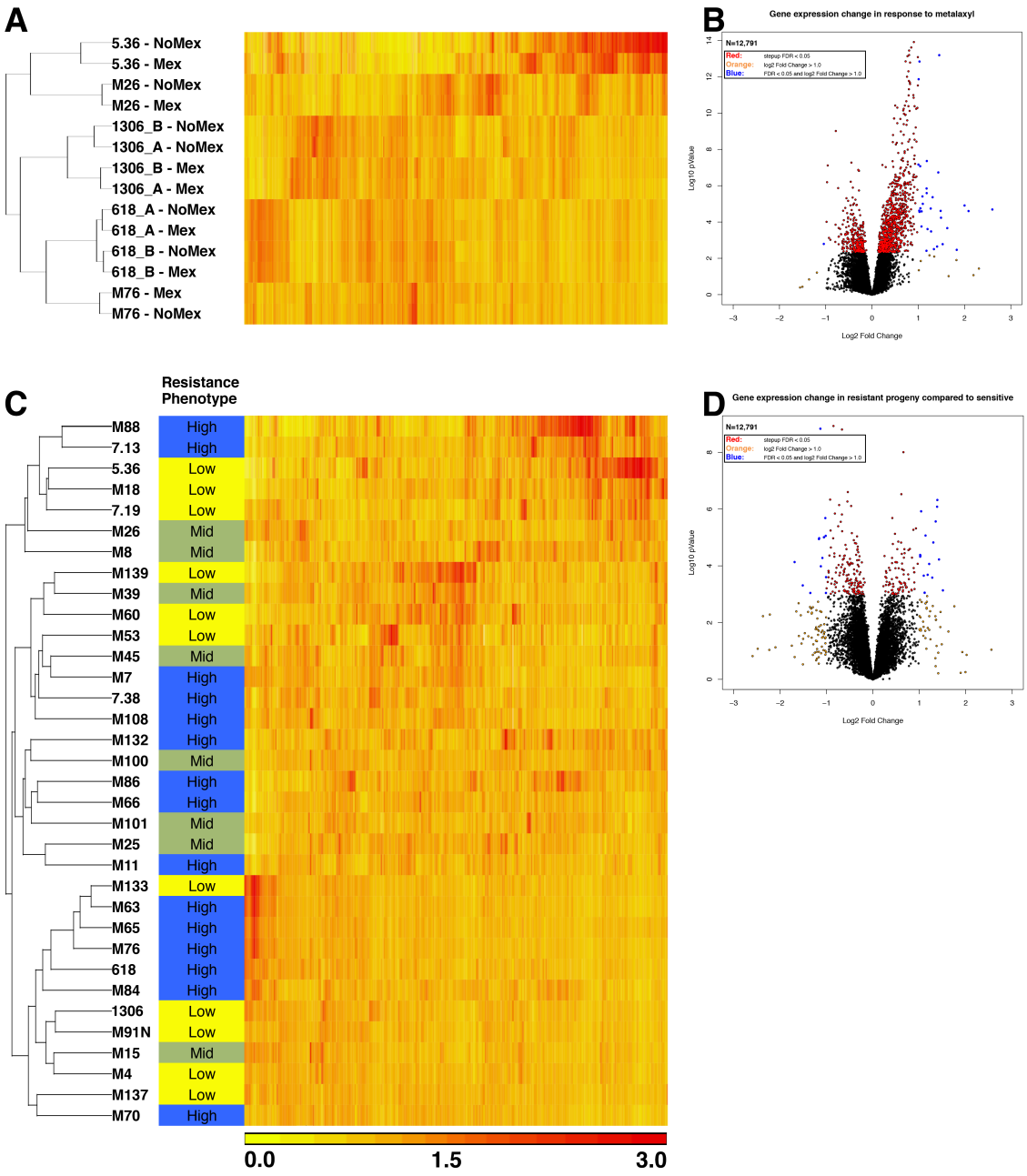

A
5.36 - NoMex
5.36 - Mex
M26 - NoMex
M26 - Mex
1306_B - NoMex
1306_A - NoMex
1306_B - Mex
1306_A - Mex
618_A - NoMex
618_A - Mex
618_B - NoMex
618_B - Mex
M76 - Mex
M76 - NoMex
B
Gene expression change in response to metalaxyl
N=12,791
Red:
Orange:
Blue:
Log10 pValue
Log2 Fold Change
C
Resistance Phenotype
M88 High
7.13 High
5.36 Low
M18 Low
7.19 Low
M26 Mid
M8 Mid
M139 Low
M39 Mid
M60 Low
M53 Low
M45 Mid
M7 High
7.38 High
M108 High
M132 High
M100 Mid
M86 High
M66 High
M101 Mid
M25 Mid
M11 High
M133 Low
M63 High
M65 High
M76 High
618 High
M84 High
1306 Low
M91N Low
M15 Mid
M4 Low
M137 Low
M70 High
0.0
1.5
3.0
D
Gene expression change in resistant progeny compared to sensitive
N=12,791
Log10 pValue
Log2 Fold Change

Table 3.6, Numbers of genes differentially expressed in relation to metalaxyl

Control	Comparison	Total DEGs	Up in resistant or present[a]	Significant: Up in resistant or present[b]	Down in resistant or present[c]	Significant: Down in resistant or present[d]	GO term summary
All progeny:	Resistant vs Sensitive	120	56	30	64	17	Sugar/carbohydrate transporters
Metalaxyl present:	Resistant vs Sensitive	3	0	0	3	1	MYSM1-like
Metalaxyl absent:	Resistant vs Sensitive	4	2	1	2	1	MYSM1-like
All progeny:	Metalaxyl present vs absent	308	278	31	30	1	rRNA and transcription
Sensitive progeny:	Metalaxyl present vs absent	342	332	46	10	2	rRNA and transcription
Resistant progeny:	Metalaxyl present vs absent	28	27	2	1	0	rRNA and transcription

[a]Genes upregulated in resistant progeny in the resistant vs sensitive comparison, or progeny treated with metalaxyl in the present vs absent comparison

[b]Genes upregulated with a log2FC >1.0 in resistant progeny in the resistant vs sensitive comparison, or progeny treated with metalaxyl in the present vs absent comparison

[c]Genes downregulated in resistant progeny in the resistant vs sensitive comparison, or progeny treated with metalaxyl in the present vs absent comparison

[d]Genes downregulated with a log2FC >1.0 in resistant progeny in the resistant vs sensitive comparison, or progeny treated with metalaxyl in the present vs absent comparison

Gene Ontology (GO) term analysis of DEGs between progeny treated with or without metalaxyl consisted of terms related rRNA binding and ribosome activity (**Table 3.7)**. Examples include PITG_17261, the 60S acidic ribosomal pritein P0, and PITG_11766, the 40S ribosomal protein S3a, both upregulated in progeny treated with metalaxyl. GO term analysis revealed that DEGs between resistant and sensitive progeny were enriched in several terms related to carbohydrate transporters (**Table 3.8**). Examples include PITG 17760 and PITG_20372, which are amino acid permeases upregulated in resistant progeny.

Discussion

A major locus linked to metalaxyl resistance

The source of metalaxyl resistance in *Phytophthora infestans* has been sought for decades, with many studies describing its incompletely-dominant behavior (Fabritius et al. 1997; Gisi and Cohen 1996; Judelson and Roberts 1999; Lamour and Hausbeck 2000; Lee et al. 1999; Shattock 1988), how major multiple loci have arisen across the globe (Fabritius et al. 1997; Judelson and Roberts 1999), a single point mutation that appears to underlie one of these major loci (Childers et al. 2014), and insights into some of the minor loci which contribute to the overall sensitivity of a given isolate (Chen et al. 2018; Childers

Table 3.7, GO term enrichments for genes differentially expressed in response to metalaxyl treatment

GOID	GO Node Size	Samples Matching Node	FDR value	Ontology	GO Term
GO:0003735	154	41	8.78E-40	Molecular Function	structural constituent of ribosome
GO:0005198	244	41	5.82E-31	Molecular Function	structural molecule activity
GO:0003723	468	34	3.36E-13	Molecular Function	RNA binding
GO:0003676	1035	39	1.91E-06	Molecular Function	nucleic acid binding
GO:0019843	45	7	0.000259623	Molecular Function	rRNA binding
GO:0006412	344	46	2.57E-28	Biological Process	translation
GO:0043043	351	46	6.46E-28	Biological Process	peptide biosynthetic process
GO:0006518	406	46	4.55E-25	Biological Process	peptide metabolic process
GO:0043604	423	46	2.80E-24	Biological Process	amide biosynthetic process
GO:0043603	501	46	4.44E-21	Biological Process	cellular amide metabolic process
GO:0010467	1523	70	4.12E-17	Biological Process	gene expression
GO:1901566	762	50	9.51E-17	Biological Process	organonitrogen compound biosynthetic process
GO:0042254	209	28	4.70E-16	Biological Process	ribosome biogenesis
GO:0022613	280	29	1.31E-13	Biological Process	ribonucleoprotein complex biogenesis
GO:0005840	138	46	8.16E-48	Cellular Component	ribosome
GO:0030529	366	57	3.16E-40	Cellular Component	intracellular ribonucleoprotein complex
GO:1990904	366	57	3.16E-40	Cellular Component	ribonucleoprotein complex
GO:0044391	67	28	3.07E-31	Cellular Component	ribosomal subunit
GO:0022626	57	24	1.20E-26	Cellular Component	cytosolic ribosome
GO:0043228	1331	71	5.20E-21	Cellular Component	non-membrane-bounded organelle
GO:0043232	1331	71	5.20E-21	Cellular Component	intracellular non-membrane-bounded organelle

Table 3.8, GO term enrichment for genes differentially expressed between metalaxyl resistant and sensitive progeny

GOID	GO Node Size	Genes Matching Node	FDR value	Ontology	GO Term
GO:0015144	94	10	1.08E-08	Molecular Function	carbohydrate transmembrane transporter activity
GO:1901476	95	10	1.20E-08	Molecular Function	carbohydrate transporter activity
GO:0005355	25	6	2.82E-07	Molecular Function	glucose transmembrane transporter activity
GO:0015145	26	6	3.65E-07	Molecular Function	monosaccharide transmembrane transporter activity
GO:0015149	26	6	3.65E-07	Molecular Function	hexose transmembrane transporter activity
GO:0051119	71	8	4.60E-07	Molecular Function	sugar transmembrane transporter activity
GO:0016781	4	3	2.94E-05	Molecular Function	phosphotransferase activity, paired acceptors
GO:0050242	4	3	2.94E-05	Molecular Function	pyruvate, phosphate dikinase activity
GO:0004553	234	10	6.74E-05	Molecular Function	hydrolase activity, hydrolyzing O-glycosyl compounds
GO:0016798	242	10	9.13E-05	Molecular Function	hydrolase activity, acting on glycosyl bonds
GO:0004650	26	4	0.00062782	Molecular Function	polygalacturonase activity
GO:0005975	592	21	1.17E-09	Biological Process	carbohydrate metabolic process
GO:0008643	99	9	3.73E-07	Biological Process	carbohydrate transport
GO:0015758	34	6	1.86E-06	Biological Process	glucose transport
GO:0008645	35	6	2.24E-06	Biological Process	hexose transport
GO:0015749	36	6	2.67E-06	Biological Process	monosaccharide transport
GO:0016052	276	12	5.01E-06	Biological Process	carbohydrate catabolic process
GO:0071555	208	10	2.38E-05	Biological Process	cell wall organization
GO:0045229	213	10	2.96E-05	Biological Process	external encapsulating structure organization
GO:0071554	253	10	0.000139353	Biological Process	cell wall organization or biogenesis
GO:0010393	42	5	0.000178899	Biological Process	galacturonan metabolic process
GO:0045488	42	5	0.000178899	Biological Process	pectin metabolic process
GO:0045490	42	5	0.000178899	Biological Process	pectin catabolic process
GO:0005576	767	14	0.002668044	Cellular Component	extracellular region
GO:0009507	190	6	0.018686533	Cellular Component	chloroplast

et al. 2014; Judelson and Senthil 2006). Our understanding of resistance is often limited by challenges of working with oomycetes, such as the relative difficulty in performing sexual crosses and the lack of inbred reference isolates. Additionally, the large genome of *Phytophthora infestans* makes genome resequencing expensive, and its highly repetitive and unstable nature introduces noise.

