UNIVERSITÉ DE NEUCHÂTEL

FACULTÈ DES SCIENCES
INSTITUTE DE BIOLOGIE

Population genomics of lifestyle and virulence evolution in the fungal wheat pathogen *Zymoseptoria tritici*

A thesis submitted to the
UNIVERSITY OF NEUCHÂTEL

For the degree of
DOCTOR OF SCIENCE

Presented by

Nikhil Kumar Singh

Master of Technology in Biotechnology,
Kalinga Institute of Industrial technology
Bhubaneswar, India

Born on 31.10.1993
Citizen of India

Accepted on the recommendation of

Prof. Dr. Daniel Croll
Prof. Dr. Pilar Junier
Dr. Remco Stam

Defended on 15.10.2021

UNIVERSITÉ DE NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Nikhil Kumar SINGH

Titre:

"Population genomics of lifestyle and virulence evolution in the fungal wheat pathogen *Zymoseptoria tritici*"

sur le rapport des membres du jury composé comme suit:

- Prof. ass. Daniel Croll, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Pilar Junier, Université de Neuchâtel, Suisse
- Dr Remco Stam, Technische Universität München, Allemagne

Neuchâtel, le 7 octobre 2021

Le Doyen, Prof. A. Bangerter

Table of Content

Résumé

Les agents pathogènes des plantes sont à l'origine d'importantes réductions de la productivité agricole dans le monde entier, et les stratégies de lutte restent inefficaces pour de nombreux agents pathogènes. Les agents pathogènes fongiques des plantes surmontent souvent rapidement l'immunité médiée par les gènes de résistance, mettant ainsi en échec les efforts de sélection coûteux. De même, le déploiement de produits chimiques de synthèse est souvent suivi de l'apparition de populations de pathogènes résistants. Cependant, nous ne comprenons toujours pas comment les différentes forces évolutives façonnent la composition génétique des populations d'agents pathogènes. De plus, la manière dont la virulence évolue au fil du temps et sa base génétique restent largement inconnues. Cette thèse de doctorat visait à étudier le mode de vie et l'évolution de la virulence chez le champignon pathogène du blé *Zymoseptoria tritici* en utilisant la génomique des populations, la cartographie d'association et l'analyse des métabolites. Nous avons séquencé en profondeur un large ensemble d'isolats de *Zymoseptoria tritici* prélevés sur de multiples cultivars de blé génétiquement distincts et répliqués dans l'espace à trois moments de la saison de croissance. L'analyse génomique des populations a révélé un niveau étonnant de diversité génétique dans un seul champ, dépassant les niveaux de diversité globale pour de nombreux autres agents pathogènes. L'étude établit le rôle de la reproduction sexuée et asexuée dans l'adaptation à court terme, car la recombinaison permet le remaniement de la variation allélique. Dans le deuxième chapitre, j'ai utilisé la cartographie d'association à l'échelle du génome pour étudier la base génétique de la virulence fongique de *Z. tritici* sur un cultivar de blé couramment utilisé. L'association la plus forte était intergénique et flanquée de gènes codant pour un effecteur prédit et une protéase de type sérine. Le locus de virulence était très dynamique, composé de plusieurs familles d'éléments transposables et de fortes signatures de mutation ponctuelle induite par répétition (RIP), contribuant à la diversification rapide du locus. Dans le troisième chapitre, nous avons étudié la base génétique

de la variation de la production de métabolites dans une seule population de *Z. tritici* sur le terrain. En utilisant un profil métabolique non ciblé et une GWAS, nous avons trouvé de fortes associations dans les gènes de proximité codant pour des fonctions liées aux fonctions endosomales et aux transporteurs. Dans l'ensemble, cette thèse de doctorat souligne le potentiel du séquençage du génome entier d'une seule population de terrain pour comprendre le mode de vie et l'évolution de la virulence d'un important phytopathogène. L'établissement d'une image complète de la base génomique de l'adaptation dans une population de pathogènes aidera à concevoir efficacement des stratégies de gestion des maladies des plantes.

Mots-clés : Génomique des populations, études d'association à l'échelle du génome (GWAS) , *Zymoseptoria tritici*, Métabolite secondaire

Summary

Plant pathogens cause significant reductions in agricultural productivity worldwide, and control strategies remain ineffective for many pathogens. Fungal plant pathogens often rapidly overcome resistance gene mediated immunity defeating expensive breeding efforts. Similarly, the deployment of synthetic chemicals is often followed by the rise of resistant pathogen populations. However, we still lack basic understanding how different evolutionary forces shape the genetic composition of pathogen populations. Additionally, how virulence evolves over time and its genetic basis remains largely unknown. In this PhD thesis was to study the lifestyle and virulence evolution in the fungal wheat pathogen *Zymoseptoria tritici* using population genomics, association mapping and metabolite analysis. We deep sequenced a large set of isolates of *Zymoseptoria tritici* sampled from multiple genetically distinct wheat cultivars replicated in space at three time points of the growing season. Population genomic analysis revealed an astonishing level of genetic diversity in a single field exceeding global diversity levels for many other pathogens. The study establishes the role of sexual and asexual reproduction in short-term adaptation, as recombination enables the reshuffling of allelic variation. In the second chapter, I used genome-wide association mapping to investigate genetic basis of fungal virulence of *Z. tritici* on a commonly used wheat cultivar. The strongest association was intergenic and flanked by genes encoding a predicted effector and a serine-type protease. The virulence loci was highly dynamic, consisting of multiple families of transposable elements and strong signatures of repeat-induced point mutation (RIP), contributing to the locus's rapid diversification. In the third chapter investigated the genetic basis of metabolites production variation in a single field *Z. tritici* population. Using untargeted metabolic profile and GWAS we found strong associations in proximity genes encoding functions related to endosomal functions and transporters. Overall this PhD thesis highlights the potential of whole genome sequencing of a single field population in understanding the

lifestyle and evolution of virulence for an important plant pathogen. Establishing a complete picture of the genomic basis of adaptation in a pathogen population will help in efficient designing of plant disease management strategies

Keywords: Fungal pathogen, Population Genomics, Genome Wide Association studies (GWAS), *Zymoseptoria tritici,* Secondary Metabolites

General Introduction

Plant diseases cause significant reductions in agricultural productivity worldwide. Among the most devastating plant pathogens are fungi that use their respective hosts for reproduction and dispersal (Fisher et al. 2012). Despite the use of control methods such as the deployment of resistant cultivars and the application of chemical compounds (e.g. fungicides, insecticides) fungal pathogens can account for at least a 10% reduction in global yields (Strange and Scott 2005; Petit-Houdenot and Fudal 2017; Rohr et al. 2017; He, Zhan, and Xie 2016). The present practice agricultural provides homogeneous genetic host environments at large regional, making it conducive for the rapid evolution and dispersal of pathogens (Möller and Stukenbrock 2017). The disastrous consequence of the rapid breakdown of resistance was evident in recent outbreaks of *Magnaporthe oryzae* and *Puccinia graminis f. sp. tritici* in Bangladesh and Italy, respectively (Bhattacharya 2017; Islam et al. 2016). The speed with which a fungal pathogen population responds to an altered host population or fungicide is determined by evolutionary forces such as mutation, gene flow, genetic drift and selection (B. A. McDonald and Linde 2002). Hence, to sustainably control fungal diseases, it becomes important to understand the evolutionary process of the pathogens.

Plant-pathogen interactions and its genetic basis

Plant pathogens have emerged as a major threat to the agricultural system causing enormous economic losses annually (Fisher et al. 2012). Fungal pathogens are amongst the most devastating plant pathogens that use plant tissues for their reproduction and dispersal. The fungal pathogens have evolved mechanisms to evade the host defence system, including physical and chemical barriers (Cook, Mesarich, and Thomma 2015). Plants can also produce specific resistance proteins targeted towards pathogen produced molecules. To avoid

recognition or suppress the plant immune system, fungi secrete small proteins called 'effectors' that act in apoplastic or symplastic spaces (Jones and Dangl 2006; Martin and Kamoun 2012). Some effectors are translocated to cellular compartments (for example, chloroplasts or nuclei), manipulating host transcription to favour pathogen invasion (Weiberg et al. 2013). Other effectors manipulate host defences by interfering with plant metabolism or hormone homeostasis (Cook, Mesarich, and Thomma 2015). To counteract pathogen invasion, plants have evolved the ability to produce resistance proteins to detect the effectors; these proteins are activated by effector-mediated signals and are responsible for signal-transduction events that activate effector-triggered immunity and block pathogen growth (Martin and Kamoun 2012). Ultimately, the infection success is determined by the pathogen's ability to overcome the immune system of the host and the ability of the plant to block pathogen invasion

Plants and their surrounding pathogens have been in a co-evolutionary arms race for millions of years. The plant immune response is mainly controlled by resistance (R) genes. The gene-for-gene (GFG) hypothesis suggests that for every host resistance gene (R-gene), there would be a corresponding avirulence (Avr) gene in the pathogen. The protein encoded by the Avr gene is recognized directly or indirectly by the product of the host R gene. Directly recognizing the Avr gene in the pathogen by the host's R gene will trigger resistance and an "arms race" between the host and pathogen. The notion that virulence is not fixed but evolves can be traced to Pasteur and Koch in the 19th century (Smith 1904; Alizon et al. 2009). The arms race coevolution model explains the continuous evolution of host-pathogen interactions. According to the model, disease pressure will favour changes in host proteins that are a target for fungal attack. In contrast, the pathogen will be pressured to evolve its virulence factors to continue to be able to attack the modified targets. In this context, virulence (at the population level) can be defined as the average ability of a pathogen population to overcome the diversity of resistance

genes present in the corresponding host population (Thrall and Burdon 2003; Hartmann et al. 2017). The interactions models have been validated in several plant-pathogen system and since then numerous Avr genes have been cloned (Palloix, Ayme, and Moury 2009; Thrall and Burdon 2003). However, a large number of Avr genes remains to be identified.

Different host genotypes vary in their composition of resistance genes, hence the selection pressure on the pathogen would also be highly heterogeneous. Studies have shown that virulent pathogens were found more frequently on highly resistant host populations, whereas avirulent pathogens dominated susceptible populations therefore, it can be inferred that evolution of pathogenicity is directly influenced by host diversity (Thrall and Burdon 2003). Nonetheless, the general influence of host diversity on the direction of selection on pathogen genes and the population structure is largely unknown.

Drivers of population structure of fungal pathogens

All living acquire resources from their surroundings throughout life and allocate them to maintain different life-cycle stages, e.g. development, survival, reproduction, and dispersal (Barrett et al. 2008). In plant pathogens these life cycle-cycle stages determine spatial and temporal dynamics of disease, thus influencing the genetic diversity and population structure (Kingsolver and Pfennig 2007). Hence pathogens would have to allot molecular machinery to express specialized traits at a specific time point of their life cycle to survive. These traits can be divided into the epidemic phase, where the pathogen has to survive within the host, and later on the survival phase, mostly outside the host. Pathogens having a narrow range of hosts are likely to be more locally adapted than their generalist counterparts (Lajeunesse and Forbes 2002). Moreover, an obligate pathogen specialised on single host is more susceptible to experience local bottleneck and recolonisation event (Lajeunesse and Forbes 2002). For

example, the obligate rust pathogen *Melampsora lini* infecting wild Australian flax *Linum marginale* follows a typical "boom-and-bust" epidemics. In a "boom-and-bust cycle" the widespread use of a single resistance gene protects multiple varieties of a grain from a disease (boom). When the pathogen overcomes this resistance gene many varieties simultaneously become susceptible (bust). Meanwhile, *Colletotrichum gloeosporioides,* a wide host range pathogen, show a clonal population structure with a slow migration rate but has evolved different genetic background for each of its host cultivars (Frézal, Jacqua, and Neema 2018; Cruz-Lagunas et al. 2020; Zhang et al. 2016).

The mode of reproduction also plays an important part in crucial the genetic structure of the pathogen population. Reproduction can be sexual, asexual, or mixed. Recent findings indicate that many fungal species known to be asexual have retained the ability to reproduce sexually and possess mixed reproduction (Nieuwenhuis and James 2016). In a mixed reproduction model, many new combinations of alleles are created through recombination during the sexual cycle, thus producing many genotypes. The most "fit" genotypes can then propagate by asexual reproduction *Puccinia graminis f.sp. tritici*, which causes wheat stripe rust, can perform sexual reproduction, thus generating a higher genetic and phenotypic diversity than a more common asexual reproduction in its original host (Leonard and Szabo 2005). The difference in the frequency of sexual and asexual reproduction may impact the genetic structure and adaptation of a pathogen population is still unclear.

It is also essential to understand the mode of transmission for correctly interpreting the genetic diversity of the pathogen population. For example, species propagating through soil or rain-splash faces a decline in the dispersal distance while wind dispersing species can propagate further (Karisto, Suffert, and Mikaberidze 2019; Barrett et al. 2008). However, in extreme

cases, it has been seen that very large distance dispersal can only occur if there is a susceptible host in the new location. In case *Puccinia graminis f.sp. tritici* spores dispersed from South Africa to Australia in a jet stream (Leonard and Szabo 2005). In such cases, the new population would suffer a loss in genetic diversity due to population bottleneck, drastically reducing the size of a population and reducing overall genetic diversity.

Role of secondary metabolites fungal plant pathogens

In their natural environment, fungal pathogens are challenged by multiple biotic and abiotic stress conditions. These range from host-pathogen interactions to nutrient deprivation to pH and temperature. As one of the responses, pathogens produce a vast number of low molecular weight compounds called secondary metabolites (SMs). SMs (except siderophores) are non-essential for growth of fungi but are involved in biological signalling and interactions (Calvo et al. 2002). SMs play vital role in the highly complex interactions of fungal plant pathogens with their respective hosts. SMs of plant pathogens often alter host cell structures and functions through different modes of action. For instance, the HC-toxin produced by the fungal pathogen of maize *Cochliobolus carbonum* inhibits histone deacetylases of both plants and mammals, disrupting gene expression (Ahn, Cheng, and Walton 2002). Genes required produce SMs are arranged in a physically linked contiguous fashion and are known as biosynthetic gene clusters (BGCs)(Keller, Turner, and Bennett 2005). According to their ecological functions, these gene clusters are co-regulated to encode the secondary metabolite (Keller, Turner, and Bennett 2005). Each BGC contains a core/backbone gene that codes for an enzyme which polymerises primary metabolites. The resulting metabolite is then further modified by the enzyme products of additional tailoring genes of the BGC to produce a final bioactive secondary metabolite. For example, polyketide synthase (PKS), a core gene, produces polyketides from acyl-CoAs (Keller, Turner, and Bennett 2005).

SMs are sub-grouped into four groups according to the precursors used by the core biosynthetic gene for their synthesis (Keller, Turner, and Bennett 2005). The two major groups are polyketides and non-ribosomal peptides (NRP). Acyl-CoAs units are converted to polyketides by polyketide synthetase enzyme (PKS). The other two groups are terpenes and alkaloids are more abundant in plants rather than fungi. Both these SMs are produced from isoprene units by terpene cyclase and dimethylallyl tryptophan synthases respectively (Pusztahelyi, Holb, and PÃ³csi 2015; Collemare and Lebrun 2011). Remarkably, each key core biosynthetic gene produces a different metabolite from similar precursors, leading to diversity in SMs (Collemare and Lebrun 2011). For example, *Gibberella fujikuroi* fumonisin biosynthetic pathway can produce four different fumonisins (FB1–FB4) (Bojja et al. 2004).

Fungal SMs can be broadly classified as host-specific or non-host-specific (Pusztahelyi, Holb, and PÃ³csi 2015; Collemare and Lebrun 2011). Host specific fungal SMs usually behave as effectors during plant-pathogen interactions. Resistance of plants to host specific SMs mainly relies on the absence of the susceptibility target protein or its modification in the host. However, the non-host specific SMs contribute to the aggressiveness of the pathogen (Collemare and Lebrun 2011). Many non-host specific SMs can induce necrosis or chlorosis on an array of plant species and provide substantial benefits to the disease progression of necrotrophic fungi. For example, the mycotoxin deoxynivalenol (DON) produced by a plant pathogen *Gibberella zeae*. Mutants defective in DON biosynthesis retained the ability to infect wheat but exhibited decreased pathogenicity (Mitchell 1984).

Even though it is understood that SMs play an important role in host pathogen interactions, very few of these SMs are fully characterized in terms of their function and mode of action.