Despite these challenges, we were able to construct a sexual population of 130 progeny between parent 618 (metalaxyl resistant, A2) and parent 1306 (metalaxyl sensitive, A1) and Illumina sequence 82 progeny to 15-30 fold coverage. The metalaxyl resistance phenotype of the 82 segregated as expected for an incompletely dominant trait (**Figure 3.1a**) and Quantitative Trait Locus (QTL) analysis revealed a highly significant peak of linkage at the end of chromosome 3 (**Figure 3.2**). Since this is the only major peak of resistance in the genome (LOD > 5.0), the gene controlling resistance in parent 618 likely resides in an interval starting at 15.75 Mb and extending to the end of the chromosome. The phased haplotypes of parent 618 in this interval also demonstrate the haplotype of origin for the resistant allele. The drop-off of marker linkage to resistance for haplotype chr3A indicates that there are fewer alternate calls among resistant progeny in that interval (**Figure 3.2d**), while the opposite is true for haplotype chr3B (**Figure 3.2e**). This suggests that the major

locus of resistance lies on the chr3B phase due to their retention in the deleted interval.

A subset of non-redundant chr3B markers was used to incorporate the resistance locus into a genetic map, and the resulting linkage group spanned 83.8 cM and included a total of 25 markers. Unfortunately, the resistance marker mapped to the very last position on the map, 17.7 cM away from the last SNP based marker at 16.89 Mb at the very end of the pseudochromosome. Thus, we cannot be certain that the major locus for metalaxyl resistance resides on chromosome 3, and not in unassembled sequence or smaller contigs. This is due to the lack of markers that appear to confidently flank the locus (**Figure 3.2**), and instead only lead up to an interval of maximum statistical association at the very end of chromosome 3, and a recombination fraction of 0.2. However, this could be due to miscalling the progeny phenotype, or also potentially due to LOH across the major locus in the time between phenotype designation and DNA isolation. It seems likely that the genes within this region probably include the one conferring to the resistance phenotype, but further experimental confirmation will have to be done (discussed below).

Three other agriculturally relevant traits were also scored for the sequenced progeny: radial growth rate, the number of sporangia produced, and growth on potato tubers (**Figure 3.1b-d, Table 3.1**). The sporangia and tuber colonization phenotypes showed considerable diversity among the progeny,

while the growth rate phenotype showed less diversity. QTL analysis could not, however, identify any loci which underlie either of these traits (**Figure 3.3**). While this is unfortunate, it seems unsurprising. Resistance to chemical control agents can be determined by a single or a small number of genes (Blum et al. 2010; Frenkel et al. 2015; Judelson and Roberts 1999), allowing the easy identification of genetically linked genomic regions segregating through a sexual population. Complex traits crucial for an organism's life cycle likely involve the interaction of hundreds of genes, with the final phenotype depending on a number of factors such as dosage or copy number, and transcriptional abundance variation.

Given that the metalaxyl resistance phenotype likely targets rRNA production, any polymerase or ribosome-associated genes at the end of chromosome 3, after 15.75 Mb, could potentially play a role in resistance. However, none of the RNA Pol I subunits were found on chromosome 3 (**Table 3.3**). While the largest subunit, RPA190, is thought to confer resistance in at least some populations (Chen et al. 2018; Randall et al. 2014), our finding is more consistent with a Netherlands field population that showed markers strongly linked to resistance, but did not reside in the same genomic regions as RPA190 (Montes et al. 2016). We identified a total of 130 genes that lie within the interval of linkage at the end of chromosome 3, and several show interesting annotations worth inspecting further. PITG_16054 is the *P. infestans* NMD3 ortholog, which has activity in the 60S rRNA processing and export to the

cytoplasm (Bai et al. 2013). Importantly, Bai et al. found that decreased NMD3 activity reduced the rate of pre-47S rRNA synthesis, a very similar effect to what metalaxyl has been shown to induce in *Phytophthora megasperma* (Davidse et al. 1983). NMD3/PITG_16504 was also found to be upregulated in response to metalaxyl treatment (**Table 3.8**). PITG_16107, encoding the ortholog to RRP5, has activity in rRNA cleavage during 18S rRNA synthesis (Lebaron et al. 2013), and depletion reduces pre-rRNA compaction and ribosome activity. The fourth subunit of RNA Polymerase III (RPC), RPC4, might be represented by PITG_16116 due to the presence of the RPC4 domain in the last 70 amino acids of protein, though the function of the rest of the protein is unclear. A study in rice showed that loss of function in an RPC4 homolog induces sterility in pollen (Nguyen et al. 2017), though it is unclear at this time how *Phytophthora* might be affected. The ribosomal protein L13a, or PITG_01883, incorporates into the ribosome through rRNA binding. However, in humans, this protein seems to have evolved into having a redundant or dispensable function as evidenced by the lack of a phenotype after RNA interference (Chaudhuri et al. 2007). This may not be the case for *Phytophthora* if its redundant partners are missing from the genome, though the search has not yet been attempted.

As one of five essential components of the exosome, Rrp42/PITG_16138 contributes to 5.8S rRNA maturation (Mitchell et al. 1997) and functions as an exonuclease homologous to RNAse PH. Mitchell et al found that depletion of

any one of the five components of the exosome complex leads to growth defects due to immature rRNA buildup. PITG_16138 is potentially found in two locations on chromosome 3, about 400kb apart. This complicates its involvement in metalaxyl resistance, because resistant mutations would likely need to occur in both paralogs in order for a fully resistant phenotype to emerge. The second putative paralog starts at position 16,817,748, and originates from a 198 Kb contig (contig220) assembled at the very end of chr3 (**Chapter 2**). Given the small size of contig220, the difficulty of assembling contigs at the end of chromosomes, and an inconsistent spike in copy number (**Figure 3.5d**), there is a strong likelihood that contig220 is redundant to the assembly to at least some degree, and thus Rrp42 remains a candidate as a target for metalaxyl and the development of resistance. Further testing to confirm this could include an Illumina read depth or qPCR copy number assay.

For any potential causal genes, the mode of action likely involves the metalaxyl molecule interacting with and inhibiting the activity of the wild type protein product. Resistance results from a mutation that leads to a conformational change in the protein, blocking the interaction site. Importantly, however, the lack of connection between metalaxyl resistance level and overall phenotypic robustness (growth rate, sporulation, infection ability, **Table 3.1**) means that the resistance gain-of-function mutation should not affect the nominal activity of the protein. Another consideration is how the causal gene responds to

changes in dosage, specifically relating to the fast-growing sectors (**Figure 3.5d**). The addition of one (or more likely several tandemly repeated copies (Judelson et al. 1991)) of RPA190 conferred a strong growth rate penalty (Randall et al. 2014), and it would be reasonable to assume the deletion of a chromosome arm containing >200 genes might affect the fitness of the individual. However, progeny such as M65 appear to endure whole chromosome aneuploidies with minimal phenotypic consequence (**Table 3.1**, **Figure 3.7a**).

Probably the most important factor determining if any of these genes have the potential to underlie the metalaxyl resistance phenotype will be if they contain non-synonymous SNPs or small indels. Considering nonsynonymous polymorphisms heterozygous in parent 618 and homozygous in parent 1306 in the five genes discussed above; PITG_16054 contains one such nonsynonymous SNP, PITG_16107 contains 21, PITG_16116 contains none, PITG_01883 contains one, and PITG_16138 contains one. One simple way to test if any of these genes and nonsynonymous SNPs, except PITG_16116, contribute to resistance is to clone out the entire putative resistant version of the gene from one of the fast-growing sector individuals, and transform back into a sensitive isolate, likely parent 1306. If PITG_16107 rescues resistance, determining which nonsynonymous SNP is the causal mutation will be challenging due to the large number of SNPs and large gene size (>5 kb).

Genome plasticity variation

While oomycetes have a base genome copy number of two, it has been established that individual lineages or progeny can vary on the whole genome level (Bertier et al. 2013; Hamed and Gisi 2013; Li et al. 2017; Shrestha et al. 2017; Yoshida et al. 2013) and even at the individual chromosome level (Carter et al. 1999; van der Lee et al. 2004). Such variation is not unique to *Phytophthora*, and aneuploidy in many species is often thought to be a rapid adaptation to various stresses (Reviewed in Berman 2016). For example, the opportunistic human fungal pathogen *Cryptococcus neoformans* develops large polyploid cells called Titan cells in response to fungicide treatment and upon colonization of a receptive human host (Gerstein et al. 2015). Polyploidy can also be seen in human liver hepatocytes (Duncan et al. 2010) and cancerous tumors show varying levels aneuploidy in a nearly ubiquitous manner across tumor types (Gordon et al. 2012). By utilizing whole genome sequencing for a large number of progeny in this study coupled with our chromosome-scale reference genome, we can examine the true extent of genome and chromosome instability on a much finer scale than previously possible for any oomycete.

For 35 CrossM progeny and parents, and 34 Cross20 progeny and parents, 31.5% and 12.9% of the total number of chromosomes existed in a copy

number greater than or equal to three, respectively (**Figure 3.7**). Six and three progeny were predicted to be triploid and tetraploid based on a majority rules definition of ploidy state, since only one of these nine individuals was wholly euploid (M46, 4N). This indicates a substantial degree of nondisjunction occurring, with 175 of 555 individual chromosomes being in a copy number greater than or equal to 3 in Cross M. Comparatively few individual progeny were predicted to be wholly diploid across all chromosomes, with only 9 progeny showing all disomic chromosomes. For comparison, a previous study generated a cross between parents Ca65 and 550 resulting in 80 total offspring and performed limited copy-number analysis using 20 RFLP markers (Carter et al. 1999). Three progeny from that cross were determined to be wholly triploid and 24 were shown to have at least one trisomic marker/chromosome. If we consider each progeny in that cross to have the base number of 15 chromosomes, the three triploid progeny contributed 45 chromosomes. For the 24 diploid but trisomic progeny, a liberal estimate of five trisomies per progeny contributes an additional 120 chromosomes. This totals 165 out of 1,200 chromosomes in a trisomic state, or around 14% of the total, comparable to the Cross20 rate and about half the CrossM rate. This may suggest that CrossM is particularly vulnerable to nondisjunction, though only a repeat of the analysis presented here on the CA65x550 cross can settle that.

Cross20 exhibited much higher levels of chromosome deletion or translocation, as evidenced by the 0.5 copy numbers (**Figure 3.7**). This is mostly due to parent 6629 exhibiting a deletion on the end of chromosome 7 and parent 550 showing a deletion on chromosome 15. These two chromosomes appeared to segregate through the progeny, but unusual allele ratio profiles and an almost complete lack of variants expected to be heterozygous in the progeny lead to uncertainty in determining the true copy-number state (**Table 3.2**). Read depth could usually be used to supplement ambiguous allele ratio calls. Outside of chromosomes 7 and 15, sporadic chromosome arm duplications, or trisomic truncations, appeared more common than in CrossM, although no clear pattern seems apparent. The same underlying biological process could contribute to both CrossM having a higher nondisjunction rate and Cross20 having a higher truncation rate.

One of the most important structural variants identified is a deletion at the end of chromosome 3 in the four progeny growing spontaneously faster on metalaxyl media (**Figure 3.5d**). The deletion of the allele contributed by parent 1306 means that only the metalaxyl-resistance allele originating from parent 618 allele is present, inducing a sudden tolerance to metalaxyl. Among these four progeny, each had a slightly different breakpoint along the chromosome, though M78F and M123F had similar breakpoints. No other progeny had deletions in this region, further supporting the idea of a spontaneous deletion event selected

for by the metalaxyl amended media, perhaps due to some particularly repetitive regions near the breakpoints leading to a enhanced intrachromsomal recombination or an enrichment of TE insertion sites.

The four pairs of normal and fast growing sector progeny also showed structural variation outside of the end of chromosome 3. For example, chromosome 7 of M64N exists as a monosomy, based on decreased read depth and loss of heterozygosity (**Figure 3.7a,g**). Chromosome 7 of M64F, however, appears to exist in two copies, with both alleles originating from either parent (**Figure 3.7h**). This case shouldn't be possible if the sampled M64F nucleus was directly derived from the M64N nucleus. Thus, the mycelia of the plate that grew the fast sector existed as a heterokaryon of many different nucleus types, and the single zoospores isolated from the different sectors of the plate were related only through either the deletion or retention of the chromosome 3 arm and not by the behavior of other chromosomes. The last common nucleus between the two could have existed a great many generations in the past. If more samples were taken and examined, more variation in other chromosomes could potentially be observed. For the other three progeny pairs, more variation can be observed across other chromosomes, suggesting that instability is common even among relatively closely related nuclei. A better understanding of the temporal behavior of instability could involve growing a freshly single zoospore derived colony, and sampling and sub-culturing the colony after a set number of days. This way,

chromosome aberrations retained or even reverted through the subcultures or could be followed.

Transcriptional variation

To our knowledge, this is one of the first efforts to examine many individuals of a sexual population for an oomycete to the scale of whole-genome sequencing. As such, it is difficult to gauge if the levels of expression diversity (**Figure 3.8**) among the progeny are comparable to other species, or if *Phytophthora* sexual progeny show a tendency of increased diversity. Additionally, the isolation of RNA from entire mycelia, instead of a specific tissue such as sporangia, further complicates analysis. A good solution to this unresolved issue would be to generate inbred or isogenic lines of *Phytophthora infestans,* perform true F1 and F2 sexual populations similar to other model species, and extract from a single tissue type growing on a standard media. However, this seems unfeasible until a protocol is developed to isolate or induce haploid tissue and artificially double the cell back to the base diploid state. Another challenge we faced is an overall low amount of differential expression differences between individuals (**Figure 3.8b,d**), as log2 fold changes never rose above an absolute value of 3.0. This could be explained by our strategy of pooling diverse progeny when comparing the two experimental setups, and

utilizing more replicates and phenotypically similar progeny may reveal more genes with high degrees of differential expression. Despite these challenges, we observe several trends across our progeny.

First, the treatment of metalaxyl induces changes in gene expression, but individual genotype still affects the transcriptional profile to a greater degree (**Figure 3.8a**). Despite this, 308 differentially expressed genes (DEGs) were found to be when comparing progeny that have and have not been treated with metalaxyl regardless of metalaxyl sensitivity, compared to 120 DEGs when comparing resistant versus sensitive progeny regardless of metalaxyl treatment. Of the 308 treatment DEGs, the vast majority of them were upregulated in progeny exposed to metalaxyl (278) compared to their status in progeny naïve to metalaxyl (30). GO term analysis revealed that genes involved in rRNA binding and ribosomal components were highly enriched in genes upregulated in response to metalaxyl. One explanation for this enrichment could be that the metalaxyl molecule is still interfering with rRNA production, but the lack of rRNA/mRNA/protein product sends a signal to the cell to increase transcription of ribosome components in an effort to boost mRNA translation. Previous work has identified ATP-Binding Cassette (ABC) Transporters as playing roles in metalaxyl resistance in *Phytophthora* and *Pythium* (Childers et al. 2014; Judelson and Senthil 2006; Lévesque et al. 2010) and six ABC transporters were noticed to be significantly differentially expressed in response to metalaxyl. Five were

upregulated in the presence of metalaxyl (PITG_17047, PITG_17048, PITG_17049, PITG_16495, and PITG_05217) while only one was downregulated (PITG_07134). The substrates of these transporters cannot be predicted from structure alone, but if any of these are having an effect on resistance, the route of action likely involves the transporters shunting the metalaxyl molecule out of the cell, away from the RNA polymerases and ribosome machinery.

In addition to transporters and other ribosome-associated proteins, several RNA polymerase subunits were upregulated in response to metalaxyl. These include the second largest Pol I subunit, RPA135 (PITG_02420), the RPABC23 subunit shared with all three polymerases (PITG_12877), the RPC10 subunit of Pol III (PITG_16526), and RPAC40 shared between Pol I and III (PITG_18777). Interestingly, RPAC40, while present in three copies of the original Broad Institute assembly (Haas et al. 2009), is present in only one copy in our updated assembly. Finally, an RNA polymerase associated gene (PITG_06706) was also upregulated, but BLAST searching does not clarify to where it belongs functionally since orthologs in other species are equally unclear.

When considering only sensitive or only resistant progeny, the sensitive progeny consistently had many more DEGs when exposed to metalaxyl, likely due to their inherent susceptibility to the compound compared to the resistant individuals. Progeny 5.36 and parent 1306, both sensitive to metalaxyl,

demonstrate this in **Figure 3.8a**, where they are more affected by metalaxyl than the other resistant individuals.

Second, genes upregulated in resistant progeny often include sugar and carbohydrate transporters. Transporters are often thought of as a major route of resistance by the exclusion of toxic compounds, but in the case of genes upregulated in resistant progeny, carbohydrate transporter activity dominate GO term enrichments (**Table 3.8**). Additionally, some genes central to signaling and metabolism such as phospholipase D (PITG_06994) and alpha-amylase (PITG_05526) are upregulated in resistant progeny. It is possible that a boost in such genes primes individuals with a more functional metabolism, allowing them to overcome some of the stresses of metalaxyl. Once the causal mutation of the major locus of metalaxyl is identified, it would be interesting to examine intermediately-sensitive individuals who lack the major locus of resistance, and see if core-metabolic genes are naturally upregulated compared to fully sensitive individuals, thus granting some resistance. These upregulated genes have the potential to underlie some of the minor loci which quantitatively modify the major resistance locus.

Overall, we found a great diversity of transcriptional variation among our progeny, the origins of which could come from several sources. One route could be the contributions of aneuploidy, which can greatly affect the dosages and stoichiometric ratios of protein complex subunits, affecting function on substrates

(Birchler and Veitia 2012). However, one would expect an individual with a highly diverged karyotype such as progeny 7.13, and an individual with only a mildly diverged karyotype, such as progeny M88, to cluster far apart from each other in terms of transcriptional profiles (**Figure 3.7a**). Instead, those two individuals do cluster together, albeit with still dissimilar profiles. Another route is through chromatin remodeling, though little is known about chromatin dynamics of oomycetes compared to other taxa. Nevertheless, it seems reasonable to assume that histone modification marks and their affects would follow a similar profile to those of plants (Li et al. 2008).

Genome plasticity and transcriptional functional consequences

Overall, our sexual population shows considerable diversity in both structural and transcriptional variation. While this may be expected from outcrossed progeny as opposed to inbred F1 progeny, the sheer level of diversity suggests other factors are at play. A similarly structured study in *Populus spp* demonstrated that structural diversity can be generated and inherited in progeny, though the source of this diversity was gamma irradiation, as opposed to natural diversity (Henry et al. 2015). Dysfunction of the centromere histone variant H3 has also been linked to chromosome elimination and shattering (Tan et al. 2015). Another mechanism for this explosion in diversity could be a burst of TE activity,

induced by sexual recombination. Genome instability is a hallmark of pathogens (Bertier et al. 2013; Dong et al. 2016; Kasuga et al. 2012; Mannaert et al. 2012; Patel 2016) due to a need for rapid evolution in the plant-pathogen and chemical control evasion arms race. DNA methylation has long been known to be associated with inactive DNA in plants, and even has roles in suppressing TEs (Miura et al. 2001; Saze et al. 2012). While oomycetes do not exhibit 5mC methylation like plants, their analogous 6mA DNA methylation could conceivably function in similar roles of TE suppression (Chen et al. 2017).

Recombination seems to share a connection with TE activity (Kent et al. 2017; Sanchez et al. 2017), and TEs are often associated with genomic instability due to their tendency to induce double strand breaks (DSBs) following insertion (Hedges and Deininger 2007; McClintock 1984). Evidence for sexual recombination, as opposed to inter- or intrachromosomal recombination, driving TE bursts or demethylation is lacking, but it is conceivable that such processes may be connected in *Phytophthora.* Studies in the sudden oak pathogen *Phytophthora ramorum* have demonstrated that host-induced stress causes genome instability and changes to colony morphology due to TE bursts (Kasuga et al. 2012) and can lead to aneuploidies (Kasuga et al. 2016). One potential way to test this theory could be to generate *de novo* assemblies of the progeny Illumina reads. Searching these assemblies for known TE motifs or simple repeats might reveal more hits in progeny with highly distorted genome

structures, especially the fast-sector progeny, when compared to progeny with few aberrations. Additionally, contigs containing repeat annotations could be mapped back to the reference assembly, and it may be possible to determine if transposon-containing contigs map to and flank known chromosome truncations.