Recent advances in chromatography techniques and genomics data availability have accelerated the prediction and validation of previously uncharacterised secondary metabolites BGCs (Keller 2019). It has been reported in many comparative genomics studies that individual BGCs are segregating distinctively in specific species or dispersed within a limited number of species (Keller 2019, 2015). Hence, the total SM gene cluster distribution and the coded metabolite can vary significantly in closely related species and even within the same species (Tralamazza et al. 2019; Lind et al. 2015). The genetic basis of the evolution of SM gene clusters are still not fully explored. In plant models genetic variations responsible for controlling natural variation in accumulation of metabolites have been studied (Jacobowitz and Weng 2020; Peng et al. 2017; Strauch et al. 2015). Exploring whole genome data of pathogen populations would help in identification of the genetic architecture related to metabolite gene clusters and resultant variation in metabolite accumulation (Pais, Li, and Xiang 2018; Firn and Jones 2000).

Genome wide association study

Association mapping aims at identifying genomic loci that mediate heritable phenotypic trait variation in natural populations. The statistical significance of the associations between phenotype and genotype variation is calculated in a group of individuals, also called a mapping population (Bergelson and Roux 2010). A genome wide association study (GWAS) is an approach that involves scanning genetic markers, specifically single nucleotide polymorphisms (SNPs), across the genome of many individuals of a species to find loci associated with a particular phenotype (Bush et al. 2012). Among various types of sequence variants, SNPs are by far the most abundant in the genome of many of the organisms (Huang and Han 2014). Moreover, they can have functional consequences, like causing amino acid changes, changes

to mRNA transcript stability, and changes to transcription factor binding affinity(Randles et al. 2013).

In GWAS, a mapping population can include individuals from one or multiple natural populations(s) and, therefore, screen for a broader range of phenotypes and genotypes (Bergelson and Roux 2010). To correct for population structure, mixed linear model are commonly used which corrects for the relatedness of the individuals (Korte and Farlow 2013). Although GWAS performs poorly in detecting genetic variants with low effect size, this issue is compensated by the advantage of fine mapping down to gene level. Additionally, the power of finding robust genotype-phenotype association depends on the level of recombination in the mapping population (Sánchez-Vallet et al. 2018; Korte and Farlow 2013). Decay of linkage disequilibrium (LD), which is the non-random association of allele, gives a strong indication of the mapping power of GWAS (Korte and Farlow 2013). GWAS has been extensively used to map the genetic basis of human diseases. Researchers have successfully identified genetic variations that contribute to risk of type 2 diabetes, Parkinson's disease, heart disorders, obesity, Crohn's disease and prostate cancer, as well as genetic variations that influence response to anti-depressant medications (Bush et al. 2012; Norrgard 2008). The recent availability of high-resolution sequence information has also made GWAS possible for many other species, including maize, rice, *Drosophila melanogaster, Plasmodium falciparum, Arabidopsis thaliana, Rhynchosporium commune* etc. (Magwire et al. 2012; Mohd-Assaad, McDonald, and Croll 2016).

The GWAS approach is still poorly used in fungal pathogens but was a powerful method to identify genetic variation underlying phenotypic variation in other systems (Bartoli and Roux 2017). Genomic region linked to virulence was identified in hundreds of isolates of a wheat

pathogen *F. graminarium* (Talas et al. 2016). GWAS on mapping population consisting of individuals from multiple population allowed to identify genetic variation associated with fungicide resistance in *Rhynchosporium commune* and *Parastagonospora nodorum* (Mohd-Assaad, McDonald, and Croll 2016; Pereira, McDonald, and Croll 2020). High throughput phenotyping platform and availability of whole genome sequence application of GWAS on fungal pathogen remains scarce (Singh, Dutta, et al. 2021; Bartoli and Roux 2017). The reason could be population structure and lack of recombination due to clonal reproduction, common in many fungal pathogens. Since the application of GWAS depends on factors like heritable phenotypic variation, weak population structure and rapid decay in linkage disequilibrium. *Z.tritici* satisfies all these conditions, and a recent study has shown the successful application of GWAS to identify a candidate effector gene related to the pathogen's virulence (Plissonneau et al. 2017; Hartmann et al. 2017; Singh, Badet, et al. 2021; Zhong et al. 2017).

The fungal wheat pathogen *Zymoseptoria tritici*

Zymoseptora tritici, also known as *Mycosphaerella graminicola* or *Septoria tritici* (anamorph), is a haploid, heterothallic ascomycete and one of the most destructive pathogens of wheat (M. C. McDonald, McDonald, and Solomon 2015). The pathogen emerged from the fertile crescent during the domestication of wheat about ~9,000 BC and is specific to wheat as a host (Stukenbrock et al. 2007). The pathogen has a biphasic life cycle. At the onset of the biotrophic phase (also denoted as non-symptomatic or latent phase) ascospores or pycnidiospores land on the wheat leaf epidermis, where they germinate, and the hyphae invades the sub-stomatal cavity via the stomatal opening. This non-symptomatic phase lasts for around 7-14 days. About 5-9 days post infection (dpi) pre-pycnidia are formed and a switch to the necrotic phase is observed. During the necrotic phase, also known as the symptomatic phase of the infection, plant cells undergo programmed cell death and the dying cells provide the fungus with the necessary

nutrients for rapid proliferation and pycnidia maturation. As a result, mature pycnidia become visible as black dots on the leaf surface (M. C. McDonald, McDonald, and Solomon 2015; Kettles and Kanyuka 2016; Steinberg 2015; Sánchez-Vallet et al. 2015). An essential aspect of Z. tritici is the sequential sexual and asexual reproduction cycle regime over the epidemic. During the sexual stage, the pathogen produces genetically distinct ascospores, which are thought to be the source of primary infection in the field and can disperse over long distances (R.-S. Chen and Mcdonald 1996; Kema et al. 1996). Once infections are established in a field, *Z. tritici* produces asexual fruiting body called pycnidiospores, which can disperse with rain splash and initiate secondary infections. The spread within fields during the epidemic is thought to be predominantly clonal (R.-S. Chen and Mcdonald 1996).

Studies have shown that the rapid switch from the biotrophic phase to the necrotrophic phase is accompanied by increased expression of genes coding for small secreted proteins (SSP). Genes like *MAPK-ZtHog1, Mg3LysM and Zt80707, AvrStb6* have been shown to play an important role during the infection (Kettles and Kanyuka 2016; Testa, Oliver, and Hane 2015; Hartmann et al. 2017; Zhong et al. 2017). There are tools available to determine the protein products of these genes and predict their possible role during the infection, for e.g. *Mg3LysM* encodes a protein that prevents chitin binding and protects against defence responses triggered by the host. (Hartmann et al. 2017) found genes coding for different enzymes like a beta-galactosidase, a major facilitator superfamily transporter and a pectinase. Most importantly identified a previously uncharacterized gene that encodes an SSP likely acting as an effector during wheat infection.

In agricultural systems, a large number of wheat cultivars are used and thus present a diverse complement in R-genes to the pathogen. For *Z. tritici* to remain successful in infecting the host,

the pathogen needs to evolve virulence towards novel R gene products deployed with novel cultivars (Hartmann et al. 2017). There is strong evidence that host resistance can be broken down rapidly, which is demonstrated in a case study of the wheat cultivar Gene. Gene is a F_4-derived selection from a cross involving the resistant cultivar Cleo, which is believed to carry the *Septoria tritici* blotch resistance gene *Stb4*. This cultivar was resistant to *Z. tritici* when it was introduced in 1992 but in just five years the pathogen adapted to become virulent towards Gene (Cowger, Hoffer, and Mundt 2000). Research from (Zhong et al. 2017) demonstrates the precise gene in the pathogen that evolved virulence to a different wheat resistance gene product (*Stb6*) thus pointing towards the rapid evolution of pathogenicity of *Z. tritici.* Taken together, this demonstrates that there are multiple pathogen loci, each of which having evolved to cause disease on a specific subset of hosts. Hence, it becomes important to understand the molecular mechanism of infections and the factors favouring rapid evolution in the pathogen.

The availability of high quality and fully assembled genomes has enabled a thorough investigation of adaptive process of the devastating pathogen (Goodwin et al. 2011; Badet et al. 2020). *Z. tritici* populations have been characterized by large effective population size, significant gene flow and high recombination rates (R. S. Chen and McDonald 1996; Zhan, Mundt, and Mcdonald 1998; Croll et al. 2015; Singh, Chanclud, and Croll 2020). Combined with mixed reproduction cycle these characteristics confer a highly adaptive potential to the fungus. Several traits like virulence, melanisation, temperature sensitivity and fungicide resistance were shown to be under selection due to local adaptation (Zhan et al. 2005). More specifically, virulence phenotype has been observed to quantitative in field population indicating that genetic basis is complex (Brown et al. 2015; E. L. Stewart and McDonald 2014). Study of secretome and differential expression of genes in-planta identified multiple effector gene which play an important role in virulence (Rudd et al. 2015; Palma-Guerrero et al. 2017).

Quantitative trait loci (QTL) mapping in four independent *Z. tritici* crosses identified genes related to virulence (E. L. Stewart et al. 2016; Gohari et al. 2015). Genome-wide association study (GWAS) and analyses of progeny populations revealed multiple effectors of *Z.tritici* (Hartmann et al. 2017; Zhong et al. 2017; Gohari et al. 2015; Meile et al. 2018; E. l. Stewart et al. 2018). Multiple studies have also explored the genes related to morphological and lifestyle switch which also regulates virulence (Francisco et al. 2019; Habig et al. 2020). However, studies describing the genetic basis of virulence evolution in a population is still lacking.

Evolutionary processes in fungal pathogens are also driven by features of the chromosomal sequence surrounding genes associated to adaptation to stress (Coghlan et al. 2005). Recent analysis of high quality global pangenomes found highly diverse chromosomal structure which was strongly associated with repeat rich region (Badet et al. 2020; Plissonneau, Hartmann, and Croll 2018). The vicinity of genes to repetitive regions, particularly transposable elements (TEs), increases the likelihood of sequence rearrangements (Plissonneau, Hartmann, and Croll 2018; Plissonneau et al. 2016; Badet et al. 2020; Hartmann et al. 2017). On chromosomal level, *Z. tritici* has eight non-essential chromosomes that is present in only in few isolates (Habig, Quade, and Stukenbrock 2017; Badet et al. 2020). Genome-wide expansion of TE families was also found in the pathogen's genome where the burst of TEs was observed in more recent population (Oggenfuss et al. 2020; Lorrain et al. 2020). This suggests that population bottlenecks played a major role in shaping TE content of the genome. The distance of genes from TEs was also differentially expressed under varying stress conditions (Fouché et al. 2020). The organisation of *Z. tritici* genome is consistent with the "two-speed genome" model with compartments of transposable element rich region and accessory chromosomes evolving faster and is suggested to be the drivers of adaptive processes (Croll and McDonald 2012; Dong, Raffaele, and Kamoun 2015; Stukenbrock et al. 2010). Life history traits in *Z. tritici* were

recently demonstrated to be governed by pleiotropic effects. (Dutta et al. 2021) found that traits related to host colonization were negatively corelated to pathogen's ability to survive in different abiotic stress. We still lack knowledge on how genetic variation in populations impacts the evolutionary trajectory and adaptation of this major wheat pathogen.

Main objective of this thesis

Despite a substantial body of research on *Z. tritici* we still lack basic understanding how different evolutionary forces shape the genetic composition of pathogen populations. Additionally, how virulence evolves over time and its genetic basis remains largely unknown. The main aim of this thesis was to improve our understanding of lifestyle and virulence evolution in the fungal wheat pathogen *Zymoseptoria tritici* in an annual epidemic season. For this I will be using population genomics, association mapping and metabolite analysis.

Chapter 1: Population-level deep sequencing reveals the interplay of clonal and sexual reproduction in the fungal wheat pathogen Zymoseptoria tritici.

The role of spatial and temporal variation in pathogen genetic diversity remains poorly understood (Barrett et al. 2008) and we still have few empirical data concerning the ways in which interactions between epidemiological and evolutionary processes influence the generation and maintenance of such variation in host– pathogen interactions (Tack et al. 2012). In this chapter I investigated the diversity in a single field population and how mixed reproduction strategies shape the population structure.

Chapter 2: Rapid sequence evolution driven by transposable elements at a virulence locus in a fungal wheat pathogen.

The rise of new virulence mechanism pathogens or the gain in resistance against chemical control is a significant challenge to agriculture (Strange and Scott 2005; McCann 2020; Subbarao, Sundin, and Klosterman 2015). The identification of the genetic basis of virulence in fungal pathogens is challenging due to the large number of many genomic features e.g presence of repeat regions or presence absence polymorphism of virulence related genes. In this chapter I performed Genome wide association study on various virulence phenotype to

map the associated loci in a single field population. I further investigated the major effect loci in global population

Chapter 3: Genome-wide association mapping to identify genes underlying population-level metabolome diversity in a fungal crop pathogen.

Fungal specialized metabolites (SMs) are involved in virulence and other biotic interactions. Structural variation related SM gene clusters can result in significant variation in metabolome production. In this chapter, I performed untargeted metabolite profile of a single field population. Using GWAS on the untargeted metabolite profile the study explored associated loci and gene functions.

Chapter 4:

Published review: Tackling microbial threats in agriculture with integrative imaging and computational approaches

Given the focus on plant pathogens in this thesis, I wanted to provide a broader context on the importance and recent developments in the high throughput phenotyping (HTP) platform. I in this short review I discussed the HTP for agricultural fields as well laboratory-based monitoring of pathogens. Later we argue how screening phenotype variation of large scale can yield fundamental insights into how pathogens face trade-offs in the adaptation to resistant crop varieties**.**

General discussion and perspectives*:* Here, I am summarizing the key insights of my thesis and highlight future directions.

References

Ahn, Joong Hoon, Yi Qiang Cheng, and Jonathan D. Walton. 2002. “An Extended Physical Map of the TOX2 Locus of Cochliobolus Carbonum Required for Biosynthesis of HC-Toxin.” *Fungal Genetics and Biology* 35 (1): 31–38. https://doi.org/10.1006/fgbi.2001.1305.

Alizon, S, A Hurford, N Mideo, and M Van Baalen. 2009. “Virulence Evolution and the Trade-off Hypothesis: History, Current State of Affairs and the Future.” *Journal of Evolutionary Biology*. https://doi.org/10.1111/j.1420-9101.2008.01658.x.

Badet, Thomas, Ursula Oggenfuss, Leen Abraham, Bruce A. McDonald, and Daniel Croll. 2020. “A 19-Isolate Reference-Quality Global Pangenome for the Fungal Wheat Pathogen Zymoseptoria Tritici.” *BMC Biology* 18 (1): 12. https://doi.org/10.1186/s12915-020-0744-3.

Barrett, Luke G., Peter H. Thrall, Jeremy J. Burdon, and Celeste C. Linde. 2008. “Life History Determines Genetic Structure and Evolutionary Potential of Host-Parasite Interactions.” *Trends in Ecology and Evolution*. Elsevier Current Trends. https://doi.org/10.1016/j.tree.2008.06.017.

Bartoli, Claudia, and Fabrice Roux. 2017. “Genome-Wide Association Studies In Plant Pathosystems: Toward an Ecological Genomics Approach.” *Frontiers in Plant Science* 8: 763. https://doi.org/10.3389/fpls.2017.00763.

Bergelson, Joy, and Fabrice Roux. 2010. “Towards Identifying Genes Underlying Ecologically Relevant Traits in Arabidopsis Thaliana.” *Nature Reviews Genetics* 11 (12): 867–79. https://doi.org/10.1038/nrg2896.

Bhattacharya, Shaoni. 2017. “Deadly New Wheat Disease Threatens Europe’s Crops.” *Nature* 542 (7640): 145–46. https://doi.org/10.1038/nature.2017.21424.

Bojja, Ravi S., Ronald L. Cerny, Robert H. Proctor, and Liangcheng Du. 2004. “Determining the Biosynthetic Sequence in the Early Steps of the Fumonisin Pathway by Use of Tree Gene-Disruption Mutants of Fusarium Verticillioides.” *Journal of Agricultural and Food Chemistry* 52 (10): 2855–60. https://doi.org/10.1021/jf035429z.

Brown, James K.M., Laëtitia Chartrain, Pauline Lasserre-Zuber, and Cyrille Saintenac. 2015. “Genetics of Resistance to Zymoseptoria Tritici and Applications to Wheat Breeding.” *Fungal Genetics and Biology* 79 (June): 33–41. https://doi.org/10.1016/j.fgb.2015.04.017.