Demethylation following the formation of hybrid progeny could lead to increased transposon accumulation. This in turn could lead to translocations or truncations through intra or inter chromosomal recombination, or errant DSB repair (Pfeiffer et al. 2000). Additionally, if DSB repair proceeds through the synthesis-dependent-strand annealing (SSDA) pathway, the use of one allele as a template strand can result in copy-neutral gene conversion (Dwivedi and Haber 2018; Johnson and Jasin 2000), a phenomenon reported in other *Phytophthora* species (Chamnanpunt et al. 2001; Lamour et al. 2012). If transposon insertion and subsequent DSB repair can also induce gene conversion, or copy neutral loss of heterozygosity (LOH), it could explain unusual allele ratio pileups in the datasets of variant ratios expected to be heterozygous (**Figure 3.7k**). One final interesting observation is how some progeny do not seem as affected by genome instability as others (**Figure 3.7a,b**). In a sexual population of *P. capsici*, 23 of 60 outcrossed F_1 individuals showed LOH in sequences flanking at least two mapping markers, and two of these individuals showed LOH spanning over 10% of the genome (Lamour et al. 2012). This parallels our results in that while structural variation and LOH are not uncommon among the progeny, it is not

ubiquitous to all progeny. An explanation for this could be that the components of the methylation maintenance system in either parent contain some loss-of-function mutants, and improper combinations of these mutations could lead to decreased maintenance effectiveness in progeny inheriting these mutation combinations. This could also explain why the parents themselves appear to remain somewhat stable in terms of their genome stability if sexual recombination-derived demethylation is an incorrect theory. Thus a conceivable explanation for chromosome structure diversity in our sexual populations may originate in the loss of transposon suppression.

This theory does not, however, address the extent of whole chromosome aneuploidy, which has been shown in many systems to be a common response to environmental, host induced, and chemical stresses (Berman 2016; Gerstein et al. 2015; Kasuga et al. 2016; Selmecki et al. 2008; Selmecki et al. 2009). Chromosome nondisjunction is the likely source of aneuploidy, though an explanation as to why it is being manifested at a considerably increased rate in some progeny but not others remains unclear. In the human pathogen *Candida albicans*, a proposed mechanism for aneuploid chromosome acquisition involved treatment of the fungus with fluconazole, which indiced tetraploid cells with extra spindle components (Harrison et al. 2014). These unstable cells subsequently underwent chromosome loss due to unequal segregation leading to aneuploidy. It is conceivable that the various cell cycle checkpoints during meiosis or mitosis,

such as spindle fiber formation and attachment or the function of Hsp90 in chromosome segregation (Kaplan and Li 2012), have deleterious heterozygous mutations in the parents, and combinations of homozygous deleterious alleles in the progeny increase the rate of nondisjunction in a similar manner to how defects in methylation maintenance pathways may lead to higher TE activity. Additionally, we lack sufficient data to inform on the timeline of aberration accumulation. Genomic DNA was typically isolated shortly (~2 weeks) after the establishment of a colony derived from a single zoospore, but phenotypic testing was not usually performed concurrently. This means that nuclei containing lots of aberrations detected in the early gDNA might actually not be present when testing specific phenotypes. The fast-growing progeny sectors on metalaxyl media illustrate this, since mycelia nuclei not containing those truncations would be outcompeted and knowledge of their presence lost. An experiment similar to the one described above following the fast-growing progeny over several subcultures could inform on the rate of chromosome aberration if performed immediately after oospore colony establishment..

In conclusion, our analysis of two *Phytophthora infestans* sexual populations reveals several intriguing phenomena. First, resistance to the fungicide metalaxyl is mostly determined by a region at the end of chromosome 3. Several genes important for ribosome activity and rRNA maturation lie within this interval, and further testing may reveal one of them to be the second major

locus of metalaxyl resistance, after RPA190. Second, genome structure in our sexual populations had vary wildly, giving rise to triploid and tetraploid progeny. This in turn likely affects the expression patterns of genes lying on affected chromosomes, leading to complex and diverse transcriptomes between the progeny. Whether or not this level of diversity seen in our sexual populations is a feature largely unique to the oomycetes or filamentous fungi remains to be seen, though many pathogenic systems exploit genome instability to evade chemical control or other stressful environments (Dong et al. 2016; Gerstein et al. 2015; Jones et al. 2014; Selmecki et al. 2009). It seems likely that sexual progeny in other systems might similarly display genome instability, though even minor chromosomal truncations or duplications can lead to unviable offspring in most animals and severely reduced fitness in plants, and so a selection bias probably applies to the viable offspring we were able to recover and culture in *P. infestans*.

This unregulated pattern of aberration in sexual progeny allows for many new combinations of gene expression, a worrisome occurrence in relation to agricultural production of potato and tomato. New sexual progeny could potentially overcome established resistance genes and chemical treatments, and thus careful monitoring the co-occurrence of the sexually compatible A1 and A2 mating type lineages in the field will continue to be important in the future. It is important to realize, however, that the vast majority of progeny will likely never be able to outcompete the established clonal lineages due to the phenotypic

consequences of so much genome instability on key traits such as sporulation and pathogenicity (**Table 3.1**). Nevertheless, the quick turnover of dominant asexual US lineages, and their likely origin from sexual populations in central Mexico (Wang et al. 2017), indicate that sexual reproduction is still an important aspect of the global control of *Phytophthora infestans.*

References

Australian Commonwealth. 2014. Threat abatement plan for disease in natural ecosystems caused by *Phytophthora cinnamomi*.

Bai, B., Moore, H. M., and Laiho, M. 2013. CRM1 and its ribosome export adaptor NMD3 localize to the nucleolus and affect rRNA synthesis. Nucleus 4:315-325.

Barchenger, D. W., Lamour, K. H., Sheu, Z. M., Shrestha, S., Kumar, S., Lin, S. W., Burlakoti, R., and Bosland, P. W. 2017. Intra- and Intergenomic variation of Ploidy and Clonality characterize *Phytophthora capsici* on *Capsicum sp* in Taiwan. Mycolo Prog 16:955-963.

Benaglia, T., Chauveau, D., Hunter, D. R., and Young, D. S. 2009. mixtools: An R Package for Analyzing Mixture Models. 2009 32:29.

Berman, J. 2016. Ploidy plasticity: a rapid and reversible strategy for adaptation to stress. FEMS Yeast Res 16:fow020-fow020.

Bertier, L., Leus, L., D'Hondt, L., de Cock, A. W., and Hofte, M. 2013. Host adaptation and speciation through hybridization and polyploidy in *Phytophthora*. PLoS One 8:e85385.

Bhat, R. G., Mcblain, B. A., and Schmitthenner, A. F. 1993. The Inheritance of Resistance to Metalaxyl and to Fluorophenylalanine in Matings of Homothallic *Phytophthora Sojae*. Mycolo Res 97:865-870.

Birchler, J. A., and Veitia, R. A. 2012. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proc Natl Acad Sci USA 109:14746-14753.

Birchler, J. A., Johnson, A. F., and Veitia, R. A. 2016. Kinetics genetics: Incorporating the concept of genomic balance into an understanding of quantitative traits. Plant Sci 245:128-134.

Blum, M., Waldner, M., and Gisi, U. 2010. A single point mutation in the novel PvCesA3 gene confers resistance to the carboxylic acid amide fungicide mandipropamid in *Plasmopara viticola*. Fun Genet Biol 47:499-510.

Bower, L. A., and Coffey, M. D. 1985. Development of laboratory tolerance to phosphorous acid, fosetyl-Al, and metalaxyl in *Phytophthora capsici*. Can J of Pla Path 7:1-6.

Browning, S. R., and Browning, B. L. 2007. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 81:1084-1097.

Brurberg, M. B., Elameen, A., Le, V. H., Naerstad, R., Hermansen, A., Lehtinen, A., Hannukkala, A., Nielsen, B., Hansen, J., Andersson, B., and Yuen, J. 2011. Genetic analysis of *Phytophthora infestans* populations in the Nordic European countries reveals high genetic variability. Fun Biol 115:335-342.

Carter, D. A., Buck, K. W., Archer, S. A., Van der Lee, T., Shattock, R. C., and Shaw, D. S. 1999. The detection of nonhybrid, trisomic, and triploid offspring in sexual progeny of a mating of *Phytophthora infestans*. Fun Genet Biol 26:198-208.

Catal, M., King, L., Tumbalam, P., Wiriyajitsomboon, P., Kirk, W. W., and Adams, G. C. 2010. Heterokaryotic nuclear conditions and a heterogeneous nuclear population are observed by flow cytometry in *Phytophthora infestans*. Cytometry A 77A:769-775.

Chamnanpunt, J., Shan, W. X., and Tyler, B. M. 2001. High frequency mitotic gene conversion in genetic hybrids of the oomycete *Phytophthora sojae*. Proc Natl Acad Sci USA 98:14530-14535.

Chaudhuri, S., Vyas, K., Kapasi, P., Komar, A. A., Dinman, J. D., Barik, S., and Mazumder, B. 2007. Human ribosomal protein L13a is dispensable for canonical ribosome function but indispensable for efficient rRNA methylation. RNA 13:2224-2237.

Chen, F., Zhou, Q., Xi, J., Li, D.-l., Schnabel, G., and Zhan, J. 2018. Analysis of RPA190 revealed multiple positively selected mutations associated with metalaxyl resistance in *Phytophthora infestans*. Pest Man Sci 74:1916-1924.

Chen, H., Shu, H., Wang, L., Zhang, F., Li, X., Ochola, S., Mao, F., Ma, H., Ye, W., Gu, T., Jiang, L., Wu, Y., Wang, Y., Kamoun, S., and Dong, S. 2017. *Phytophthora* methylomes modulated by expanded 6mA methyltransferases are associated with adaptive genome regions. bioRxiv doi.org/10.1101/217646.

Childers, R., Danies, G., Myers, K. L., Fei, Z., Small, I. M., and Fry, W. 2014. Acquired resistance to mefenoxam in sensitive isolates of *Phytophthora infestans*. Phytopathology 105:342-349.

Cohen, Y., and Coffey, M. D. 1986. Systemic fungicides and the control of oomycetes. Ann Rev Phytopath 24:311-338.

Davidse, L. C., Hofman, A. E., and Velthuis, G. C. M. 1983. Specific interference of metalaxyl with endogenous RNA polymerase activity in isolated-nuclei from *Phytophthora-megasperma f. sp medicaginis*. Exp Mycol 7:344-361.

Davidse, L. C., Gerritsma, O. C. M., Ideler, J., Pie, K., and Velthuis, G. C. M. 1988. Antifungal modes of action of metalaxyl, cyprofuram, benalaxyl and oxadixyl in phenylamide-sensitive and phenylamide-resistant strains of *Phytophthora Megasperma* f. sp *medicaginis* and *Phytophthora infestans*. Crop Protection 7:347-355.

Delignette-Muller, M. L., and Dutang, C. 2015. fitdistrplus: An R Package for Fitting Distributions. J of Stat Soft 64:34.

Dong, Y., Li, Y., Qi, Z., Zheng, X., and Zhang, Z. 2016. Genome plasticity in filamentous plant pathogens contributes to the emergence of novel effectors and their cellular processes in the host. Curr Gen 62:47-51.

Dowley, L. J., and Osullivan, E. 1981. Metalaxyl resistant strains of *Phytophthora-Infestans* (Mont) de Bary in Ireland. Potato Res 24:417-421.

Drenth, A., Janssen, E. M., and Govers, F. 1995. Formation and survival of oospores of *Phytophthora infestans* under natural conditions. Plant Pathol 44:86-94.

Duncan, A. W., Taylor, M. H., Hickey, R. D., Hanlon Newell, A. E., Lenzi, M. L., Olson, S. B., Finegold, M. J., and Grompe, M. 2010. The ploidy conveyor of mature hepatocytes as a source of genetic variation. Nature 467:707.

Dwivedi, G., and Haber, J. E. 2018. Chapter Eight: Assaying mutations associated with gene conversion repair of a double strand break. Pages 145-160 in: Meth enzymology, vol. 601. M. Spies and A. Malkova, eds. Academic Press.

Fabritius, A. L., and Judelson, H. S. 1997. Mating-type loci segregate aberrantly in *Phytophthora infestans* but normally in *Phytophthora parasitica*: implications for models of mating-type determination. Curr Gen 32:60-65.

Fabritius, A. L., Shattock, R. C., and Judelson, H. S. 1997. Genetic analysis of metalaxyl insensitivity loci in *Phytophthora infestans* using linked DNA markers. Phytopathology 87:1034-1040.

Fisher, D. J., and Hayes, A. L. 1984. Studies of mechanisms of metalaxyl fungitoxicity and resistance to metalaxyl. Crop Protection 3:177-185.

Fong, A. M. V. A., and Judelson, H. S. 2003. Cell cycle regulator Cdc14 is expressed during sporulation but not hyphal growth in the fungus-like oomycete *Phytophthora infestans*. Mol Micro 50:487-494.

Frenkel, O., Cadle-Davidson, L., Wilcox, W. F., and Milgroom, M. G. 2015. Mechanisms of resistance to an azole fungicide in the grapevine powdery mildew fungus, *Erysiphe necator*. Phytopathology 105:370-377.

Frost, C., and Dowley, L. 1979. Metalaxyl-a new fungicide. Technical Bulletin, agriculture series, an foras taluntais, Dublin 4 3:1-6.

Gel, B., and Serra, E. 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33:3088-3090.

Gerstein, A. C., Fu, M. S., Mukaremera, L., Li, Z., Ormerod, K. L., Fraser, J. A., Berman, J., and Nielsen, K. 2015. Polyploid titan cells produce haploid and aneuploid progeny to promote stress adaptation. MBio 6:e01340-01315.

Gisi, U., and Cohen, Y. 1996. Resistance to phenylamide: a case study with *Phytophthora infestans* involving mating type and race structure. Ann Rev Phytopath 34:549-572.

Gómez-Rubio, V. 2017. ggplot2 - Elegant Graphics for Data Analysis (2nd Edition). J Stat Soft 77:3.

Gordon, D. J., Resio, B., and Pellman, D. 2012. Causes and consequences of aneuploidy in cancer. Nat Rev Genet 13:189-203.

Gordon, T. R. 2017. *Fusarium oxysporum* and the *Fusarium* wilt syndrome. Ann Rev Phytopath 55:23-39.

Guo, M., Davis, D., and Birchler, J. A. 1996. Dosage effects on gene expression in a maize ploidy series. Genetics 142:1349-1355.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., 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., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., 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., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Hamed, B. H., and Gisi, U. 2013. Generation of pathogenic F1 progeny from crosses of *Phytophthora infestans* isolates differing in ploidy. Plant Pathol 62:708-718.

Harrison, B. D., Hashemi, J., Bibi, M., Pulver, R., Bavli, D., Nahmias, Y., Wellington, M., Sapiro, G., and Berman, J. 2014. A tetraploid intermediate precedes aneuploid formation in yeasts exposed to fluconazole. PLoS Biol 12:e1001815.

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Visser, R. G. F., and van der Vossen, E. A. G. 2008. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. Potato Res 51:47-57.

Hedges, D. J., and Deininger, P. L. 2007. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. Mut Res and Mol Mech of Mutagen 616:46-59.

Henry, I. M., Zinkgraf, M. S., Groover, A. T., and Comai, L. 2015. A system for sosage-based functional genomics in poplar. Plant Cell 27:2370-2383.

Hohn, T. 2013. Plant pararetroviruses: interactions of cauliflower mosaic virus with plants and insects. Curr Opin Virol 3:629-638.

Jiang, R. H., and Tyler, B. M. 2012. Mechanisms and evolution of virulence in oomycetes. Ann Rev Phytopath 50:295-318.

Johnson, R. D., and Jasin, M. 2000. Sister chromatid gene conversion is a prominent double-strand break repair pathway in mammalian cells. EMBO J 19:3398-3407.

Jones, L., Riaz, S., Morales-Cruz, A., Amrine, K. C., McGuire, B., Gubler, W. D., Walker, M. A., and Cantu, D. 2014. Adaptive genomic structural variation in the grape powdery mildew pathogen, *Erysiphe necator*. BMC Geno 15:1081.

Judelson, H. S., and Yang, G. 1998. Recombination pathways in *Phytophthora infestans*: polyploidy resulting from aberrant sexual development and zoospore-mediated heterokaryosis. Mycolo Res 102:1245-1253.

Judelson, H. S., and Roberts, S. 1999. Multiple loci determining insensitivity to phenylamide fungicides in P*hytophthora infestans*. Phytopathology 89:754-760.

Judelson, H. S., and Senthil, G. 2006. Investigating the role of ABC transporters in multifungicide insensitivity in *Phytophthora infestans*. Mol Plan Path 7:17-29.

Judelson, H. S., Tyler, B. M., and Michelmore, R. W. 1991. Transformation of the oomycete pathogen, *Phytophthora infestans*. Mol Plan Micro Interac 4:602-607.

Kamoun, S., Furzer, O., Jones, J. D., Judelson, H. S., Ali, G. S., Dalio, R. J., Roy, S. G., Schena, L., Zambounis, A., Panabieres, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X. R., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B. M., Grunwald, N. J., Mukhtar, M. S., Tome, D. F., Tor, M., Van Den Ackerveken, G., McDowell, J., Daayf, F., Fry, W. E., Lindqvist-Kreuze, H., Meijer, H. J., Petre, B., Ristaino, J., Yoshida, K., Birch, P. R., and Govers, F. 2015. The Top 10 oomycete pathogens in molecular plant pathology. Mol Plan Path 16:413-434.

Kaplan, K. B., and Li, R. 2012. A prescription for ‘stress’ – the role of Hsp90 in genome stability and cellular adaptation. Trends Cell Biol 22:576-583.

Kasuga, T., Kozanitas, M., Bui, M., Hüberli, D., Rizzo, D. M., and Garbelotto, M. 2012. Phenotypic diversification is associated with host-induced transposon derepression in the Sudden Oak Death pathogen *Phytophthora ramorum*. PLoS One 7:e34728.

Kasuga, T., Bui, M., Bernhardt, E., Swiecki, T., Aram, K., Cano, L. M., Webber, J., Brasier, C., Press, C., Grünwald, N. J., Rizzo, D. M., and Garbelotto, M. 2016. Host-induced aneuploidy and phenotypic diversification in the Sudden Oak Death pathogen *Phytophthora ramorum*. BMC Geno 17:385.

Kent, T. V., Uzunović, J., and Wright, S. I. 2017. Coevolution between transposable elements and recombination. Philo Transof the Royal Society: Biol Sci 372:20160458.

Lamour, K. H., and Hausbeck, M. K. 2000. Mefenoxam insensitivity and the sexual stage of *Phytophthora capsici* in Michigan cucurbit fields. Phytopathology 90:396-400.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S., Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. Mol Plan Micro Interac 25:1350-1360.

Layton, A. C., and Kuhn, D. N. 1988. Heterokaryon formation by protoplast fusion of drug-resistant mutants in *Phytophthora megasperma f.*sp. *glycinea*. Exp Mycol 12:180-194.

Lebaron, S., Segerstolpe, A., French, S. L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A. L., Wieslander, L., and Tollervey, D. 2013. Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. Mol Cell 52:707-719.

Lee, T. Y., Mizubuti, E., and Fry, W. E. 1999. Genetics of metalaxyl resistance in *Phytophthora infestans*. Fun Genet Biol 26:118-130.

Lévesque, C. A., Brouwer, H., Cano, L., Hamilton, J. P., Holt, C., Huitema, E., Raffaele, S., Robideau, G. P., Thines, M., Win, J., Zerillo, M. M., Beakes, G. W., Boore, J. L., Busam, D., Dumas, B., Ferriera, S., Fuerstenberg, S.

I., Gachon, C. M., Gaulin, E., Govers, F., Grenville-Briggs, L., Horner, N., Hostetler, J., Jiang, R. H., Johnson, J., Krajaejun, T., Lin, H., Meijer, H. J., Moore, B., Morris, P., Phuntmart, V., Puiu, D., Shetty, J., Stajich, J. E., Tripathy, S., Wawra, S., van West, P., Whitty, B. R., Coutinho, P. M., Henrissat, B., Martin, F., Thomas, P. D., Tyler, B. M., De Vries, R. P., Kamoun, S., Yandell, M., Tisserat, N., and Buell, C. R. 2010. Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. Genome Biol 11:R73.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.

Li, X., Wang, X., He, K., Ma, Y., Su, N., He, H., Stolc, V., Tongprasit, W., Jin, W., Jiang, J., Terzaghi, W., Li, S., and Deng, X. W. 2008. High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between dna methylation, histone methylation, and gene expression. Plant Cell 20:259-276.

Li, Y., Shen, H., Zhou, Q., Qian, K., van der Lee, T., and Huang, S. W. 2017. Changing ploidy as a strategy: the Irish potato famine pathogen shifts ploidy in relation to its sexuality. Mol Plan Micro Interac 30:45-52.