Bush, William S., Jason H. Moore, J Li, SK McDonnell, and KG Rabe. 2012. “Chapter 11: Genome-Wide Association Studies.” Edited by Fran Lewitter and Maricel Kann. *PLoS Computational Biology* 8 (12): e1002822. https://doi.org/10.1371/journal.pcbi.1002822.

Calvo, Ana M, Richard A Wilson, Jin Woo Bok, and Nancy P Keller. 2002. “Relationship between Secondary Metabolism and Fungal Development.” *Microbiology and Molecular Biology Reviews* 66 (3): 447–59. https://doi.org/10.1128/mmbr.66.3.447-459.2002.

Chen, R S, and B A McDonald. 1996. “Sexual Reproduction Plays a Major Role in the Genetic Structure of Populations of the Fungus Mycosphaerella Graminicola.” *Genetics* 142 (4): 1119–27. http://www.ncbi.nlm.nih.gov/pubmed/8846892.

Chen, Ruey-Shyang, and Bruce A Mcdonald. 1996. “Sexual Reproduction Plays a Major Role in the Genetic Structure of Populations of the Fungus Mycosphaerella Graminicola.” https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1207112/pdf/ge14241119.pdf.

Coghlan, Avril, Evan E Eichler, Stephen G Oliver, Andrew H Paterson, and Lincoln Stein. 2005. “Chromosome Evolution in Eukaryotes: A Multi-Kingdom Perspective.” *Trends in Genetics*. Elsevier. https://doi.org/10.1016/j.tig.2005.09.009.

Collemare, Jérôme, and Marc-Henri Lebrun. 2011. “Fungal Secondary Metabolites: Ancient Toxins and Novel Effectors in Plant-Microbe Interactions.” In *Effectors in Plant-Microbe Interactions*, 377–400. Oxford, UK: Wiley-Blackwell. https://doi.org/10.1002/9781119949138.ch15.

Cook, David E., Carl H. Mesarich, and Bart P.H.J. Thomma. 2015. “Understanding Plant Immunity as a Surveillance System to Detect Invasion.” *Annual Review of Phytopathology* 53 (1): 541–63. https://doi.org/10.1146/annurev-phyto-080614-120114.

Cowger, C., M. E. Hoffer, and C. C. Mundt. 2000. “Specific Adaptation by Mycosphaerella Graminicola to a Resistant Wheat Cultivar.” *Plant Pathology* 49 (4): 445–51. https://doi.org/10.1046/j.1365-3059.2000.00472.x.

Croll, Daniel, Mark H Lendenmann, Ethan Stewart, and Bruce A McDonald. 2015. “The Impact of Recombination Hotspots on Genome Evolution of a Fungal Plant Pathogen.” *Genetics* 201 (3): 1213–28. https://doi.org/10.1534/genetics.115.180968.

Croll, Daniel, and Bruce A. McDonald. 2012. “The Accessory Genome as a Cradle for Adaptive Evolution in Pathogens.” *PLoS Pathogens* 8 (4). https://doi.org/10.1371/journal.ppat.1002608.

Cruz-Lagunas, Blas, Santo Ángel Ortega-Acosta, Guadalupe Reyes-García, Jeiry Toribio-Jiménez, Porfirio Juárez-López, Dagoberto Guillén-Sánchez, Agustín Damián-Nava, Yanet Romero-Ramírez, and Francisco Palemón-Alberto. 2020. “Colletotrichum Gloeosporioides Causes Anthracnose on Grapefruit (Citrus Paradisi) in Mexico.” *Australasian Plant Disease Notes* 15 (1): 31. https://doi.org/10.1007/s13314-020-00401-z.

Dong, Suomeng, Sylvain Raffaele, and Sophien Kamoun. 2015. “The Two-Speed Genomes of Filamentous Pathogens: Waltz with Plants.” *Current Opinion in Genetics and Development* 35: 57–65. https://doi.org/10.1016/j.gde.2015.09.001.

Dutta, Anik, Fanny E. Hartmann, Carolina Sardinha Francisco, Bruce A. McDonald, and Daniel Croll. 2021. “Mapping the Adaptive Landscape of a Major Agricultural Pathogen Reveals Evolutionary Constraints across Heterogeneous Environments.” *ISME Journal*, January, 1–18. https://doi.org/10.1038/s41396-020-00859-w.

Firn, Richard D., and Clive G. Jones. 2000. “The Evolution of Secondary Metabolism - A Unifying Model.” *Molecular Microbiology* 37 (5): 989–94. https://doi.org/10.1046/j.1365-2958.2000.02098.x.

Fisher, Matthew C., Daniel. A. Henk, Cheryl J. Briggs, John S. Brownstein, Lawrence C. Madoff, Sarah L. McCraw, and Sarah J. Gurr. 2012. “Emerging Fungal Threats to Animal, Plant and Ecosystem Health.” *Nature* 484 (7393): 186–94. https://doi.org/10.1038/nature10947.

Fouché, Simone, Thomas Badet, Ursula Oggenfuss, Clémence Plissonneau, Carolina Sardinha Francisco, and Daniel Croll. 2020. “Stress-Driven Transposable Element De-Repression Dynamics and Virulence Evolution in a Fungal Pathogen.” Edited by Irina Arkhipova. *Molecular Biology and Evolution* 37 (1): 221–39. https://doi.org/10.1093/molbev/msz216.

Francisco, Carolina Sardinha, Xin Ma, Maria Manuela Zwyssig, Bruce A. McDonald, and Javier Palma-Guerrero. 2019. “Morphological Changes in Response to Environmental Stresses in the Fungal Plant Pathogen Zymoseptoria Tritici.” *Scientific Reports* 9 (1): 9642. https://doi.org/10.1038/s41598-019-45994-3.

Frézal, Lise, Guy Jacqua, and Claire Neema. 2018. “Adaptation of a Fungal Pathogen to Host Quantitative Resistance.” *Frontiers in Plant Science* 871 (October). https://doi.org/10.3389/fpls.2018.01554.

Gohari, Amir Mirzadi, Sarah B. Ware, Alexander H. J. Wittenberg, Rahim Mehrabi, Sarrah Ben M’Barek, Els C. P. Verstappen, Theo A. J. van der Lee, et al. 2015. “Effector

Discovery in the Fungal Wheat Pathogen Zymoseptoria Tritici." *Molecular Plant Pathology* 16 (9): 931. https://doi.org/10.1111/MPP.12251.

Goodwin, Stephen B., Sarrah Ben M'Barek, Braham Dhillon, Alexander H.J. Wittenberg, Charles F. Crane, James K. Hane, Andrew J. Foster, et al. 2011. "Finished Genome of the Fungal Wheat Pathogen Mycosphaerella Graminicola Reveals Dispensome Structure, Chromosome Plasticity, and Stealth Pathogenesis." Edited by Harmit S. Malik. *PLoS Genetics* 7 (6): e1002070. https://doi.org/10.1371/journal.pgen.1002070.

Habig, Michael, Sharon Marie Bahena-Garrido, Friederike Barkmann, Janine Haueisen, and Eva Holtgrewe Stukenbrock. 2020. "The Transcription Factor Zt107320 Affects the Dimorphic Switch, Growth and Virulence of the Fungal Wheat Pathogen Zymoseptoria Tritici." *Molecular Plant Pathology* 21 (1): 124–38. https://doi.org/10.1111/mpp.12886.

Habig, Michael, Jakob Quade, and Eva Holtgrewe Stukenbrock. 2017. "Forward Genetics Approach Reveals Host Genotype-Dependent Importance of Accessory Chromosomes in the Fungal Wheat Pathogen Zymoseptoria Tritici." *MBio* 8 (6): e01919-17. https://doi.org/10.1128/mBio.01919-17.

Hartmann, Fanny E, Andrea Sánchez-Vallet, Bruce A McDonald, and Daniel Croll. 2017. "A Fungal Wheat Pathogen Evolved Host Specialization by Extensive Chromosomal Rearrangements." *The ISME Journal* 11 (5): 1189–1204. https://doi.org/10.1038/ismej.2016.196.

He, Dun Chun, Jia Sui Zhan, and Lian Hui Xie. 2016. "Problems, Challenges and Future of Plant Disease Management: From an Ecological Point of View." *Journal of Integrative Agriculture* 15 (4): 705–15. https://doi.org/10.1016/S2095-3119(15)61300-4.

Huang, Xuehui, and Bin Han. 2014. "Natural Variations and Genome-Wide Association Studies in Crop Plants." *Annual Review of Plant Biology* 65 (November 2013): 531–51. https://doi.org/10.1146/annurev-arplant-050213-035715.

Islam, M. Tofazzal, Daniel Croll, Pierre Gladieux, Darren M. Soanes, Antoine Persoons, Pallab Bhattacharjee, Md. Shaid Hossain, et al. 2016. "Emergence of Wheat Blast in Bangladesh Was Caused by a South American Lineage of Magnaporthe Oryzae." *BMC Biology* 14 (1): 84. https://doi.org/10.1186/s12915-016-0309-7.

Jacobowitz, Joseph R., and Jing Ke Weng. 2020. "Exploring Uncharted Territories of Plant Specialized Metabolism in the Postgenomic Era." *Annual Review of Plant Biology* 71: 631–58. https://doi.org/10.1146/annurev-arplant-081519-035634.

Jones, Jonathan D. G., and Jeffery L. Dangl. 2006. "The Plant Immune System." *Nature* 444 (7117): 323–29. https://doi.org/10.1038/nature05286.

Karisto, Petteri, Frédéric Suffert, and Alexey Mikaberidze. 2019. "Spatially-Explicit Modeling Improves Empirical Characterization of Dispersal: Theory and a Case Study." *BioRxiv*, 1–65. https://doi.org/10.1101/789156.

Keller, Nancy P. 2019. "Fungal Secondary Metabolism: Regulation, Function and Drug Discovery." *Nature Reviews Microbiology* 17 (3): 167–80. https://doi.org/10.1038/s41579-018-0121-1.

Keller, Nancy P., Geoffrey Turner, and Joan W. Bennett. 2005. "Fungal Secondary Metabolism — from Biochemistry to Genomics." *Nature Reviews Microbiology* 3 (12): 937–47. https://doi.org/10.1038/nrmicro1286.

Keller, Nancy P. 2015. "Translating Biosynthetic Gene Clusters into Fungal Armor and Weaponry." *Nature Chemical Biology* 11 (9): 671–77. https://doi.org/10.1038/nchembio.1897.

Kema, G H, E C Verstappen, M Todorova, and C Waalwijk. 1996. "Successful Crosses and Molecular Tetrad and Progeny Analyses Demonstrate Heterothallism in Mycosphaerella Graminicola." *Current Genetics* 30 (3): 251–58. http://www.ncbi.nlm.nih.gov/pubmed/8753655.

Kettles, Graeme J., and Kostya Kanyuka. 2016. “Dissecting the Molecular Interactions between Wheat and the Fungal Pathogen Zymoseptoria Tritici.” *Frontiers in Plant Science* 7 (April): 508. https://doi.org/10.3389/fpls.2016.00508.

Kingsolver, Joel G., and David W. Pfennig. 2007. “Patterns and Power of Phenotypic Selection in Nature.” *BioScience*. Oxford Academic. https://doi.org/10.1641/B570706.

Korte, Arthur, and Ashley Farlow. 2013. “The Advantages and Limitations of Trait Analysis with GWAS: A Review.” *Plant Methods* 9: 29. https://doi.org/10.1186/1746-4811-9-29.

Lajeunesse, Marc J., and Mark R. Forbes. 2002. “Host Range and Local Parasite Adaptation.” *Proceedings of the Royal Society B: Biological Sciences* 269 (1492): 703–10. https://doi.org/10.1098/rspb.2001.1943.

Leonard, Kurt J., and Les J. Szabo. 2005. “Stem Rust of Small Grains and Grasses Caused by Puccinia Graminis.” *Molecular Plant Pathology*. https://doi.org/10.1111/j.1364-3703.2005.00273.x.

Lind, Abigail L., Jennifer H. Wisecaver, Timothy D. Smith, Xuehuan Feng, Ana M. Calvo, and Antonis Rokas. 2015. “Examining the Evolution of the Regulatory Circuit Controlling Secondary Metabolism and Development in the Fungal Genus Aspergillus.” Edited by Geraldine Butler. *PLOS Genetics* 11 (3): e1005096. https://doi.org/10.1371/journal.pgen.1005096.

Lorrain, Cécile, Alice Feurtey, Mareike Möller, Janine Haueisen, and Eva Stukenbrock. 2020. “Dynamics of Transposable Elements in Recently Diverged Fungal Pathogens: Lineage-Specific Transposable Element Content and Efficiency of Genome Defences.” *BioRxiv*. https://doi.org/10.1101/2020.05.13.092635.

Magwire, Michael M., Daniel K. Fabian, Hannah Schweyen, Chuan Cao, Ben Longdon, Florian Bayer, and Francis M. Jiggins. 2012. “Genome-Wide Association Studies Reveal a Simple Genetic Basis of Resistance to Naturally Coevolving Viruses in Drosophila Melanogaster.” Edited by Marta L. Wayne. *PLoS Genetics* 8 (11): e1003057. https://doi.org/10.1371/journal.pgen.1003057.

Martin, Francis, and Sophien Kamoun. 2012. *Effectors in Plant – Microbe Interactions*.

McCann, Honour C. 2020. “Skirmish or War: The Emergence of Agricultural Plant Pathogens.” *Current Opinion in Plant Biology* 56 (August): 147–52. https://doi.org/10.1016/J.PBI.2020.06.003.

McDonald, Bruce A., and Celeste Linde. 2002. “P ATHOGEN P OPULATION G ENETICS , E VOLUTIONARY P OTENTIAL, AND D URABLE R ESISTANCE.” *Annual Review of Phytopathology* 40 (1): 349–79. https://doi.org/10.1146/annurev.phyto.40.120501.101443.

McDonald, Megan C, Bruce A McDonald, and Peter S Solomon. 2015. “Recent Advances in the Zymoseptoria Tritici-Wheat Interaction: Insights from Pathogenomics.” *Frontiers in Plant Science* 6: 102. https://doi.org/10.3389/fpls.2015.00102.

Meile, Lukas, Daniel Croll, Patrick C. Brunner, Clémence Plissonneau, Fanny E. Hartmann, Bruce A. McDonald, and Andrea Sánchez-Vallet. 2018. “A Fungal Avirulence Factor Encoded in a Highly Plastic Genomic Region Triggers Partial Resistance to Septoria Tritici Blotch.” *New Phytologist* 219 (3): 1048–61. https://doi.org/10.1111/nph.15180.

Mitchell, R E. 1984. “The Relevance of Non-Host-Specific Toxins in the Expression of Virulence by Pathogens.” *Annual Review of Phytopathology* 22 (1): 215–45. https://doi.org/10.1146/annurev.py.22.090184.001243.

Mohd-Assaad, Norfarhan, Bruce A. McDonald, and Daniel Croll. 2016. “Multilocus Resistance Evolution to Azole Fungicides in Fungal Plant Pathogen Populations.” *Molecular Ecology* 25 (24): 6124–42. https://doi.org/10.1111/mec.13916.