Ma, L. J., van der Does, H. C., Borkovich, K. A., Coleman, J. J., Daboussi, M. J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P. M., Kang, S., Shim, W. B., Woloshuk, C., Xie, X., Xu, J. R., Antoniw, J., Baker, S. E., Bluhm, B. H., Breakspear, A., Brown, D. W., Butchko, R. A., Chapman, S., Coulson, R., Coutinho, P. M., Danchin, E. G., Diener, A., Gale, L. R., Gardiner, D. M., Goff, S., Hammond-Kosack, K. E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C. D., Koehrsen, M., Kumar, L., Lee, Y. H., Li, L., Manners, J. M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S. Y., Proctor, R. H., Regev, A., Ruiz-Roldan, M. C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D. C., Turgeon, B. G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C. A., Kistler, H. C., and Rep, M. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367-373.

Malcolmson, J. F., and Black, W. 1966. New r genes in *Solanum demissum lindl* and their complementary races of *Phytophthora infestans* (mont) de bary. Euphytica 15:199-203.

Mannaert, A., Downing, T., Imamura, H., and Dujardin, J. C. 2012. Adaptive mechanisms in pathogens: universal aneuploidy in *Leishmania*. Trends Parasit 28:370-376.

Matson, M. E. H., Small, I. M., Fry, W. E., and Judelson, H. S. 2015. Metalaxyl resistance in *Phytophthora infestans*: assessing role of RPA190 gene and diversity within clonal lineages. Phytopathology 105:1594-1600.

McClintock, B. 1984. The significance of responses of the genome to challenge. Science 226:792-801.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. 1997. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell 91:457-466.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. 2001. Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature 411:212-214.

Montes, M. S., Nielsen, B. J., Schmidt, S. G., Bødker, L., Kjøller, R., and Rosendahl, S. 2016. Population genetics of *Phytophthora infestans* in Denmark reveals dominantly clonal populations and specific alleles linked to metalaxyl-M resistance. Plant Pathol 65:744-753.

Nguyen, G. N., Yamagata, Y., Shigematsu, Y., Watanabe, M., Miyazaki, Y., Doi, K., Tashiro, K., Kuhara, S., Kanamori, H., Wu, J., Matsumoto, T., Yasui, H., and Yoshimura, A. 2017. Duplication and loss of function of genes encoding RNA polymerase III subunit C4 causes hybrid incompatibility in rice. G3 7:2565-2575.

Patel, S. 2016. Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. Infect Genet Evol 45:151-164.

Pfeiffer, P., Goedecke, W., and Obe, G. 2000. Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. Mutagenesis 15:289-302.

Pscheidt, J. W., and Ocamb, C. M. 1999. Pacific Northwest plant disease control handbook. Oregon State University Press.

Quesada-Ocampo, L. M., and Hausbeck, M. K. 2010. Resistance in tomato and wild relatives to crown and root rot caused by *Phytophthora capsici*. Phytopathology 100:619-627.

Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., Jiang, R. H., Zody, M. C., Kunjeti, S. G., Donofrio, N. M., Meyers, B. C., Nusbaum, C., and Kamoun, S. 2010. Genome evolution following host jumps in the Irish potato famine pathogen lineage. Science 330:1540-1543.

Randall, E., Young, V., Sierotzki, H., Scalliet, G., Birch, P. R., Cooke, D. E., Csukai, M., and Whisson, S. C. 2014. Sequence diversity in the large subunit of RNA polymerase I contributes to mefenoxam insensitivity in *Phytophthora infestans*. Mol Plan Path 15:664-676.

Reuveni, M., Eyal, H., and Cohen, Y. 1980. Development of resistance to metalaxyl in *Pseudoperonospora cubensis*. Plant Dis 64:1108-1109.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.

Sanchez, D. H., Gaubert, H., Drost, H.-G., Zabet, N. R., and Paszkowski, J. 2017. High-frequency recombination between members of an LTR retrotransposon family during transposition bursts. Nat Comm 8:1283.

Saville, A., Pearce, C., and Ristaino, J. B. 2012. USABlight and fungicide sensitivity of recent genotypes of *Phytophthora infestans* to oomycete-targeted compounds. Phytopathology 102:106-106.

Saze, H., Tsugane, K., Kanno, T., and Nishimura, T. 2012. DNA methylation in plants: relationship to small rnas and histone modifications, and functions in transposon inactivation. Plant Cell Physi 53:766-784.

Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., and Berman, J. 2008. An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. Mol Micro 68:624-641.

Selmecki, A. M., Dulmage, K., Cowen, L. E., Anderson, J. B., and Berman, J. 2009. Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet 5:e1000705.

Shattock, R. C. 1988. Studies on the Inheritance of resistance to metalaxyl in *Phytophthora infestans*. Plant Pathol 37:4-11.

Sheltzer, J. M., and Amon, A. 2011. The aneuploidy paradox: costs and benefits of an incorrect karyotype. Trends Genet 27:446-453.

Shrestha, S. K., Miyasaka, S. C., Shintaku, M., Kelly, H., and Lamour, K. 2017. *Phytophthora colocasiae* from Vietnam, China, Hawaii and Nepal: intra- and inter-genomic variations in ploidy and a long-lived, diploid Hawaiian lineage. Mycol Prog 16:893-904.

Tan, E. H., Henry, I. M., Ravi, M., Bradnam, K. R., Mandakova, T., Marimuthu, M. P., Korf, I., Lysak, M. A., Comai, L., and Chan, S. W. 2015. Catastrophic chromosomal restructuring during genome elimination in plants. Elife 4:e06516.

TW, H. B., and Girke, T. 2016. systemPipeR: NGS workflow and report generation environment. BMC Bioinfo 17:388.

Urech, P. A., Schwinn, F., and Staub, T. 1977. CGA 48988, a novel fungicide for the control of late blight, downy mildews and related soil-borne diseases. Proc Br Crop Prot Conf Pests & Dis. 2:623–632.

van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., and DePristo, M. A. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Current Protoc Bioinf 11:1-33.

van der Lee, T., Testa, A., van 't Klooster, J., van den Berg-Velthuis, G., and Govers, F. 2001. Chromosomal deletion in isolates of *Phytophthora infestans* correlates with virulence on R3, R10, and R11 potato lines. Mol Plan Micro Interac 14:1444-1452.

van der Lee, T., Testa, A., Robold, A., van 't Klooster, J., and Govers, F. 2004. High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics 167:1643-1661.

Vleeshouwers, V. G., Raffaele, S., Vossen, J. H., Champouret, N., Oliva, R., Segretin, M. E., Rietman, H., Cano, L. M., Lokossou, A., Kessel, G., Pel, M. A., and Kamoun, S. 2011. Understanding and exploiting late blight resistance in the age of effectors. Ann Rev Phytopath 49:507-531.

Wang, J., Fernandez-Pavia, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodriguez-Alvarado, G., Grunwald, N. J., and Goss, E. M. 2017. High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. Mol Ecol 26:1091-1107.

Whittaker, S. L. S., R.C.; Shaw, D.S. . 1991. Variation in DNA content of nuclei of *Phytophthora infestans* as measured by a microfluorimetric method using the fluorochrome DAPI. Mycolo Res 95:602-610.

Xin, X. F., and He, S. Y. 2013. *Pseudomonas syringae* pv. tomato DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. Ann Rev Phytopath 51:473-498.

Xu, S. 2013. Mapping quantitative trait loci by controlling polygenic background effects. Genetics 195:1209-1222.

Yoshida, K., Schuenemann, V. J., Cano, L. M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F. N., Kamoun, S., Krause, J., Thines, M., Weigel, D., and Burbano, H. A. 2013. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. Elife 2:e00731.

Conclusions

A hallmark of pathogens is their ability to rapidly adapt to dynamic and stressful conditions (Dong et al. 2015; Patel 2016; Selmecki et al. 2009). Their comparatively short generation times compared to plants and higher animals allows them to test vast numbers of adaptations, either through acquired mutations or through genome structure instability. This rapid diversification is crucial to the success of pathogens, and the arms race between host and pathogen will likely never cease. While much is known about the mechanisms influencing how pathogens can adapt to stresses in regards to genome instability (Berman 2016), few studies of sexual populations of pathogens have addressed genome instability, which has the potential to incur an even more rapid burst of diversification in pathogen species which contain a sexual cycle in nature, such as the oomycetes.

The oomycetes are a diploid group of filamentous organisms that resemble fungi in lifestyle, but actually belong to the Stramenopile/Heterokont kingdom of life (Gunderson et al. 1987). While some oomycetes live as saprotrophs, many live parasitically and can infect a diverse range of organisms. Examples of oomycetes include *Saprolegnia parasitica*, which is a major production issue in fish aquaculture (Hatai and Hoshiai 1992), *Pythium insidiosum*, which causes pythiosis in mammals (De Cock et al. 1987), and

Phytophthora ramorum, which currently causing a major decline of oak species on the forested North American west coast due to sudden oak death (Rizzo et al. 2005). Arguably though, *Phytophthora infestans,* infamous for its role in the Irish potato famine of the 1840s, is the most damaging plant pathogen in history (Haverkort et al. 2008; Kamoun et al. 2015) and is still a major pest to this day.

Despite being a narrow host range pathogen, the destructiveness of *P. infestans* comes from several factors, including the extreme rapidness of its spread given ideal environmental conditions, a lack of effective resistance genes bred into its tomato and potato hosts, and the ability to quickly overcome control pressures originating from both host defensive genes and chemical control applications. However, chemical control remains as the most effective preventative and reactionary method for dealing with outbreaks and epidemics. While some oomycetes, such as the cucurbit pathogen *Phytophthora capsici,* are observed to readily undergo sexual reproduction in nature (Carlson et al. 2017; Lamour and Hausbeck 2000; Lamour et al. 2012), *P. infestans* usually proliferates as asexual lineages (Hu et al. 2012; Saville et al. 2016; Yoshida et al. 2013). These asexual lineages throughout North America are thought to originate from sexual metapopulations in Mexico (Wang et al. 2017) and a trend toward triploidy seems common in most lineages (Li et al. 2017). While sexual populations are uncommon in nature (Brurberg et al. 2011; Danies et al. 2014;

Yuen and Andersson 2013; Zhu et al. 2015), their presence has the potential to introduce new diversity and severely challenge existing systems of control.

This dissertation explores two major topics in *P. infestans* biology, namely the basis of resistance to the fungicide metalaxyl, and the extent of genome instability in sexual populations. Insight into both of these areas were aided by the construction of a new chromosome-scale reference assembly. Chapter I described a sexual population that was performed between metalaxyl-resistant parent 618, and metalaxyl sensitive parent 1306. This particular cross had been performed previously, but was expanded to a total of 130 members with the current iteration (CrossM)(Fabritius et al. 1997; Matson et al. 2015). Resistance is usually observed to segregate as a single major locus in an incompletely dominant manner, a trend that was again seen in this cross. Notably through, resistance segregation was unlinked to the segregation of single nucleotide polymorphism (SNP) markers within the RPA190 gene, which was previously shown to confer resistance in some UK populations of *P. infestans* (Randall et al. 2014). Instead, we found linkage to a genetically separate SNP marker, derived from a previously described randomly amplified polymorphic DNA (RAPD) marker (Judelson and Roberts 1999). Our genetic evidence presents a compelling argument for a second major locus of resistance, separate from RPA190 in parent 618. Chapter I also examines diversity in RPA190 genotypes across isolates belonging to the contemporary US-23 and US-24 clonal lineages

and their relationship to metalaxyl resistance. A key SNP implicated by another laboratory in resistance, T1145A, appeared to not correlate with resistance level at all in our isolates, although other interacting mutations in RPA190 could still contribute to resistance in these populations (Chen et al. 2018).