Möller, Mareike, and Eva H. Stukenbrock. 2017. “Evolution and Genome Architecture in Fungal Plant Pathogens.” *Nature Reviews Microbiology* 15 (12): 756–71.

https://doi.org/10.1038/nrmicro.2017.76.
Nieuwenhuis, Bps P S, and T Y James. 2016. “The Frequency of Sex in Fungi.” *Phil. Trans. R. Soc. B* 371 (1706): 20150540. https://doi.org/10.1098/rstb.2015.0540.
Norrgard, Karen. 2008. “Genetic Variation and Disease: GWAS.” *Nature Education* 1(1):87. https://www.nature.com/scitable/topicpage/genetic-variation-and-disease-gwas-682.
Oggenfuss, Ursula, Thomas Badet, Thomas Wicker, Fanny E Hartmann, Nikhil Kumar Singh, Leen N Abraham, Petteri Karisto, et al. 2020. “A Population-Level Invasion by Transposable Elements in a Fungal Pathogen.” *BioRxiv*, 2020.02.11.944652. https://doi.org/10.1101/2020.02.11.944652.
Pais, Andrew L, Xu Li, and Qiu-Yun Xiang. 2018. “Discovering Variation of Secondary Metabolite Diversity and Its Relationship with Disease Resistance in Cornus Florida L.” *Ecology and Evolution* 8: 5619–36. https://doi.org/10.1002/ece3.4090.
Palloix, A., V. Ayme, and B. Moury. 2009. “Durability of Plant Major Resistance Genes to Pathogens Depends on the Genetic Background, Experimental Evidence and Consequences for Breeding Strategies.” *New Phytologist* 183 (1): 190–99. https://doi.org/10.1111/j.1469-8137.2009.02827.x.
Palma-Guerrero, Javier, Xin Ma, Stefano F F Torriani, Marcello Zala, Carolina S. Francisco, Fanny E. Hartmann, Daniel Croll, and Bruce A. McDonald. 2017. “Comparative Transcriptome Analyses in Zymoseptoria Tritici Reveal Significant Differences in Gene Expression Among Strains During Plant Infection.” *Molecular Plant-Microbe Interactions* 30 (3): 231–44. https://doi.org/10.1094/MPMI-07-16-0146-R.
Peng, Meng, Raheel Shahzad, Ambreen Gul, Hizar Subthain, Shuangqian Shen, Long Lei, Zhigang Zheng, et al. 2017. “Differentially Evolved Glucosyltransferases Determine Natural Variation of Rice Flavone Accumulation and UV-Tolerance.” *Nature Communications* 8 (1): 1975. https://doi.org/10.1038/s41467-017-02168-x.
Pereira, Danilo, Bruce A McDonald, and Daniel Croll. 2020. “The Genetic Architecture of Emerging Fungicide Resistance in Populations of a Global Wheat Pathogen.” Edited by Li-Jun Ma. *Genome Biology and Evolution* 12 (12): 2231–44. https://doi.org/10.1093/gbe/evaa203.
Petit-Houdenot, Yohann, and Isabelle Fudal. 2017. “Complex Interactions between Fungal Avirulence Genes and Their Corresponding Plant Resistance Genes and Consequences for Disease Resistance Management.” *Frontiers in Plant Science* 8 (June). https://doi.org/10.3389/fpls.2017.01072.
Plissonneau, Clémence, Juliana Benevenuto, Norfarhan Mohd-Assaad, Simone Fouché, Fanny E. Hartmann, and Daniel Croll. 2017. “Using Population and Comparative Genomics to Understand the Genetic Basis of Effector-Driven Fungal Pathogen Evolution.” *Frontiers in Plant Science* 8 (February): 1–15. https://doi.org/10.3389/fpls.2017.00119.
Plissonneau, Clémence, Fanny E Hartmann, and Daniel Croll. 2018. “Pangenome Analyses of the Wheat Pathogen Zymoseptoria Tritici Reveal the Structural Basis of a Highly Plastic Eukaryotic Genome.” *BMC Biology* 16 (1). https://doi.org/10.1186/s12915-017-0457-4.
Plissonneau, Clémence, Alessandra Stürchler, Daniel Croll, and John W Taylor. 2016. “The Evolution of Orphan Regions in Genomes of a Fungal Pathogen of Wheat.” https://doi.org/10.1128/mBio.01231-16.
Pusztahelyi, TÃ¼nde, Imre J. Holb, and IstvÃ¡n PÃ³csi. 2015. “Secondary Metabolites in Fungus-Plant Interactions.” *Frontiers in Plant Science* 6 (August): 573. https://doi.org/10.3389/fpls.2015.00573.
Randles, Lucy G., Gwen J.S. Dawes, Beth G. Wensley, Annette Steward, Adrian A. Nickson, and Jane Clarke. 2013. “Understanding Pathogenic Single-Nucleotide Polymorphisms in

Multidomain Proteins - Studies of Isolated Domains Are Not Enough." *FEBS Journal* 280 (4): 1018–27. https://doi.org/10.1111/febs.12094.

Rohr, Jason R., Jenise Brown, William A. Battaglin, Taegan A. McMahon, and Rick A. Relyea. 2017. "A Pesticide Paradox: Fungicides Indirectly Increase Fungal Infections." *Ecological Applications : A Publication of the Ecological Society of America* 27 (8): 2290. https://doi.org/10.1002/EAP.1607.

Rudd, Jason J., Kostya Kanyuka, Keywan Hassani-Pak, Mark Derbyshire, Ambrose Andongabo, Jean Devonshire, Artem Lysenko, et al. 2015. "Transcriptome and Metabolite Profiling of the Infection Cycle of *Zymoseptoria Tritici* on Wheat Reveals a Biphasic Interaction with Plant Immunity Involving Differential Pathogen Chromosomal Contributions and a Variation on the Hemibiotrophic Lifest." *Plant Physiology* 167 (3): 1158–85. https://doi.org/10.1104/pp.114.255927.

Sánchez-Vallet, Andrea, Fanny E. Hartmann, Thierry C. Marcel, and Daniel Croll. 2018. "Nature's Genetic Screens: Using Genome-Wide Association Studies for Effector Discovery." *Molecular Plant Pathology*. Wiley/Blackwell (10.1111). https://doi.org/10.1111/mpp.12592.

Sánchez-Vallet, Andrea, Megan C. McDonald, Peter S. Solomon, and Bruce A. McDonald. 2015. "Is Zymoseptoria Tritici a Hemibiotroph?" *Fungal Genetics and Biology* 79: 29–32. https://doi.org/10.1016/j.fgb.2015.04.001.

Singh, Nikhil Kumar, Thomas Badet, Leen Abraham, and Daniel Croll. 2021. "Rapid Sequence Evolution Driven by Transposable Elements at a Virulence Locus in a Fungal Wheat Pathogen." *BMC Genomics* 22 (1): 393. https://doi.org/10.1186/s12864-021-07691-2.

Singh, Nikhil Kumar, Emilie Chanclud, and Daniel Croll. 2020. "Population-Level Deep Sequencing Reveals the Interplay of Clonal and Sexual Reproduction in the Fungal Wheat Pathogen Zymoseptoria Tritici." *BioRxiv*, July, 2020.07.07.191510. https://doi.org/10.1101/2020.07.07.191510.

Singh, Nikhil Kumar, Anik Dutta, Guido Puccetti, and Daniel Croll. 2021. "Tackling Microbial Threats in Agriculture with Integrative Imaging and Computational Approaches." *Computational and Structural Biotechnology Journal* 19 (January): 372–83. https://doi.org/10.1016/J.CSBJ.2020.12.018.

Smith, Theobald. 1904. "Some Problems in the Life History of Pathogenic Microorganisms." *Science*. https://doi.org/10.1126/science.20.520.817.

Steinberg, Gero. 2015. "Cell Biology of Zymoseptoria Tritici: Pathogen Cell Organization and Wheat Infection." *Fungal Genetics and Biology : FG & B* 79 (June): 17–23. https://doi.org/10.1016/j.fgb.2015.04.002.

Stewart, Ethan l., Daniel Croll, Mark H. Lendenmann, Andrea Sanchez-Vallet, Fanny E. Hartmann, Javier Palma-Guerrero, Xin Ma, and Bruce A. McDonald. 2018. "Quantitative Trait Locus Mapping Reveals Complex Genetic Architecture of Quantitative Virulence in the Wheat Pathogen *Zymoseptoria Tritici*." *Molecular Plant Pathology* 19 (1): 201–16. https://doi.org/10.1111/mpp.12515.

Stewart, Ethan L., Christina Heber Hagerty, Alexey Mikaberidze, Chris Mundt, Ziming Zhong, and Bruce A. McDonald. 2016. "An Improved Method for Measuring Quantitative Resistance to the Wheat Pathogen Zymoseptoria Tritici Using High Throughput Automated Image Analysis." *Phytopathology* 106 (7): 782–88. https://doi.org/10.1094/PHYTO-01-16-0018-R.

Stewart, Ethan L, and Bruce A McDonald. 2014. "Measuring Quantitative Virulence in the Wheat Pathogen Zymoseptoria Tritici Using High-Throughput Automated Image Analysis." *Phytopathology* 104 (9): 985–92. https://doi.org/10.1094/PHYTO-11-13-0328-R.

Strange, Richard N., and Peter R. Scott. 2005. “Plant Disease: A Threat to Global Food Security.” *Annual Review of Phytopathology* 43 (Figure 1): 83–116. https://doi.org/10.1146/annurev.phyto.43.113004.133839.

Strauch, Renee C, Elisabeth Svedin, Brian Dilkes, Clint Chapple, and Xu Li. 2015. “Discovery of a Novel Amino Acid Racemase through Exploration of Natural Variation in Arabidopsis Thaliana.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (37): 11726–31. https://doi.org/10.1073/pnas.1503272112.

Stukenbrock, Eva H., Søren Banke, Mohammad Javan-Nikkhah, and Bruce A. McDonald. 2007. “Origin and Domestication of the Fungal Wheat Pathogen Mycosphaerella Graminicola via Sympatric Speciation.” *Molecular Biology and Evolution* 24 (2): 398–411. https://doi.org/10.1093/molbev/msl169.

Stukenbrock, Eva H., Frank G. Jørgensen, Marcello Zala, Troels T. Hansen, Bruce A. McDonald, and Mikkel H. Schierup. 2010. “Whole-Genome and Chromosome Evolution Associated with Host Adaptation and Speciation of the Wheat Pathogen Mycosphaerella Graminicola.” *PLoS Genetics* 6 (12): 1–13. https://doi.org/10.1371/journal.pgen.1001189.

Subbarao, Krishna V, George W Sundin, and Steven J Klosterman. 2015. “Focus Issue Articles on Emerging and Re-Emerging Plant Diseases.” https://doi.org/10.1094/PHYTO-105-7-0001.

Tack, A. J.M., P. H. Thrall, L. G. Barrett, J. J. Burdon, and A. L. Laine. 2012. “Variation in Infectivity and Aggressiveness in Space and Time in Wild Host-Pathogen Systems: Causes and Consequences.” *Journal of Evolutionary Biology* 25 (10): 1918–36. https://doi.org/10.1111/j.1420-9101.2012.02588.x.

Talas, Firas, Rasha Kalih, Thomas Miedaner, and Bruce A. McDonald. 2016. “Genome-Wide Association Study Identifies Novel Candidate Genes for Aggressiveness, Deoxynivalenol Production, and Azole Sensitivity in Natural Field Populations of Fusarium Graminearum.” *Molecular Plant-Microbe Interactions* 29 (5): 417–30. https://doi.org/10.1094/MPMI-09-15-0218-R.

Testa, Alison, Richard Oliver, and James Hane. 2015. “Overview of Genomic and Bioinformatic Resources for Zymoseptoria Tritici.” *Fungal Genetics and Biology* 79: 13–16. https://doi.org/10.1016/j.fgb.2015.04.011.

Thrall, Peter H., and Jeremy J. Burdon. 2003. “Evolution of Virulence in a Plant Host-Pathogen Metapopulation.” *Science* 299 (5613): 1735–37. https://doi.org/10.1126/science.1080070.

Tralamazza, Sabina Moser, Liliana Oliveira Rocha, Ursula Oggenfuss, Benedito Corrêa, and Daniel Croll. 2019. “Complex Evolutionary Origins of Specialized Metabolite Gene Cluster Diversity among the Plant Pathogenic Fungi of the Fusarium Graminearum Species Complex.” Edited by Laura Rose. *Genome Biology and Evolution* 11 (11): 3106–22. https://doi.org/10.1093/gbe/evz225.

Weiberg, Arne, Ming Wang, Feng-Mao Lin, Hongwei Zhao, Zhihong Zhang, Isgouhi Kaloshian, Hsien-Da Huang, and Hailing Jin. 2013. “Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways.” *Science (New York, N.Y.)* 342 (6154): 118–23. https://doi.org/10.1126/science.1239705.

Zhan, J., C. C. Linde, T. J??rgens, U. Merz, F. Steinebrunner, and B. A. McDonald. 2005. “Variation for Neutral Markers Is Correlated with Variation for Quantitative Traits in the Plant Pathogenic Fungus Mycosphaerella Graminicola.” *Molecular Ecology* 14 (9): 2683–93. https://doi.org/10.1111/j.1365-294X.2005.02638.x.

Zhan, J, C C Mundt, and B A Mcdonald. 1998. “Ecology and Population Biology Measuring Immigration and Sexual Reproduction in Field Populations of Mycosphaerella Graminicola.” https://apsjournals.apsnet.org/doi/pdf/10.1094/PHYTO.1998.88.12.1330.

Zhang, Qing-Yu, Li-Qing Zhang, Li-Li Song, Ke Duan, Na Li, Yan-Xiu Wang, and Qing-Hua Gao. 2016. “The Different Interactions of Colletotrichum Gloeosporioides with Two Strawberry Varieties and the Involvement of Salicylic Acid.” *Horticulture Research* 3 (1): 16007. https://doi.org/10.1038/hortres.2016.7.

Zhong, Ziming, Thierry C. Marcel, Fanny E. Hartmann, Xin Ma, Cl??mence Plissonneau, Marcello Zala, Aur??lie Ducasse, et al. 2017. “A Small Secreted Protein in Zymoseptoria Tritici Is Responsible for Avirulence on Wheat Cultivars Carrying the Stb6 Resistance Gene.” *New Phytologist* 214 (2): 619–31. https://doi.org/10.1111/nph.14434.

Chapter 1

Population-level deep sequencing reveals the interplay of clonal and sexual reproduction in the fungal wheat pathogen *Zymoseptoria tritici*

Nikhil Kumar Singh[1], Petteri Karisto[2], Daniel Croll[1]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland

[2] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland.

Published in Microbial Genomics: Volume 7, Issue 10

https://doi.org/10.1099/mgen.0.000678

Data summary

All Illumina sequence data is available from the NCBI SRA BioProject number PRJNA596434 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA596434). The Supplementary Tables S1-S3 list the exact strain names, collection location, genotype and genetic diversity indices.

CHAPTER 1

Abstract

Pathogens cause significant challenges to global food security. On annual crops, pathogens must re-infect from environmental sources at every growing season. Fungal pathogens evolved mixed reproductive strategies to cope with the distinct challenges of colonizing growing plants. However, how pathogen diversity evolves during growing seasons remains largely unknown. Here, we performed a deep hierarchical sampling in a single experimental wheat field infected by the major fungal pathogen *Zymoseptoria tritici*. We analyzed whole genome sequences of 177 isolates collected from twelve distinct cultivars replicated in space at three time points of the growing season to maximize capturing genetic diversity. The field population was highly diverse with 37 SNPs per kilobase and a linkage disequilibrium decay within 200-700 bp. Using experimental infections, we tested a subset of the collected isolates on the dominant cultivar planted in the field. However, we found no significant difference in virulence of isolates collected from the same cultivar compared to isolates collected on other cultivars. About 20% of the isolate genotypes were grouping into 15 clonal groups. Pairs of clones were disproportionally found at short distances (<5m) consistent experimental estimates for per-generation dispersal distances performed in the same field. This confirms a predominant leaf-to-leaf transmission during the growing season. Surprisingly, levels of clonality did not increase over time in the field although reproduction is thought to be exclusively asexual during the growing season. Our study shows that the pathogen establishes vast and stable gene pools in single fields. Monitoring short-term evolutionary changes in crop pathogens will inform more durable strategies to contain diseases.

Introduction

Infectious diseases have a major impact on plant and animal populations including humans. Designing effective strategies for disease prevention requires knowledge of the evolutionary dynamics of pathogens. The evolutionary response of pathogen populations to novel resistance factors and changing environmental conditions are critical (Gandon *et al.*, 2013). Responses are often a function of genetic diversity in pathogens with consequences for disease severity and the spread in heterogeneous host populations (Read & Taylor, 2001; Thrall & Burdon, 2003; Vignuzzi et al., 2006). Highly resistant host populations tend to harbour the most virulent pathogens, whereas less virulent pathogens dominate susceptible populations (Thrall & Burdon, 2003). In addition to selection by host genotypes, pathogen populations are subject to selection at life cycle stages outside of the host (Walther & Ewald, 2004; Linde, 2010; Hamelin *et al.*, 2011) and other environmental factors including the application of chemical compounds such as fungicides (He *et al.*, 2016; Sundin *et al.*, 2016). Pathogens of annual plants must survive periods of time in the absence of a host and succeed in re-infecting new plants (Laine & Barrès, 2013; Mariette *et al.*, 2016). During the growing season of the host, pathogens are selected to colonize additional plants. Hence, evolutionary trajectories are expected to vary in space and time depending on selective pressures imposed by both the host and the environment (Tack *et al.*, 2012; Becheler *et al.*, 2016). Beyond selection, small pathogen populations are likely experiencing genetic bottlenecks or gene flow. Hence, it remains largely unknown how pathogen populations evolve over short time periods (Barrett et al., 2008).

In agricultural crop-pathogen systems, infections begin with the exposure of a susceptible host to pathogen strains at the beginning of the growing season. The intensity of an infection depends on the level of host resistance, the amount of pathogen inoculum and environmental

factors (Burdon *et al.*, 2016; Suffert *et al.*, 2019). If the infection is successful, pathogens complete their life cycle within the host and then transmit to another host. The genetic architecture of traits to successfully passage between infection stages is likely distinct (Jones & Dangl, 2006; Dodds & Thrall, 2009). Entering a host plant requires the manipulation of the host immune system (De Torres *et al.*, 2006; Sun *et al.*, 2006; Nomura *et al.*, 2006). Proliferation and reproduction inside the plant depends likely on optimal resource exploitation and continued suppression of immunity (Chisholm *et al.*, 2006; Steinberg, 2015). Dispersal between hosts is governed by propagule properties (Peay & Bruns, 2014; Parratt *et al.*, 2016). As a consequence, pathogen populations likely experience complex selection pressures over the life cycle. Depending on the mode of between-season transmission and environmental conditions, plant pathogen populations experience severe demographic bottlenecks during the off-season restricting the transmission of genetic variation between growing seasons (Barrett *et al.*, 2008; Susi *et al.*, 2020). After a filtering of genotypes for persistence in the off-season environment, the surviving genotypes undergo selection for infectivity and spread among hosts during the subsequent epidemic phase. In any such scenario, pathogens must re-infect crops annually and successfully spread among plants in a field. Therefore, considering the time-scale of pathogen population turnover and demographic effects is important.

The mode of dispersal is often tightly associated with specific reproductive modes. Pathogens often have mixed reproductive strategies meaning that sexual and asexual reproduction can alternate over the course of the life cycle adding to the complexity of recognizing species and challenging genetic diversity analyses (McDonald, 1997; Billiard *et al.*, 2012). Fungal crop pathogens often reproduce asexually during the growing season to maximize dispersal at low population density. In turn, the asexual reproduction during the epidemic phase has likely a strong influence on genetic structure of pathogen populations (Barrett *et al.*, 2008). On

senescent plants, sexual reproduction can lead to favourable allele combinations for long-distance dispersal and survival (McDonald & Linde, 2002). An important pathogen with a mixed reproductive strategy is the fungal wheat pathogen *Zymoseptoria tritici*. The fungus is a haploid ascomycete and one of the most destructive pathogens of wheat with yield losses of 5–30% (Fones & Gurr, 2015). Populations worldwide harbour significant variation in pathogenicity and diversity at loci underlying host specificity (Zhong *et al.*, 2017; Hartmann *et al.*, 2017; Hartmann & Croll, 2017; Krishnan *et al.*, 2018). The pathogen follows typically a sequential regime of sexual and asexual reproduction cycles over the growing season. During the sexual stage, the pathogen produces asci containing four pairs of genetically distinct ascospores, which are thought to be the source of primary infection in the field and can disperse over long distances (Chen & Mcdonald, 1996; Kema *et al.*, 1996). Once infections are established in a field, *Z. tritici* produces asexual pycnidiospores, which can disperse with rain splash and initiate secondary infections nearby. The spread within fields during the growing season is thought to be predominantly clonal (Chen & Mcdonald, 1996). Population genomics analyses of worldwide pathogen samples showed high genetic diversity within fields and extensive gene flow (Hartmann *et al.*, 2017, 2018; McDonald *et al.*, 2019). Multi-year studies of *Z. tritici* populations revealed trade-offs between the intra- and interannual scales in the evolution of aggressiveness (Suffert *et al.*, 2018) with re-infections at the beginning of the growing season imposing no detectable genetic bottleneck (Morais *et al.*, 2019). Consistent with the maintenance of high levels of diversity, individual wheat leaves are frequently infected by multiple different genotypes (Linde *et al.*, 2002). However, whole genome analyses of hierarchical field collections to resolve the spatial and temporal changes in genotypic diversity have been lacking.

In this study, we have whole-genome sequenced 177 distinct isolates of *Z. tritici* from a single wheat field covering three time points over the growing season. We analyzed twelve distinct cultivars replicated in space to maximize capturing genetic diversity. Based on genome-wide polymorphisms, we reconstructed how different chromosomal compartments contributed to overall diversity. Using information about the spatial organization in the field and experimental assessments of dispersal distances, we quantified the extent of clonal spread in space and time.

Results

Pathogen field population shows extensive sequence variation

A total of 177 isolates of *Z. tritici* were collected from an experimental wheat field in Switzerland planted with 335 wheat cultivars in 1.2-by-1.7m plots (Supplementary Table S1; Karisto *et al.*, 2018). We collected isolates at three time points during the wheat growing season representing different phases in the infection life cycle of the pathogen (Fig 1A; Karisto *et al.*, 2018). The isolates from the first collection (C1) were exposed to the application of demethylation inhibitor (DMI) fungicides. Isolates from the second and third collection were exposed additionally to a succinate dehydrogenase inhibitor (SDHI) and additional DMIs (Fig 1A). We collected isolates from 12 winter wheat cultivars (n = 1-43 isolates per cultivar) which were planted in two replicate plots separated by ~100 m (Fig 1B-C). The cultivar Claro was sampled on two additional plots. The collection was performed when the wheat plants were in growth stages GS 41 (C1), GS 75 (C2) and GS 85 (C3), respectively (Meier, 2001; Karisto *et al.*, 2018). We Illumina sequenced whole genomes for all 177 isolates producing 0.8-15 million reads per isolate with an average read count of 4.2 million (Fig 1D). The isolates were sequenced to a depth of 5-77 X and a mean coverage of 21.4 X. We detected a total of 1’496’037 high-confidence SNPs. The average genotyping rate was 97.85 % across loci (Fig 1E). We detected on average 37 SNPs per kb showing that the field population is highly polymorphic. In line with this, we found that the minor allele (*i.e.* the less frequent allele at a locus) frequency spectrum showed a strong skew towards rare alleles in the population, which suggests that the population did not experience any recent bottlenecks (Fig 1F). Population bottlenecks are expected to generate an excess of high-frequency alleles. Based on the finding that our population shows substantial genetic variation, we aimed to estimate the total genetic diversity in the field. For this, we performed random down-sampling analysis and we found that subsets of 10% and 50% of the population harboured on average 763'765 and 1'234'846

segregating SNPs representing 51% and 82.5% of the total detected SNPs, respectively (Fig. 1G). We observed only a weak plateau effect for subsets close to 100% indicating that the total polymorphism in the field population is likely substantially higher.

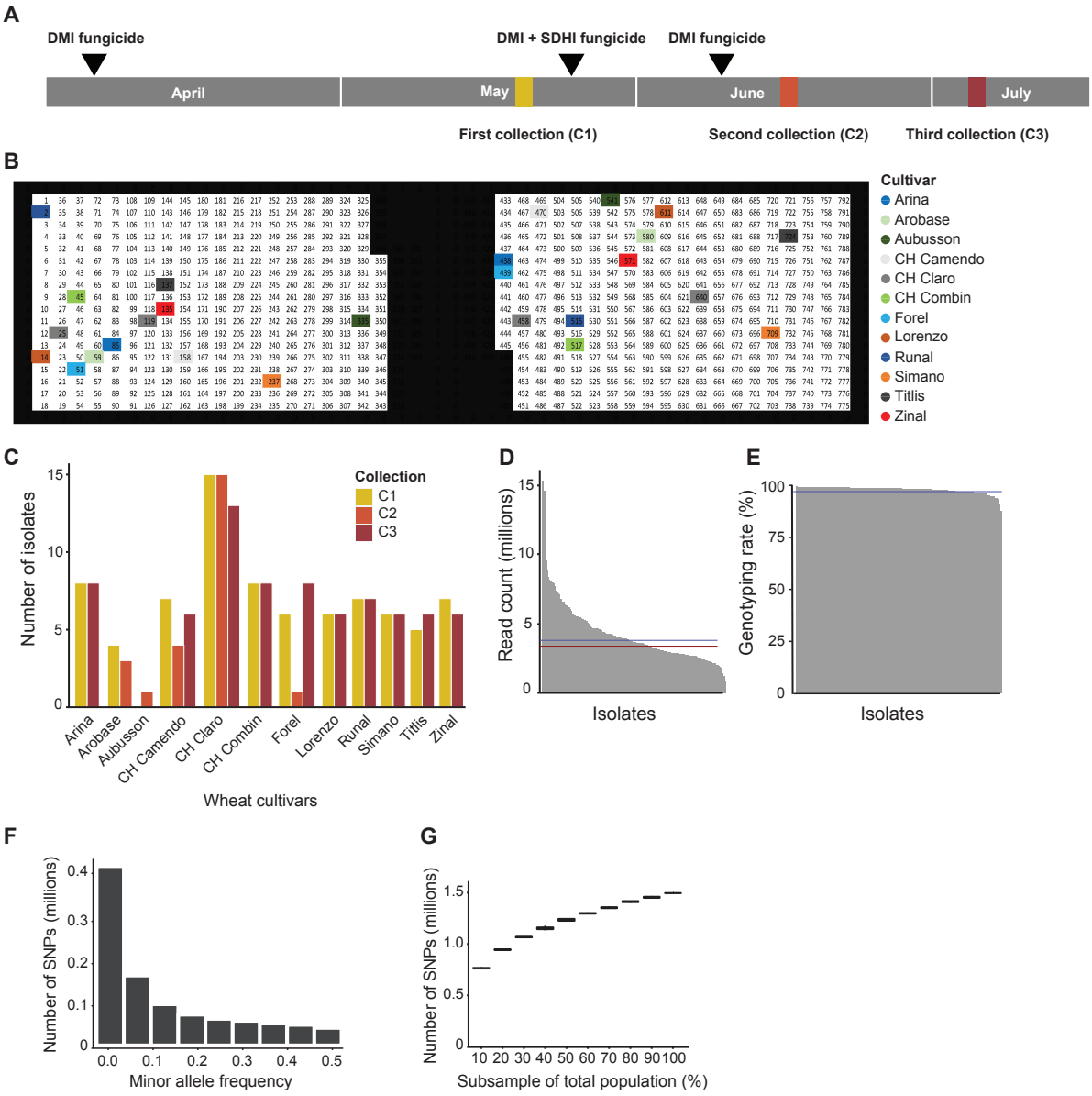


Figure 1: Genome sequencing of 177 *Zymoseptoria tritici* isolates from a single field. A) Time scale showing the three collection points from May to July. The experimental field was treated with fungicides thrice (see Methods for details). B) Graphical representation of the experimental wheat field showing the sampled cultivars (colored background). Each number indicates a plot of 1.2-by-1.7 meters. Wheat cultivar color codes are maintained throughout the manuscript. Sampled cultivar numbers are also shown in Supplementary Table S4. C) Number of isolates collected from each of the 12 cultivars at each collection time point. D) Trimmed read count for each isolate. Lines represent the mean (blue) and median (red) read counts. E) Genotyping rate of all the 177 isolates (% loci genotyped). The blue line represents the mean genotyping rate. F) Minor allele frequency (frequency at which the second most common allele

occurs in a given population) spectrum of 1'496'037 SNPs genotyped in 177 isolates. G) Boxplot showing the number of segregating SNPs in subsets of the 177 isolates. Error bars show 10 replicates of resampled isolates.

Accessory chromosome polymorphism

Z. tritici harbours the most extensive set of accessory chromosomes known for plant pathogens (Croll & McDonald, 2012). We analyzed chromosome-wide read coverage to identify chromosomal presence-absence variation. All core chromosomes (1-13) were consistently detected in all isolates as expected. In contrast, we found 139 isolates (~79%) missing one or more accessory chromosomes (Fig 2A). We found no isolate that lacked all accessory chromosomes, but we identified two isolates with only a single accessory chromosome out of eight. Accessory chromosome 16 was the most frequent (174/177 isolates) and chromosome 18 was the rarest being present in less than half of the isolates (Fig 2B). We identified several accessory chromosomes showing evidence for partial deletions or duplications based on normalized read coverage (Supplementary Figure S1). Specifically, we identified two partial deletions of chromosome 17 and 20 (Fig 2C), as well as chromosomal duplications of chromosome 17, 18 19 and 21 (Fig 2C). In a next step, we analyzed how polymorphism was structured among core and accessory chromosomes. We found that the accessory chromosomes show higher SNP densities, but lower total SNP counts compared to core chromosomes (Fig 2D-E). Overall, smaller chromosomes tend to have higher SNP densities (Fig 2F). This is consistent with relaxed selection leading to accumulation of mutations on these chromosomes (Stukenbrock *et al.*, 2010). It is important to note though that short read-based SNP calling is not feasible in highly repetitive chromosomal regions. Consistent with this, we found gaps in high-confidence SNP calls along all chromosomes but the effect was particularly pronounced on accessory chromosomes (Supplementary Figure S2). Hence, our polymorphism analyses likely underestimate SNP densities in particular on accessory chromosomes.

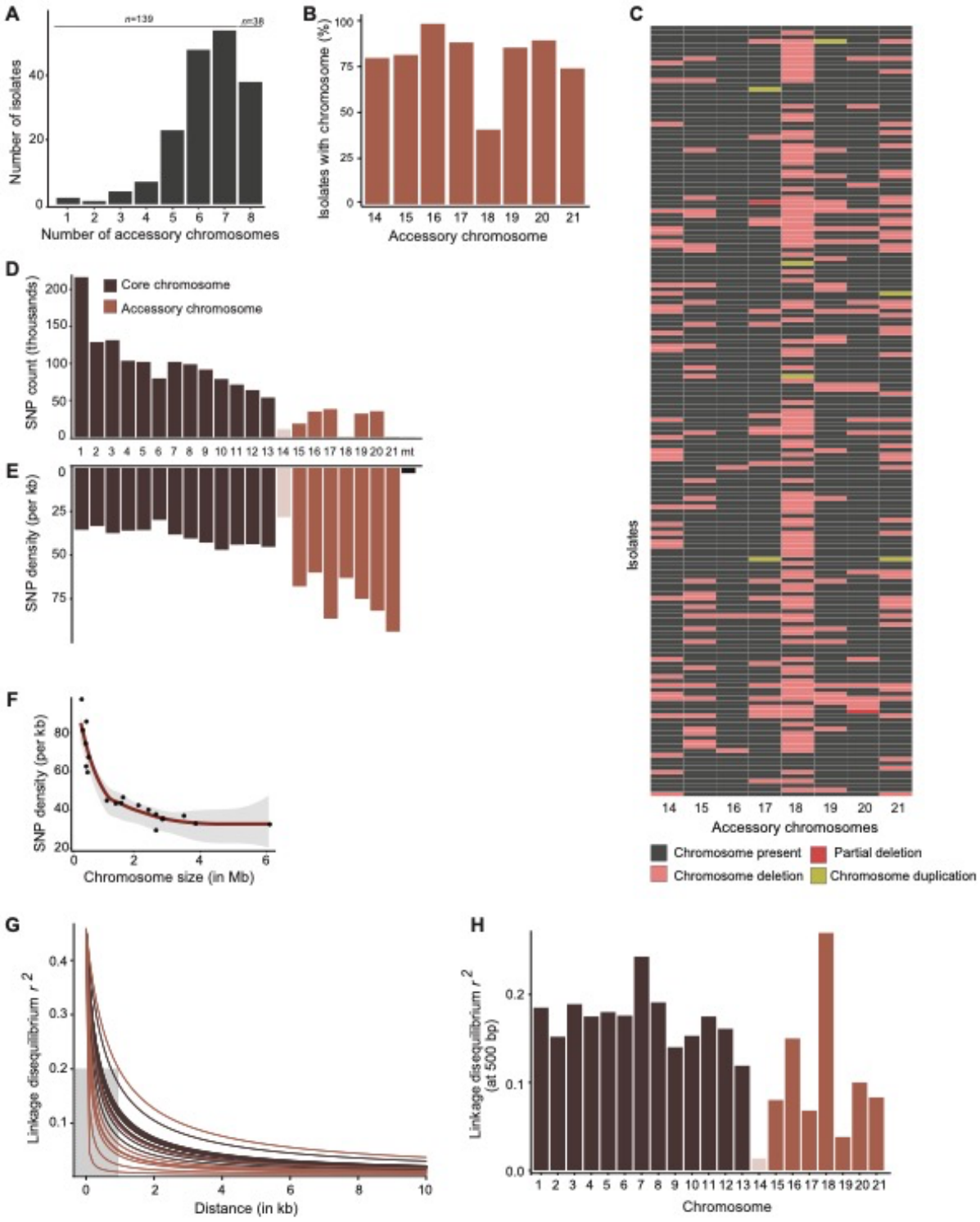


Figure 2: Genome-wide polymorphism and accessory chromosomes. A) Total number of accessory chromosomes per isolate. The numbers above the bars indicate the isolates missing at least one accessory chromosome or carry all chromosomes. B) Percentage presence of all the accessory chromosomes within the population calculated on the basis of normalized chromosome coverage. C) Heat map showing the presence-absence variation in accessory chromosomes across isolates assessed by normalized read depth. D) SNP count and E) SNP density per chromosome. F) Correlation plot of SNP density and chromosome size. G) Pairwise linkage disequilibrium decay among all pairs of SNPs within a fixed window size of 10'000 bp

for each chromosome. A non-linear model was fitted using the equation of Ingvarsson (2005). The grey shading indicates the maximum distance needed for all chromosomes to reach r^2 = 0.2. H) Linkage disequilibrium r^2 for each chromosome at 500 bp. Light shade: chromosome 14 was omitted from per chromosome genetic diversity analyses due to the heterogeneous distribution of regions with robust SNP calls.