Chapter II undertook an effort to construct a chromosome-scale reference genome for *P. infestans*. Motivations for this were twofold: first, our inability to precisely map the metalaxyl resistance trait in the previously available and incompletely assembled reference genome from the Broad Institute (Haas et al. 2009) meant that a more contiguous assembly would be required to perform effective quantitative trail locus (QTL) analysis. Second, a global survey of structural variation would be possible with a more complete assembly, and instances of chromosome-scale aneuploidy could be observed. To accomplish this goal, we used the approach of Illumina whole-genome sequencing 84 parents and progeny from CrossM and 34 parents (6629x550) and progeny from Cross20 (Judelson et al. 1995) to generate four individual genetic maps corresponding to each parent. We also collaborated with Dr. Stefano Lonardi at UCR to develop a superior input assembly to be incorporated with the genetic maps. This input scaffold was generated through PacBio sequencing followed by *de novo* assembly, and two scaffolding technologies contributing to the improvement of the PacBio assembly: a Dovetail Genomics Chicago library and a Bionano genomics optical map. Most challenges to this approach involved

removing and correcting genotyping errors caused by the relatively shallow sequencing read coverage. However, other challenges such as aneuploid progeny and contig redundancy contributed to issues with our assembly. Despite this, we were able to confidently anchor 218 of the total 248 Mb (88.5%) assembly sequence into 15 pseudochromosome scaffolds, of which 153 Mb was unambiguously oriented. While this is not the first instance of an oomycete reference genome being assembled to the chromosome-scale, it is the first done using third-generation sequencing technology, and thus likely is of high quality compared to other oomycete references. We were also able to locate gene models originating from the original Broad assembly, and subsequently improved by our laboratory to a total of 18,672 genes (Shrivistava 2017), and generated 20,030 mapped genes, 19,013 of which reside on the assembled pseudochromosomes. This annotation, however, should be further investigated due to the 7,397 genes present in the original annotations but are lacking loci on the new assembly, and 3,149 present in two more copies. A more robust approach, using a program such as Maker (Cantarel et al. 2008) could substantially improve the annotations and further reveal the relationship between gene and repeat locations and how they relate to structural diversity between isolates or progeny.

Finally, chapter III utilized this new reference assembly to examine the extent of genetic loci underlying four major phenotypic traits, major structural

diversity among the progeny, as well as their transcriptomic diversity. To follow up on chapter I, we employed QTL analysis on the 82 CrossM progeny that were Illumina-sequenced, and found a 7-8 Mb interval of significant linkage (LOD > 11.5) at the end of chromosome three. Interestingly, this interval also coincided with a major hemizygous truncation of chromosome 3 in four individuals that spontaneously grew faster on media amended with metalaxyl. This truncation was shown to have preferentially deleted the allele that originated from resistant parent 618 through a combined qPCR and high resolution melt (HRM) approach. While metalaxyl resistance is thought to be associated with RNA Polymerase I (RPA) activity via temperance of rRNA elongation or maturation (Davidse et al. 1988; Davidse et al. 1983; Randall et al. 2014), none of the RPA subunits could be mapped to chromosome 3. However, 130 genes lie within the interval defined by the truncation, and four in particular show annotations relating to ribosome function and rRNA maturation. These could be excellent targets for follow-up. Aside from metalaxyl resistance, we also generated phenotypes for three other agriculturally important traits: overall radial growth rate, sporulation proficiency, and pathogenicity on potato tubers. However, none of these phenotypes revealed any significant QTL in the genome, though this is unsurprising given that these traits are likely controlled by many individual genes each contributing to the metabolic network which affects the final phenotype.

Consistent with the notion that pathogens can display considerable levels of genome plasticity (Barchenger et al. 2017; Berman 2016; Li et al. 2017; Patel 2016), cases of progeny showing chromosomal copy numbers varying from one to four were quite common in both CrossM and Cross20. Three putatively tetraploid and six triploid progeny were identified in CrossM, and three triploid progeny were seen in Cross20. Only nine CrossM progeny were wholly diploid, though this becomes 12 if including progeny with only a single trisomy, and of the 555 total possible chromosomes across the 37 parents and progeny, 173 (31.2%) were in a copy number greater than or equal to three. Cross20 contained fewer whole-chromosome aneuploidies, with only 65 of 510 total chromosomes existing in a copy number greater than or equal to three. However, chromosomal deletions in both parents 6629 and 550 segregated through the progeny, and the rate of truncations was higher than in CrossM as well.

This impressive rate of nondisjunction appears to be a unique finding due to the lack of similar studies in other systems. While aneuploidy can be a common response to environmental stresses (Berman 2016), it is not clear if the level seen in our populations is due to sexual recombination or if they would have naturally diverged in copy number given stresses such as prolonged exposure to metalaxyl, growth on a plant host, or growth on a minimal nutrient-depleted media (Hall 1959). We have already shown that progeny growing on metalaxyl

amended media have the potential to lose chromosome arms to gain an advantage, and future studies of both sexual and field-collected individuals growing in such conditions, coupled with a strict temporal collection schedule, may be able to resolve the frequency of both chromosomal arm duplication/loss and whole chromosome aneuploidy.

The ultimate source of the observed genome plasticity is unclear at this time, though a plausible explanation may center on the activity of transposable elements (TEs)(Faino et al. 2016). *P. infestans* is unusual among the oomycetes in that it has a relatively large (238 Mb) and repeat-rich genome (> 74% repeats) compared to other *Phytophthora* genomes (Haas et al. 2009). Given that TEs often contribute to bloated genome size (Chénais et al. 2012; Raffaele and Kamoun 2012), it is conceivable that *P. infestans* contains only a partially repressed repertoire of TEs. The sudden oak death pathogen, *Phytophthora ramorum*, has been shown to undergo TE induced genome instability (Kasuga et al. 2016; Kasuga et al. 2012) and our results here could follow a similar mechanism. In general, whole chromosome aneuploidy appeared to be a more common manifestation of genome instability in our CrossM population, while chromosome truncation was more common in the Cross20 population.

One route of TE suppression may include 6mA methylation (Chen et al. 2017) in the place of 5mC methylation, which *Phytophthora* lacks, but which is a component of TE repression in plants (Miura et al. 2001; Saze et al. 2012).

Thus, a plausible route of genome instability could be that each parent contains deleterious heterozygous mutations in one or more genes responsible for methylation maintenance, and some sexual progeny inherit a homozygous deleterious gene, triggering TE activity and structural aberrations through the double strand breaks induced by TE activity (Hedges and Deininger 2007). Since some of these genes may confer a lethal phenotype if loss of function, this hypothesis would probably predict reduced function mutations as opposed to loss-of-function. This could also explain why some progeny seem unaffected by genome instability, since they show few aberrations in chromosome structure. One route of aneuploidy, as opposed to chromosome structure instability, can originate from defects in spindle assembly or mutations in the centromere specific histone variants (Tan et al. 2015), though little is known of *Phytophthora* centromere biology. Aneuploidy is a common response to environmental stresses (Berman 2016; Kaplan and Li 2012; Kasuga et al. 2016; Selmecki et al. 2009), and additional copies of the targets of fungicides can increase resistance through sequestering of the compound, and more copies of enzymes efluxing the fungicide, such as ABC transporters, reduce the intracellular concentration of the fungicide (Judelson and Senthil 2006; Selmecki et al. 2008).

Our final contribution is the examination of genome-wide transcriptional variation in 34 CrossM parents and progeny. Genes differentially expressed between two main comparisons were sought: genes differentially expressed in

individuals growing in the presence versus absence of metalaxyl, and between resistant and sensitive progeny. Interestingly, progeny genotype seemed to contribute to transcriptional variation much more than the environment (metalaxyl or no metalaxyl) since no single genotype clustered to any other genotype except for its environmental counterpart. Unsurprisingly though, metalaxyl-sensitive progeny were more responsive to transcriptional shifts due to the presence of metalaxyl. In total, we found 120 genes differentially expressed between sensitive and resistant progeny, regardless of growth condition, and 308 genes differentially expressed between individuals grown with and without metalaxyl present. Sensitive individuals grown in the presence often showed enrichments in rRNA, ABC transporter, and ribosome biogenesis functions. While these kinds of annotations appear to be associated with the route of action for metalaxyl, it seems unlikely that these genes would be upregulated in a metalaxyl-sensitive individual. Instead, the stress that metalaxyl impacts on the cell functions could send a signal boosts ribosome transcription in an attempt to make up for reduced translation.

Overall, this dissertation provides insight into the *P. infestans* genome structure on a level not previously possible. Our chromosome-scale reference assembly granted us the ability to perform effective QTL analysis on an important agricultural trait, the resistance to the fungicide metalaxyl, and we discovered a region strongly associated with resistance at the end of chromosome 3.

Additionally, we observed a considerable level of genome plasticity among our progeny, which in turn likely contributed to considerable phenotypic variation observed in the transcriptomes, which finally manifested in our diverse phenotypes observed across four traits. This reveals the great potential of wild *P. infestans* sexual populations to induce and experiment with a great level of phenotypic diversity. While many, if not, all progeny in this study probably will not be competitive if released into nature, the sheer number of sexual oospores that could potentially be generated and released into the environment presents a considerable risk if one of the new phenotypic features of a sexual progeny is an ability to overcome established control measures such a resistance genes or chemical controls. Thus, monitoring of contemporary lineage distribution and traits, such as when lineages of opposite mating types are found in close proximity, will continue to be crucially important to the global fight against late blight.

References

Barchenger, D. W., Lamour, K. H., Sheu, Z. M., Shrestha, S., Kumar, S., Lin, S. W., Burlakoti, R., and Bosland, P. W. 2017. Intra- and Intergenomic variation of Ploidy and Clonality characterize *Phytophthora capsici* on *Capsicum sp* in Taiwan. Mycolo Prog 16:955-963.

Berman, J. 2016. Ploidy plasticity: a rapid and reversible strategy for adaptation to stress. FEMS Yeast Res 16:fow020-fow020.

Brurberg, M. B., Elameen, A., Le, V. H., Naerstad, R., Hermansen, A., Lehtinen, A., Hannukkala, A., Nielsen, B., Hansen, J., Andersson, B., and Yuen, J. 2011. Genetic analysis of *Phytophthora infestans* populations in the Nordic European countries reveals high genetic variability. Fun Biol 115:335-342.

Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., and Yandell, M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188-196.