Another important component of genetic variation is the extent of linkage disequilibrium. We estimated linkage disequilibrium decays for each chromosome separately and found that most accessory chromosomes showed a faster decay compared to core chromosomes (r^2 = 0.2 within 100 bp, Fig 2G), with the exception of chromosome 16 and 18 showing a decay of r^2 to 0.2 within 300 and 900 bp. Interestingly, chromosome 7 showed the slowest decay of all core chromosomes (r^2 = 0.2 at ~700 bp, Fig 2H). This chromosome was proposed to have originated at least partially from an accessory chromosome (Schotanus *et al.*, 2015). Chromosome 18 shows a particularly low degree of collinearity among genomes of the same species (Badet *et al.*, 2020a). (Supplementary Figure S3). It remains to be investigated whether structural variation had an impact on linkage disequilibrium assessments, in particular on accessory chromosomes. Taken together, we show that the accessory chromosomes accumulated more sequence variation compared to core chromosome, but the level of conservation varies substantially among accessory chromosomes.

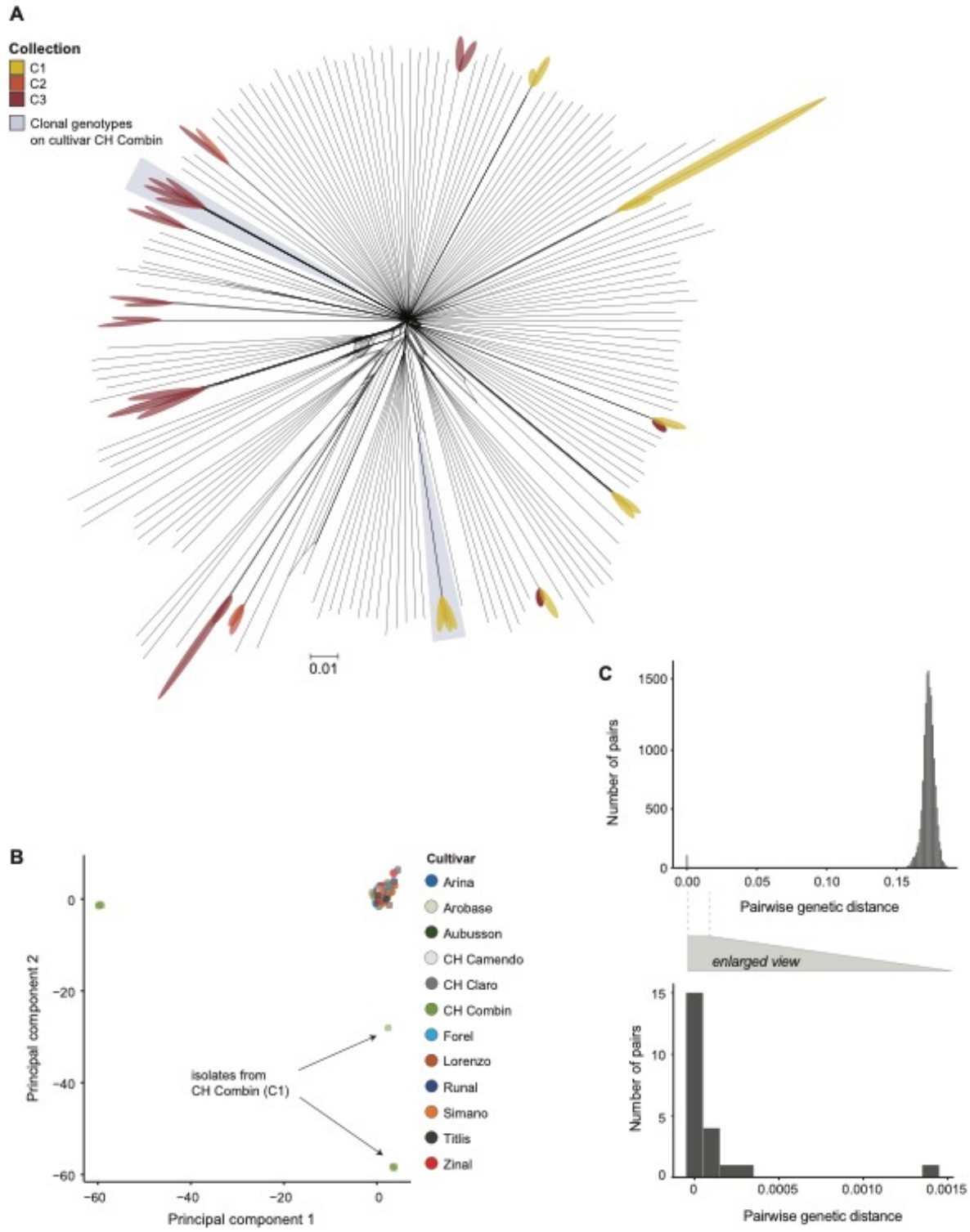


Figure 3: Genetic differentiation and clonality within the field. A) Phylogenetic network constructed using SplitsTree. Groups of clonal genotypes are marked with highlights identifying the collection of origin. B) The first two principal components (PC) from a PC analysis. Isolates are color coded by the cultivar of origin. C) Histograms showing the distribution of pairwise genetic distances among genotypes. A distance of 1 corresponds to the total number of SNPs. D) Histogram of pairwise genetic distances < 0.01.

Population clonality in space and time

To track the evolution of genotypic diversity in space and time, we first constructed an unrooted phylogenetic network based on SplitsTree and found that most pairs of genotypes are at a similar genetic distance (Fig 3A). The star-like structure of the network is consistent with frequent recombination. A subset of genotypes showed much shorter genetic distances suggesting recent ancestry or clonal reproduction. We also performed a PCA to determine the degree of differentiation in the population (Fig 3B). We detected no meaningful population subdivision except for seven isolates, which clustered into 3 groups (Fig 3B). Interestingly, all seven isolates were collected from cultivar CH Combin (arrows in Fig 3B). To clarify how clonal reproduction impacts the genetic structure within the field, we identified groups of clonal genotypes based on pairwise genetic distances. We found that most isolates were at a relative pairwise distance of more than 0.15 where a value of 1 corresponds a genetic difference at every analyzed SNP position (Fig 3C). We defined isolate pairs with a genetic distance of <0.01 as being clones (Fig 3D). We assigned groups of clones into four categories according to the distance between field plots and collection time point: Clones originating from the same plot and same collection time point (category A), different plot but same collection time point (category B), same plot but different collection time point (category C) and different plot and different collection time point (category D; Fig 4A, Supplementary table S2).

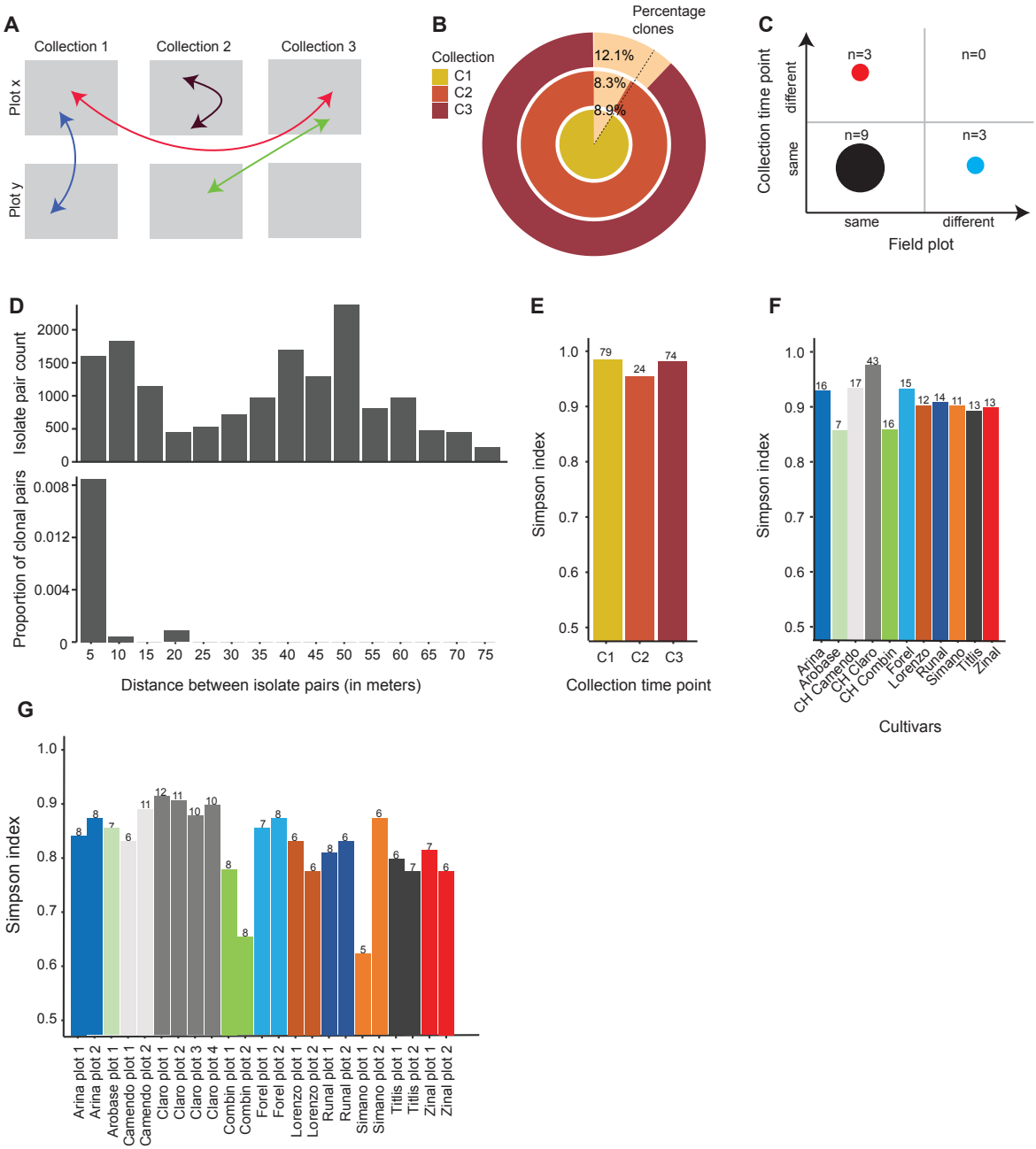


Figure 4: Genetic diversity through space and time. A) Schematic representation of different clone categories with pairs identified from the same or different plots and time points. B) Pie chart showing the change in clonality over time. The dotted line indicates the average clonality of the whole population. C) Number of clone groups categorized based on differences in space and time. The numbers indicate the clone groups in each category. D) Analyses of all possible pairwise comparisons of isolates with the distance between collection points shown as a histogram. The proportion of clone pairs is shown as a proportion of the total number of isolate pairs. E) Simpson diversity index between collections C1-C3, F) between cultivars and G) plots.

We identified a total of 15 distinct groups of clones (~8.5% of the isolates; Fig 4B; Supplementary table S2). The proportion of clones ranged 8.9% in first collection to 12.1% in third collection but the difference was not statistically significant (Fisher exact test, *p-value* = 0.244; Fig 4B). The majority of the clone groups ($n = 9$) were sampled from the same plot and same collection time point (category A; Fig 4C; Supplementary Table S2). Interestingly, isolates collected from the cultivar CH Combin showed a high degree of clonality (Supplementary Table S2). We also found that 3 clonal groups each were either collected from the same plot at different time points or different plots at the same time point (category B and C; Fig 4C). The maximum distance at which we identified clonal pairs was 20 meters (Supplementary Table S2) although distances between plots ranged up to 120 m (Fig 5A).

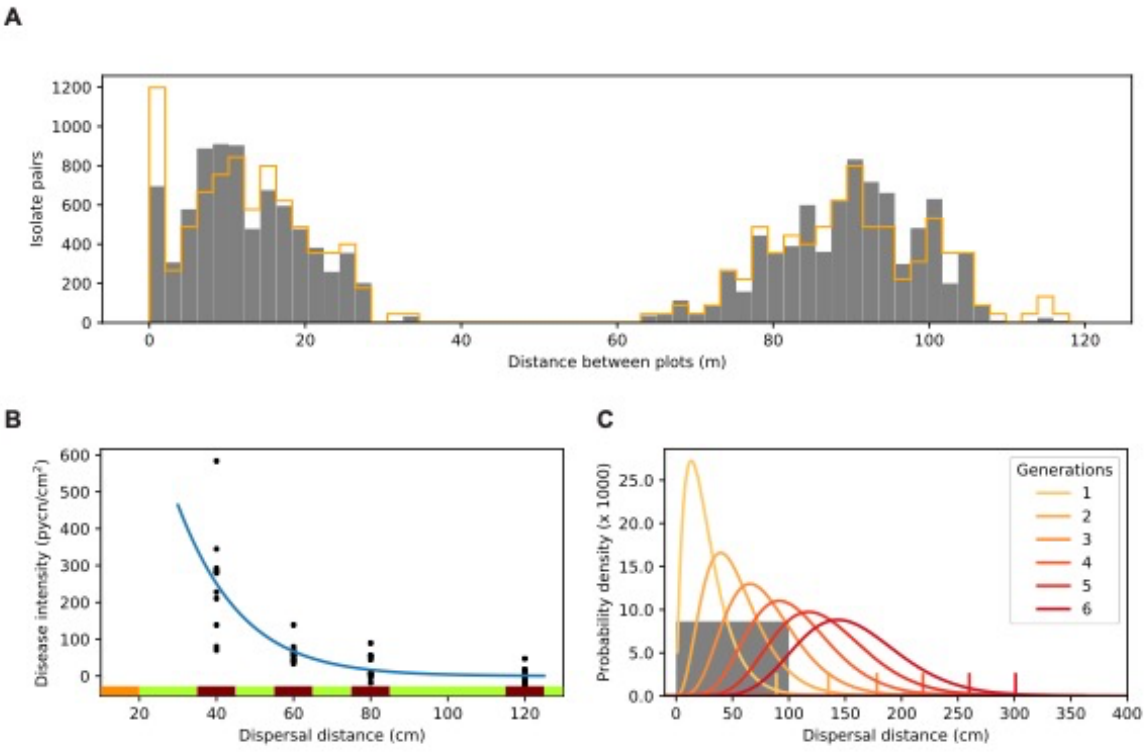


Figure 5. Estimation of dispersal distances. A) Observed distribution of distances between all isolates (grey) and the expected distribution assuming a uniform sampling (orange). The bimodal distribution results from the two-block design of the experiment. B) Experimental assessment of dispersal distances based on data generated by Karisto et al. 2021 (DOI pending) showing the observed primary disease gradient from an infected source (orange) to measurement lines (red) at specific distances from the source. C) Spatial distribution of clonal

pairs (grey) compared to distributions of the expected dispersal distances after successive generations.

Estimating dispersal distances of clonal genotypes

We analyzed all pairwise distances of isolate sampling locations across the field. Given the spatial arrangement of the analyzed plots, we found a bimodal distribution of pairwise distances (Fig 5A). Next, we used data collected from an asexual spore dispersal experiment conducted in the same field (Karisto et al. 2021, bioRxiv DOI pending). In this experiment, a focal infection with a known genotype was established and the radiation of the infection focus was monitored. Disease intensity substantially decreased already at the scale of a meter (Fig 5C). Based on this information, we compared the distribution of the detected clonal pairs against the experimentally estimated dispersal distances in successive generations of asexual propagation (Karisto et al. 2021, DOI pending; Fig 5B). We could not re-estimate the dispersal kernel of the pathogen due to low number of clonal isolates and low spatial resolution of the sampling. Comparing the distances between the observed clonal isolates to the expected dispersal distance distributions suggests that most clonal pairs are likely originating from recent asexual dispersal. However, the longest distance of 20 m between clones is very unlikely to have occurred by splash dispersal. Human or animal movements within the field are the most likely explanations.

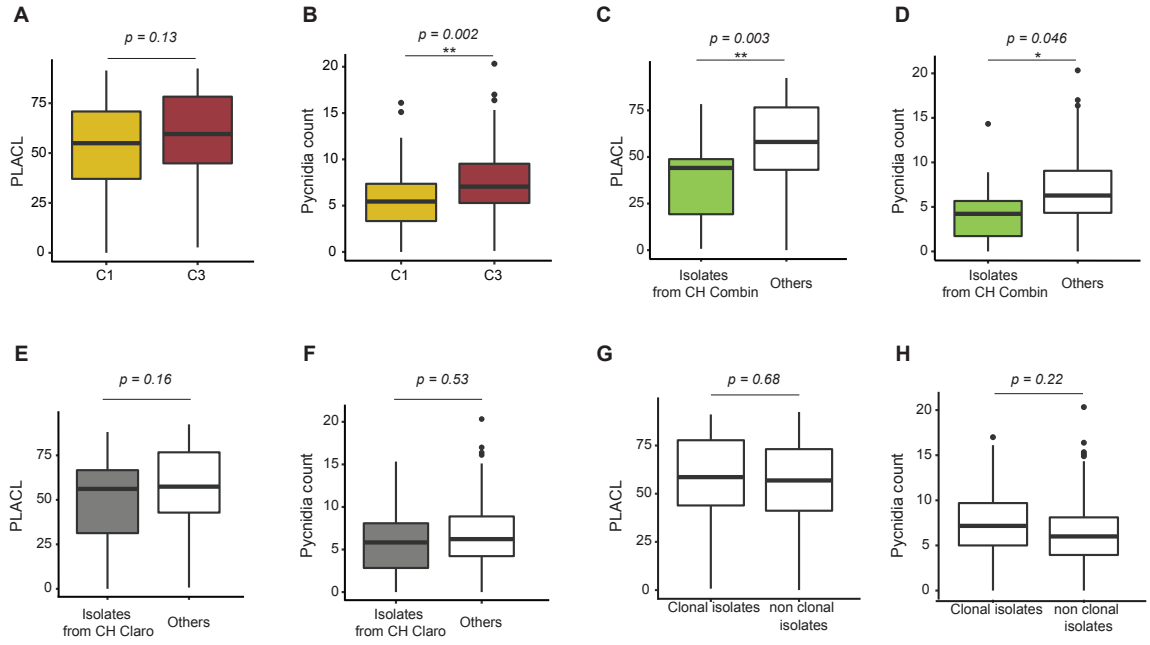


Figure 6: Experimental assessment of virulence on the wheat cultivar CH Claro. Percentage of leaf area covered by lesions (PLACL) and pycnidia counts are show. A-B) Isolates from different collection time points. C-D) Isolates originally collected from the wheat cultivar CH Combin compared to isolates collected from other cultivars. E-F) Isolates originating the from wheat cultivar CH Claro compared to isolates collected from other cultivars. G-H) Isolates with a genotype grouped into a clone group compared to other isolates.

Genotypic diversity assessments

We analyzed the genotypic diversity in the field across time points using diversity indices. Overall, we found less diversity in the third collection compared to the first collection (Fig 4E Supplementary Table S3). The second collection contains too few isolates for comparison. The genotypic diversity was highest for cultivar Claro (0.97, Fig 4F) constituting also the best sampled cultivar in the field. Isolates from cultivar Combin had the lowest diversity (0.86, Fig 4F) due to the fact that almost all recovered genotypes were clones. Both field plots cultivar Combin showed low diversity with a Simpson index of 0.78 and 0.66, respectively (Fig 4G). We did not consider the wheat cultivar Aubusson due to low sampling size (n = 1). Overall, clonal genotypes are an important constituent of the genetic structure in the field. However, the exclusive asexual reproduction during the growing season does not meaningfully affect the trajectory of genetic diversity.

Experimental analyses of virulence on cultivar Claro

We tested using greenhouse assays whether isolates sampled from different time points or cultivars showed differences in virulence on the most frequent cultivar (CH Claro) planted in the field. For this, we analyzed data generated for a genome-wide association study from a subset (n = 120) of the isolates (Singh et al. 2021, DOI pending; Supplementary Table S5). We found that isolates from the third collection to produce significantly more pycnidia (a proxy for pathogen reproduction) compared to earlier collection (one-way ANOVA, p-value = 0.002) (Figure 6A-B). Clonal isolates recovered from the cultivar CH Combin were found to be causing significantly less lesion damage (*i.e.* percent leaf area covered by lesions; one-way ANOVA, p-value = 0.003) and produced less pycnidia (one-way ANOVA, p-value = 0.046; Figure 6C-D). However, we found no significant difference in virulence traits between isolates sampled from CH Claro compared to other wheat cultivars (Figure 6E-F). We found no significant difference in virulence traits between isolates found in clone groups or not (Figure 6G-H)

Discussion

The mode of reproduction and levels of genetic diversity play an important role in the rapid evolution of plant pathogens. We analyzed a large collection of *Z. tritici* genomes obtained from a single wheat field and found that the population was highly diverse with a rapid decay in linkage disequilibrium. We found also a minor degree of clonal structure most likely driven by splash dispersal among neighbouring wheat plants.

Genome-wide levels of genetic diversity in a single field

The analyzed field population shows a dominant sexual reproduction regime and is highly polymorphic. Using hierarchical sampling from the same field plots and wheat cultivars across different time points, we found that ~80% of the isolates were genetically distinct. The allele frequency spectrum and the rapid decay in linkage disequilibrium suggest that there has been no obvious recent population bottleneck. Linkage disequilibrium decayed to $r^2 < 0.2$ within 1.5 kb across all chromosomes and often at a much shorter distance. Since the average distance between genes in the genome of *Z. tritici* is approximately 1 kb (Goodwin *et al.*, 2011), polymorphism in any gene is expected to evolve largely independently. Hence, polymorphism in genes playing roles in adaptation to either host resistance or environmental adaptation should respond independently to selection pressures allowing the population to adapt very rapidly (Slatkin, 2008; Barton, 2010). We have also shown that the rate of linkage disequilibrium decay is not the same in all chromosomes with accessory chromosomes showing faster decays. In *Z. tritici*, accessory chromosomes are small, low in G/C and gene content and have higher degrees of repetitive DNA (Goodwin *et al.*, 2011). In line with this, accessory chromosomes have been shown to have higher recombination rates than core chromosomes (Croll *et al.*, 2015). Analyses of polymorphism and linkage disequilibria may have been affected by the low degree of synteny (or collinearity) among accessory chromosomes and the high repetitive

sequence content (Badet *et al.*, 2020a,b). Genotyping using short read data is generally only possible in the least repetitive sections of accessory chromosomes. Substantial insertion and deletion polymorphism changes the physical distances among SNP loci, which could also have affected estimates of linkage disequilibrium. Rapid decay in linkage disequilibrium is an important property also in other rapidly evolving pathogens such as *Leptosphaeria maculans* causing blackleg disease in oilseed rape. Rapid allelic diversification of avirulence genes driven by sexual reproduction can lead to resistance breakdowns of the host (Daverdin *et al.*, 2012). Similarly, the poplar rust *Melampsora larici-populina* shows signs of rapid sweeps of virulent alleles associated with population replacements (Persoons *et al.*, 2017).

Clonal genotype contributions in space and time

Sexual ascospores originating from crop residues of the previous cropping season are thought to be the source of primary infections in the field (Chen & Mcdonald, 1996; Kerdraon *et al.*, 2019). As a result of this, a wheat field infected by *Z. tritici* is expected to carry essentially no clonal genotypes at its initial colonization stage. Clonal genotypes are expected to increase in frequency as a result of splash-dispersed asexual pycnidiospores. However, we found no significant increase in clonality despite the fact that asexual reproduction should dominate over the course of a growing season. In addition to the sampling period, we found that the likelihood to observe clonal genotypes did not meaningfully change with the cultivar or physical distance. This means that even at the level of individual plots of ~2 m^2, wheat was colonised nearly entirely by genetically distinct isolates. The pairs of clonal isolates found at the shortest distance (0 and 2-3m distance classes) are most likely the result of splash dispersal. This was supported by the experimental assessment of splash dispersal distances in the same field showing a per generation dispersal distance of well under a meter. Notable exceptions included the clonal genotypes associated with the cultivar CH Combin and clonal genotypes on different

cultivars physically separated by as much as 20 meters. Given our estimates of splash dispersal distances, human (or animal) movement within the field is a more plausible explanation for these unusual dispersal event. We also identified clone groups with isolates collected in the same plot but from different collection time points. Hence, successful genotypes are able to persist in individual field plots at high enough levels through time to be resampled.

It is currently unknown how adaptation to environmental factors (including host genotypes) influences the evolution of genetic diversity in infected fields. Selection due to seasonal changes can lead to short-term increases in aggressiveness at the annual scale (Suffert *et al.*, 2015), however this effect can disappear at the interannual scale (Suffert *et al.*, 2018). Genotypes carrying beneficial alleles to overcome host immunity should increase in frequency. However, evolutionary constraints may slow down the rise of such beneficial alleles (Dutta *et al.* 2020). In the absence of sexual recombination during the growing season, this should lead to the expansion of successful clonal genotypes. We found overall no signature of clonal expansion except for clones associated with cultivar CH Combin. The expansion of clones on CH Combin may indeed be the result of highly successful genotypes rising in frequency. Host damage (lesions) caused by these isolates is however lower compared to other isolates when tested on a different cultivar. Neutral processes such as population bottlenecks could well underlie the observed clonal expansion associated with the cultivar CH Combin. Consistent with selection on (asexual) pathogen reproduction over the course of the growing season is the significantly increased pycnidia production of isolates collected at C3 compared to C1. Matching changes in genotype frequencies with changes in frequencies of beneficial alleles for host colonization over the course of a growing season will help to disentangle neutral from selective processes.

The high degree of genotypic diversity originating from sexual reproduction is consistent with recent studies based on microsatellite markers showing spatial genetic structure but no meaningful loss in diversity over consecutive growing seasons (Siah *et al.*, 2018; Hassine *et al.*, 2019). Uniformity in host genotypes in agricultural ecosystems enables clonal pathogen populations to establish infections (Drenth *et al.*, 2019). However, the introduction of new host resistance genes can trigger the local extinction of such pathogen populations (Möller & Stukenbrock, 2017). Similarly, the application of fungicides can lead to local extinctions of clonal populations if populations lack resistance mutations. In contrast, in sexually reproducing populations recombination produces a multitude of new genotypes increasing the likelihood for adaptive traits to evolve (Burdon & Silk, 1997). Managing virulence and fungicide resistance emergence in pathogens has become a pressing issue to ensure sustainable agricultural production (Fisher *et al.*, 2012). Hence, understanding how pathogen populations respond in the short term (*i.e.* over the scale of weeks during the growing season) to challenges becomes a critical area of investigation. Detecting changes in pathogen populations *in situ* may become a powerful approach to identify the emergence of previously unknown adaptive mutations, changes in reproductive modes or the introduction of foreign genotypes. Deep population sequencing approaches revealing rapid changes in allele frequencies and population structure will become key tools in such endeavours.

Materials and methods

Field collection and storage

Z. tritici isolates were collected from the Field Phenotyping Platform (FIP) site of the ETH Zurich, Switzerland (coordinates 47.449°N, 198 8.682°E) (Kirchgessner *et al.*, 2017). A total of 335 European winter wheat varieties were grown in two replicates (Fig 1B) during the 2015–16 growing season separated each by approximately 100 m. During the collection season in 2016 the following fungicide applications were made: 4-8 April (300 g/l Spiroxamin, 160 g/l Prothioconazole), 25 May (Aviator Xpro with 75 g/l Bixafen + 150 g/l Prothioconazol), 6 June (Osiris with 56.25 g/l epoxiconazole and 41.25 g/l metconazole). Wheat was also subjected to regular applications of fertilizers and herbicides. A total of 177 isolates of *Z. tritici* were isolated from 12 winter wheat cultivars that are commonly grown in Switzerland (Levy et al., 2017) over three collection time points (fig 1A). To isolate strains, individual cirri from pycnidia identified on infected leaves were streak-plated on a yeast sucrose broth (YSB) solid media plate and incubated at 18° C for a week. Media were supplemented with 50 µg/ml kanamycin to prevent bacterial growth. After the incubation period, a single colony from the media plate was inoculated into 35 ml YSB liquid media and incubated on a shaking incubator at 18^{O}C for 8 days (140-180 rpm). After the incubation period, a dense culture of blastospores was obtained. The culture was washed twice with sterile distilled water. The spores were stored as both glycerol or silica stock for future use. For the preparation of glycerol stocks, equal amounts of washed blastospores and sterile glycerol were mixed in a cryotube and stored at -80 ^{O}C for future use. For the preparation of silica stocks, 200 µl of washed blastospores were poured onto 1 ml of sterile silica powder in a cryotube. Silica stocks were stored at 80 ^{O}C.

Culture preparation and seedling infection assay

Isolates were revived from glycerol stock by adding 50 µl stock solution to a 50 ml conical flask containing 35 ml liquid YSB medium. The inoculated flasks were incubated in the dark at 18° C and 140-180 rpm. After 8 days of incubation, the cultures were passed through four layers of meshed cheesecloth and washed twice with sterile water to remove media traces. The filtering step also largely eliminated hyphal biomass but retained spores. We analyzed pathogenicity profiles for a subset of 120 isolates out of the total number of isolates analyzed in this study. The pathogenicity assay dataset was generated in order to perform genome-wide association studies (Singh et al. 2021, bioRxiv *DOI pending*). In summary, seedlings of the cultivar Claro were grown under controlled conditions as follows: 16/8 hours day/night periods at 18°C throughout the experiment. The growth chamber was maintained at 70% humidity. Plants were grown for three weeks before infection with *Z. tritici*. To establish infections, washed spores were diluted to 2 x 10^5 spores/ml in 15 ml of sterile water containing 0.1% TWEEN20. Twenty-one days post inoculation (dpi), the second leaf of each plant was cut and fixed on a barcoded white paper. Leaves were scanned immediately using a flatbed scanner at 1200 dpi. The scanned images were batch-processed using automated image analysis (Karisto *et al.*, 2018). We recorded the percent leaf area covered by lesions and pycnidia counts.

Whole-genome sequencing and variant calling

Approximately 100 mg of lyophilized spores were used to extract high-quality genomic DNA using the Qiagen DNeasy Plant Mini Kit according to the manufacturer's protocol. We sequenced paired-end reads of 100 bp each with an insert size of ~550 bp on the Illumina HiSeq 4000 platform. Raw reads are available on NCBI Short Read Archive under the BioProject PRJNA596434 (Oggenfuss *et al.*, 2020). We performed sequencing quality checks using FastQC v. 0.11.9. (Andrews S., 2010. www.bioinformatics.babraham.ac.uk/projects/fastqc) and extracted read counts. Sequencing reads were then trimmed for adapter sequences and

sequencing quality using Trimmomatic v0.39 (Bolger et al., 2014) using the following settings: illuminaclip=TruSeq3-PE.fa:2:30:10, leading=10, trailing = 10, sliding-window = 5:10 and minlen = 50. Trimmed sequencing reads were aligned to the reference genome IPO323 (Goodwin et al., 2011) using Bowtie2 v. 2.4.1 (Langmead & Salzberg, 2012). Multi-sample joint variant calling was performed using the HaplotypeCaller and GenotypeGVCF tools of the GATK package version 4.0.1.2 (McKenna et al., 2010). We retained only SNP variants (excluding indels) and proceeded to hard filtered using the GATK VariantFiltration tool based on the following cutoffs: QD < 5.0; QUAL < 1000.0; MQ < 20.0; -2 > ReadPosRankSum > 2.0; -2 > MQRankSum > 2.0; -2 > BaseQRankSum > 2.0. After filtering for locus level genotyping rate (>80%) and minor allele count (MAC) of 1 using VCFtools v. 0.1.15 (Danecek et al., 2011). We analyzed genotyping rate at the isolate level using the --missing function of PLINK version 1.07 (Purcell et al., 2007). PLINK analyses were possible only for bi-allelic sites.