Carlson, M. O., Gazave, E., Gore, M. A., and Smart, C. D. 2017. Temporal Genetic Dynamics of an Experimental, Biparental Field Population of *Phytophthora capsici*. Front in Gen 8.

Chen, F., Zhou, Q., Xi, J., Li, D.-l., Schnabel, G., and Zhan, J. 2018. Analysis of RPA190 revealed multiple positively selected mutations associated with metalaxyl resistance in *Phytophthora infestans*. Pest Man Sci 74:1916-1924.

Chen, H., Shu, H., Wang, L., Zhang, F., Li, X., Ochola, S., Mao, F., Ma, H., Ye, W., Gu, T., Jiang, L., Wu, Y., Wang, Y., Kamoun, S., and Dong, S. 2017. *Phytophthora* methylomes modulated by expanded 6mA methyltransferases are associated with adaptive genome regions. bioRxiv doi.org/10.1101/217646.

Chénais, B., Caruso, A., Hiard, S., and Casse, N. 2012. The impact of transposable elements on eukaryotic genomes: From genome size increase to genetic adaptation to stressful environments. Gene 509:7-15.

Danies, G., Myers, K., Mideros, M. F., Restrepo, S., Martin, F. N., Cooke, D. E., Smart, C. D., Ristaino, J. B., Seaman, A. J., Gugino, B. K., Grunwald, N. J., and Fry, W. E. 2014. An ephemeral sexual population of *Phytophthora infestans* in the Northeastern United States and Canada. PLoS One 9:e116354.

Davidse, L. C., Hofman, A. E., and Velthuis, G. C. M. 1983. Specific interference of metalaxyl with endogenous RNA polymerase activity in isolated-nuclei from *Phytophthora-megasperma f. sp medicaginis*. Exp Mycol 7:344-361.

Davidse, L. C., Gerritsma, O. C. M., Ideler, J., Pie, K., and Velthuis, G. C. M. 1988. Antifungal modes of action of metalaxyl, cyprofuram, benalaxyl and oxadixyl in phenylamide-sensitive and phenylamide-resistant strains of *Phytophthora Megasperma* f. sp *medicaginis* and *Phytophthora infestans*. Crop Protection 7:347-355.

De Cock, A. W., Mendoza, L., Padhye, A. A., Ajello, L., and Kaufman, L. 1987. *Pythium insidiosum* sp. nov., the etiologic agent of pythiosis. J Clin Microbiol 25:344-349.

Dong, S., Raffaele, S., and Kamoun, S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. Curr Opin Gen Dev 35:57-65.

Fabritius, A. L., Shattock, R. C., and Judelson, H. S. 1997. Genetic analysis of metalaxyl insensitivity loci in *Phytophthora infestans* using linked DNA markers. Phytopathology 87:1034-1040.

Faino, L., Seidl, M. F., Shi-Kunne, X., Pauper, M., van den Berg, G. C., Wittenberg, A. H., and Thomma, B. P. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome Res 26:1091-1100.

Gunderson, J. H., Elwood, H., Ingold, A., Kindle, K., and Sogin, M. L. 1987. Phylogenetic relationships between chlorophytes, chrysophytes, and oomycetes. Proc Natl Acad Sci USA 84:5823-5827.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., Bollmann, S. R., Bos, J. I., Bulone, V., Cai, G., 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., Jones, J. D., Jones, R. W., Karlsson, E. K., Kunjeti, S. G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M. C., McDonald, H., McWalters, J., Meijer, H. J., Morgan, W., Morris, P. F., Munro, C. A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., 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., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B. C., van West, P., Ristaino, J., Govers, F., Birch, P. R., Whisson, S. C., Judelson, H. S., and Nusbaum, C. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-398.

Hall, A. M. 1959. The culture of *Phytophthora infestans* in artificial media. Trans of the B Mycol Soc 42:15-26.

Hatai, K., and Hoshiai, G. 1992. Mass mortality in cultured coho salmon (*Oncorhynchus kisutch*) due to S*aprolegnia parasitica* coker. J Wildlife Dis 28:532-536.

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Visser, R. G. F., and van der Vossen, E. A. G. 2008. Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. Potato Res 51:47-57.

Hedges, D. J., and Deininger, P. L. 2007. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. Mut Res and Mol Mech of Mutagen 616:46-59.

Hu, C. H., Perez, F. G., Donahoo, R., McLeod, A., Myers, K., Ivors, K., Secor, G., Roberts, P. D., Deahl, K. L., Fry, W. E., and Ristaino, J. B. 2012. Recent genotypes of *Phytophthora infestans* in the eastern united states reveal clonal populations and reappearance of mefenoxam sensitivity. Plant Dis 96:1323-1330.

Judelson, H. S., and Roberts, S. 1999. Multiple loci determining insensitivity to phenylamide fungicides in P*hytophthora infestans*. Phytopathology 89:754-760.

Judelson, H. S., and Senthil, G. 2006. Investigating the role of ABC transporters in multifungicide insensitivity in *Phytophthora infestans*. Mol Plan Path 7:17-29.

Judelson, H. S., Spielman, L. J., and Shattock, R. C. 1995. Genetic mapping and non Mendelian segregation of mating-type loci in the oomycete, *Phytophthora infestans*. Genetics 141:503-512.

Kamoun, S., Furzer, O., Jones, J. D., Judelson, H. S., Ali, G. S., Dalio, R. J., Roy, S. G., Schena, L., Zambounis, A., Panabieres, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X. R., Hulvey, J., Stam, R., Lamour, K., Gijzen, M., Tyler, B. M., Grunwald, N. J., Mukhtar, M. S., Tome, D. F., Tor, M., Van Den Ackerveken, G., McDowell, J., Daayf, F., Fry, W. E., Lindqvist-Kreuze, H., Meijer, H. J., Petre, B., Ristaino, J., Yoshida, K., Birch, P. R., and Govers, F. 2015. The Top 10 oomycete pathogens in molecular plant pathology. Mol Plan Path 16:413-434.

Kaplan, K. B., and Li, R. 2012. A prescription for 'stress' – the role of Hsp90 in genome stability and cellular adaptation. Trends Cell Biology 22:576-583.

Kasuga, T., Kozanitas, M., Bui, M., Hüberli, D., Rizzo, D. M., and Garbelotto, M. 2012. Phenotypic diversification is associated with host-induced transposon derepression in the Sudden Oak Death pathogen *Phytophthora ramorum*. PLoS One 7:e34728.

Kasuga, T., Bui, M., Bernhardt, E., Swiecki, T., Aram, K., Cano, L. M., Webber, J., Brasier, C., Press, C., Grünwald, N. J., Rizzo, D. M., and Garbelotto, M. 2016. Host-induced aneuploidy and phenotypic diversification in the Sudden Oak Death pathogen *Phytophthora ramorum*. BMC Geno 17:385.

Lamour, K. H., and Hausbeck, M. K. 2000. Mefenoxam insensitivity and the sexual stage of *Phytophthora capsici* in Michigan cucurbit fields. Phytopathology 90:396-400.

Lamour, K. H., Mudge, J., Gobena, D., Hurtado-Gonzales, O. P., Schmutz, J., Kuo, A., Miller, N. A., Rice, B. J., Raffaele, S., Cano, L. M., Bharti, A. K., Donahoo, R. S., Finley, S., Huitema, E., Hulvey, J., Platt, D., Salamov, A., Savidor, A., Sharma, R., Stam, R., Storey, D., Thines, M., Win, J., Haas, B. J., Dinwiddie, D. L., Jenkins, J., Knight, J. R., Affourtit, J. P., Han, C. S., Chertkov, O., Lindquist, E. A., Detter, C., Grigoriev, I. V., Kamoun, S., and Kingsmore, S. F. 2012. Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. Mol Plan Micro Interac 25:1350-1360.

Li, Y., Shen, H., Zhou, Q., Qian, K., van der Lee, T., and Huang, S. W. 2017. Changing ploidy as a strategy: the Irish potato famine pathogen shifts ploidy in relation to its sexuality. Mol Plan Micro Interac 30:45-52.

Matson, M. E. H., Small, I. M., Fry, W. E., and Judelson, H. S. 2015. Metalaxyl resistance in *Phytophthora infestans*: assessing role of RPA190 gene and diversity within clonal lineages. Phytopathology 105:1594-1600.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. 2001. Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature 411:212-214.

Patel, S. 2016. Drivers of bacterial genomes plasticity and roles they play in pathogen virulence, persistence and drug resistance. Infect Genet Evol 45:151-164.

Raffaele, S., and Kamoun, S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. Nat Rev Micro 10:417.

Randall, E., Young, V., Sierotzki, H., Scalliet, G., Birch, P. R., Cooke, D. E., Csukai, M., and Whisson, S. C. 2014. Sequence diversity in the large subunit of RNA polymerase I contributes to mefenoxam insensitivity in *Phytophthora infestans*. Mol Plan Path 15:664-676.

Rizzo, D. M., Garbelotto, M., and Hansen, E. M. 2005. *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. Ann Rev Phytopath 43:309-335.

Saville, A. C., Martin, M. D., and Ristaino, J. B. 2016. Historic late blight outbreaks caused by a widespread dominant lineage of *Phytophthora infestans* (mont.) de bary. PLoS One 11:e0168381.

Saze, H., Tsugane, K., Kanno, T., and Nishimura, T. 2012. DNA methylation in plants: relationship to small rnas and histone modifications, and functions in transposon inactivation. Plant Cell Physi 53:766-784.

Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., and Berman, J. 2008. An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. Mol Micro 68:624-641.

Selmecki, A. M., Dulmage, K., Cowen, L. E., Anderson, J. B., and Berman, J. 2009. Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet 5:e1000705.

Shrivistava, J. 2017. Gene structure and variation in *Phytopthora infestans*: Resources for understanding and managing a global food threat to food security. PhD Thesis, University of California, Riverside.

Tan, E. H., Henry, I. M., Ravi, M., Bradnam, K. R., Mandakova, T., Marimuthu, M. P., Korf, I., Lysak, M. A., Comai, L., and Chan, S. W. 2015. Catastrophic chromosomal restructuring during genome elimination in plants. Elife 4:e06516.

Wang, J., Fernandez-Pavia, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodriguez-Alvarado, G., Grunwald, N. J., and Goss, E. M. 2017. High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. Mol Ecol 26:1091-1107.

Yoshida, K., Schuenemann, V. J., Cano, L. M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F. N., Kamoun, S., Krause, J., Thines, M., Weigel, D., and Burbano, H. A. 2013. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. Elife 2:e00731.

Yuen, J. E., and Andersson, B. 2013. What is the evidence for sexual reproduction of *Phytophthora infestans* in Europe? Plant Pathol 62:485-491.

Zhu, W., Yang, L. N., Wu, E. J., Qin, C. F., Shang, L. P., Wang, Z. H., and Zhan, J. 2015. Limited sexual reproduction and quick turnover in the population genetic structure of *Phytophthora infestans* in Fujian, China. Sci Rep 5:10094.