Population genetic analyses

We performed a down-sampling analysis to estimate the effect of sample size on the total number of segregating SNPs in the population. We randomly created subsets of 177 isolates from the population and counted the number of segregating SNPs in the subset. Presence-absence polymorphism of accessory chromosomes was analyzed based on chromosome-level coverage using BAMstats v. 1.25 (https://sourceforge.net/projects/bamstats/). We calculated the standardized coverage of all chromosomes by normalizing to the mean coverage of the core chromosomes. We categorized a chromosome as present in an isolate if the standardized coverage was within the range of 0.625-1.5. Otherwise, the chromosome was categorized as partially deleted (coverage of 0.25-0.625), absent (coverage <0.25) or duplicated (coverage >1.5). We estimated SNP density on each accessory chromosome by adjusting the "--max-

missing" filter for SNPs to 80% of the total number of observed chromosomes. We used the loess model implemented in the geom_smooth function from the R package ggplot2 (Wickham, 2016).

Linkage disequilibrium and population structure analyses

We analyzed the decay in linkage disequilibrium for each chromosome separately. For this, we used all SNPs with a minor allele frequency >5% in the population. We calculated the linkage disequilibrium r^2 between marker pairs using the option–hap-r2 in VCFtools v. 0.1.15 (Danecek *et al.*, 2011) with --ld-window-bp of 10000. The decay of linkage disequilibrium with physical distance was estimated using a nonlinear regression model (Remington *et al.*, 2001; Ingvarsson, 2005). We performed and visualised principal component analyses (PCA) using the R packages vcfR v. 1.8.0 (Knaus & Grünwald, 2017), adegenet v. 2.1.1 (Jombart & Ahmed, 2011), and ggplot2 v. 3.1.0 (Wickham, 2016). We generated an unrooted phylogenetic network using SplitsTree v4.14.6 (Huson, 1998). File format conversions were performed using PGDSpider v2.1.1.5 (Lischer & Excoffier, 2011). To identify groups of clonal genotypes, we calculated the pairwise genetic distances between all genotypes using the function dist.dna() included in the R package ape v. 5.3 (Paradis & Schliep, 2019). Isolate pairs with a pairwise genetic distance below 0.01 were considered as clones for further analyses. The Simpson diversity index was calculated using the R package vegan v. 2.5 (Oksanen *et al.*, 2019).

Spatial analyses and dispersal kernel estimation

The spatial coverage of isolate sampling and distances between the observed clonal isolate pairs were analysed. First, distances between all possible isolate pairs were calculated based on distance between plots: 2m along the columns and 1.5m along the rows (plot dimensions plus gaps between plots). The expected distribution of distances between isolates was

calculated assuming balanced sampling of isolates from each plot. The distribution of detected clonal pairs was compared to expected dispersal distances in successive generations of asexual propagation. The dispersal distance distribution was calculated assuming an exponential dispersal location kernel parametrized with dispersal kernel estimate from Karisto et al (2021, DOI pending). In brief, the dispersal kernel estimate were obtained by analyzing an experimentally set up infection focus in the same experimental wheat field at the beginning of the growing season. The infection was started from a previously genotyped isolate. Over the course of the growing season, infected leaves in immediate proximity (meter scale) of the infection focus were genotyped to determine the presence of the isolate introduced during the experiment. The shape parameter of the exponential dispersal location kernel was estimated to be α=13.5cm. The expected total dispersal distance distribution was calculated as convolutions of the dispersal distance kernel neglecting dispersal direction. The analyses were implemented in Python 3.7, using packages NumPy v1.17.3 (Harris *et al.*, 2020), SciPy v1.3.1 (Virtanen *et al.*, 2020), pandas v0.25.3 (Mckinney, 2010) and Matplotlib v.3.2.1 (Hunter, 2007).

Acknowledgement

Luzia Stalder and Ursula Oggenfuss provided helpful comments on a previous version of the manuscript. Emilie Chanclud provided advice on bioinformatics. The DNA library for the sequencing was prepared at the Genetic Diversity Centre (GDC) facility of ETH, Zurich. Sequencing was performed at the iGE3 platform of the University of Geneva.

References

Badet T, Fouché S, Hartmann FE, Croll D. 2020a. Machine-learning predicts genomic determinants of meiosis-driven structural variation in a eukaryotic pathogen. *bioRxiv*: 2020.10.23.352468.

Badet T, Oggenfuss U, Abraham L, McDonald BA, Croll D. 2020b. A 19-isolate reference-quality global pangenome for the fungal wheat pathogen Zymoseptoria tritici. *BMC Biology* 18: 12.

Barrett LG, Thrall PH, Burdon JJ, Linde CC. 2008. Life history determines genetic structure and evolutionary potential of host–parasite interactions. *Trends in Ecology & Evolution* 23: 678–685.

Barton NH. 2010. Genetic linkage and natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2559–2569.

Becheler R, Xhaard C, Klein EK, Hayden KJ, Frey P, De Mita S, Halkett F. 2016. Genetic signatures of a range expansion in natura: when clones play leapfrog. *Ecology and Evolution* 6: 6625–6632.

Billiard S, López-Villavicencio M, Hood ME, Giraud T. 2012. Sex, outcrossing and mating types: Unsolved questions in fungi and beyond. *Journal of Evolutionary Biology* 25: 1020–1038.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30: 2114–20.

Burdon JJ, Silk J. 1997. Sources and patterns of diversity in plant-pathogenic fungi. *Phytopathology* 87: 664–669.

Burdon JJ, Zhan J, Barrett LG, Papaix J, Thrall PH. 2016. Addressing the challenges of pathogen evolution on the world's arable crops. *Phytopathology* 106: 1117–1127.

Chen R-S, Mcdonald BA. 1996. Sexual Reproduction Plays a Major Role in the Genetic Structure of Populations of the Fungus *Mycosphaerella graminicola*. *Genetics* 142: 1119–1127.

Chisholm ST, Coaker G, Day B, Staskawicz BJ. 2006. Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell* 124: 803–814.

Croll D, Lendenmann MH, Stewart E, McDonald BA. 2015. The impact of recombination hotspots on genome evolution of a fungal plant pathogen. *Genetics* 201: 1213–1228.

Croll D, McDonald BA. 2012. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathogens* 8.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Daverdin G, Rouxel T, Gout L, Aubertot JN, Fudal I, Meyer M, Parlange F, Carpezat J, Balesdent MH. 2012. Genome Structure and Reproductive Behaviour Influence the Evolutionary Potential of a Fungal Phytopathogen. *PLoS Pathogens* 8: e1003020.

Dodds P, Thrall P. 2009. Recognition events and host-pathogen co-evolution in gene-for-gene resistance to flax rust. *Functional Plant Biology* 36: 395–408.

Drenth A, McTaggart AR, Wingfield BD. 2019. Fungal clones win the battle, but recombination wins the war. *IMA Fungus* 10: 18.

Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ. 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484: 186–194.

Fones H, Gurr S. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology* 79: 3–7.

Gandon S, Hochberg ME, Holt RD, Day T. 2013. What limits the evolutionary emergence of

pathogens? *Philosophical Transactions of the Royal Society B: Biological Sciences* 368: 20120086.
Goodwin SB, M'Barek S Ben, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, van der Lee TAJ, Grimwood J, Aerts A, *et al.* 2011. Finished genome of the fungal wheat pathogen Mycosphaerella graminicola reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genetics* 7: e1002070.
Hamelin FM, Castel M, Poggi S, Andrivon D, Mailleret L. 2011. Seasonality and the evolutionary divergence of plant parasites. *Ecology* 92: 2159–2166.
Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, *et al.* 2020. Array programming with NumPy. *Nature* 585: 357–362.
Hartmann FE, Croll D. 2017. Distinct trajectories of massive recent gene gains and losses in populations of a microbial eukaryotic pathogen. *Molecular Biology and Evolution* 34: 2808–2822.
Hartmann FE, McDonald BA, Croll D. 2018. Genome-wide evidence for divergent selection between populations of a major agricultural pathogen. *Molecular Ecology* 27: 2725–2741.
Hartmann FE, Sánchez-Vallet A, McDonald BA, Croll D. 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME Journal* 11: 1189–1204.
Hassine M, Siah A, Hellin P, Cadalen T, Halama P, Hilbert JL, Hamada W, Baraket M, Yahyaoui A, Legrève A, *et al.* 2019. Sexual reproduction of *Zymoseptoria tritici* on durum wheat in Tunisia revealed by presence of airborne inoculum, fruiting bodies and high levels of genetic diversity. *Fungal Biology* 123: 763–772.
He DC, Zhan JS, Xie LH. 2016. Problems, challenges and future of plant disease management: From an ecological point of view. *Journal of Integrative Agriculture* 15: 705–715.
Hunter JD. 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* 9: 90–95.
Huson DH. 1998. SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics* 14: 68–73.
Ingvarsson PK. 2005. Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., *salicaceae*). *Genetics* 169: 945–953.
Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.
Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* 444: 323–329.
Karisto P, Hund A, Yu K, Anderegg J, Walter A, Mascher F, McDonald BA, Mikaberidze A. 2018. Ranking quantitative resistance to septoria tritici blotch in elite wheat cultivars using automated image analysis. *Phytopathology* 108: 568–581.
Kema GH, Verstappen EC, Todorova M, Waalwijk C. 1996. Successful crosses and molecular tetrad and progeny analyses demonstrate heterothallism in Mycosphaerella graminicola. *Current genetics* 30: 251–8.
Kerdraon L, Laval V, Suffert F. 2019. Microbiomes and pathogen survival in crop residues, an ecotone between plant and soil. *Phytobiomes Journal* 3: 246–255.
Kirchgessner N, Liebisch F, Yu K, Pfeifer J, Friedli M, Hund A, Walter A. 2017. The ETH field phenotyping platform FIP: A cable-suspended multi-sensor system. *Functional Plant Biology* 44: 154–168.
Knaus BJ, Grünwald NJ. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources. 17: 44–53.

Krishnan P, Ma X, McDonald BA, Brunner PC. 2018. Widespread signatures of selection for secreted peptidases in a fungal plant pathogen. *BMC evolutionary biology* 18: 7.
Laine A-L, Barrès B. 2013. Epidemiological and evolutionary consequences of life-history trade-offs in pathogens. *Plant Pathology* 62: 96–105.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* 9: 357–9.
Levy L, Courvoisier N, Rechsteiner S, Herrera J, Brabant C, Hund A, Weissflog T, Dierauer H, Pellet D. 2017. Winterweizen: Bilanz aus 15 Jahren Sortenprüfung unter extensiven Anbaubedingungen. *Agrarforschung Schweiz* 8: 300–309.
Linde CC. 2010. Population genetic analyses of plant pathogens: New challenges and opportunities. *Australasian Plant Pathology* 39: 23–28.
Linde CC, Zhan J, McDonald BA. 2002. Population structure of *Mycosphaerella graminicola*: From lesions to continents. *Phytopathology* 92: 946–955.
Mariette N, Mabon R, Corbière R, Boulard F, Glais I, Marquer B, Pasco C, Montarry J, Andrivon D. 2016. Phenotypic and genotypic changes in French populations of *Phytophthora infestans*: Are invasive clones the most aggressive? *Plant Pathology* 65: 577–586.
McDonald BA. 1997. The Population Genetics of Fungi: Tools and Techniques. *Phytopathology* 87: 448–453.
McDonald BA, Linde C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* 40: 349–379.
McDonald MC, Renkin M, Spackman M, Orchard B, Croll D, Solomon PS, Milgate A. 2019. Rapid Parallel Evolution of Azole Fungicide Resistance in Australian Populations of the Wheat Pathogen *Zymoseptoria tritici*. *Applied and Environmental Microbiology* 85.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al.* 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20: 1297–303.
Mckinney W. 2010. *Data Structures for Statistical Computing in Python*. DOI: https://doi.org/10.25080/Majora-92bf1922-00a
Meier U. 2001. Growth stages of mono-and dicotyledonous plants BBCH Monograph Edited by Uwe Meier Federal Biological Research Centre for Agriculture and Forestry.
Morais D, Duplaix C, Sache I, Laval V, Suffert F, Walker AS. 2019. Overall stability in the genetic structure of a *Zymoseptoria tritici* population from epidemic to interepidemic stages at a small spatial scale. *European Journal of Plant Pathology* 154: 423–436.
Nomura K, DebRoy S, Lee YH, Pumplin N, Jones J, He SY. 2006. A bacterial virulence protein suppresses host innate immunity to cause plant disease. *Science* 313: 220–223.
Oggenfuss U, Badet T, Wicker T, Hartmann FE, Singh NK, Abraham N, Karisto P, Vonlanthen T, Mundt CC, Mcdonald BA, *et al.* 2020. A population-level invasion by transposable. *bioRxiv*. DOI: https://doi.org/10.1101/2020.02.11.944652
Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, O'hara RB, Simpson GL, Solymos P, *et al.* 2019. *Package 'vegan' Title Community Ecology Package Version 2.5-6*.
Paradis E, Schliep K. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.
Parratt SR, Numminen E, Laine A-L. 2016. Infectious Disease Dynamics in Heterogeneous Landscapes. *Annu. Rev. Ecol. Evol. Syst* 47: 283–306.
Peay KG, Bruns TD. 2014. Spore dispersal of basidiomycete fungi at the landscape scale is driven by stochastic and deterministic processes and generates variability in plant-fungal interactions. *New Phytologist* 204: 180–191.

Persoons A, Hayden KJ, Fabre B, Frey P, De Mita S, Tellier A, Halkett F. 2017. The escalatory Red Queen: Population extinction and replacement following arms race dynamics in poplar rust. *Molecular Ecology* 26: 1902–1918.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ, *et al.* 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559–575.

Read AF, Taylor LH. 2001. The ecology of genetically diverse infections. *Science* 292: 1099–1102.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler IV ES. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the United States of America* 98: 11479–11484.

Schotanus K, Soyer JL, Connolly LR, Grandaubert J, Happel P, Smith KM, Freitag M, Stukenbrock EH. 2015. Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and accessory chromosomes. *Epigenetics & Chromatin* 8: 41.

Siah A, Bomble M, Tisserant B, Cadalen T, Holvoet M, Hilbert JL, Halama P, Reignault P. 2018. Genetic structure of zymoseptoria tritici in northern France at region, field, plant, and leaf layer scales. *Phytopathology* 108: 1114–1123.

Slatkin M. 2008. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature reviews. Genetics* 9: 477–85.

Steinberg G. 2015. Cell biology of Zymoseptoria tritici: Pathogen cell organization and wheat infection. *Fungal genetics and biology* 79: 17–23.

Stukenbrock EH, Jørgensen FG, Zala M, Hansen TT, McDonald BA, Schierup MH. 2010. Whole-Genome and Chromosome Evolution Associated with Host Adaptation and Speciation of the Wheat Pathogen Mycosphaerella graminicola. *PLoS Genetics* 6: e1001189.

Suffert F, Delestre G, Gélisse S. 2019. Sexual Reproduction in the Fungal Foliar Pathogen *Zymoseptoria tritici* Is Driven by Antagonistic Density Dependence Mechanisms. *Microbial Ecology* 77: 110–123.

Suffert F, Goyeau H, Sache I, Carpentier F, Gélisse S, Morais D, Delestre G. 2018. Epidemiological trade-off between intra- and interannual scales in the evolution of aggressiveness in a local plant pathogen population. *Evolutionary Applications* 11: 768–780.

Suffert F, Ravigné V, Sachec I. 2015. Seasonal changes drive short-term selection for fitness traits in the wheat pathogen Zymoseptoria tritici. *Applied and Environmental Microbiology* 81: 6367–6379.

Sun W, Dunning FM, Pfund C, Weingarten R, Bent AF. 2006. Within-species flagellin polymorphism in *Xanthomonas campestris* pv *campestris* and its impact on elicitation of *Arabidopsis* FLAGELLIN SENSING2-dependent defenses. *Plant Cell* 18: 764–779.

Sundin GW, Castiblanco LF, Yuan X, Zeng Q, Yang C-H. 2016. Bacterial disease management: challenges, experience, innovation and future prospects: Challenges in Bacterial Molecular Plant Pathology. *Molecular plant pathology* 17: 1506–1518.

Susi H, Burdon JJ, Thrall PH, Nemri A, Barrett LG. 2020. Genetic analysis reveals long-standing population differentiation and high diversity in the rust pathogen melampsora lini. *PLoS Pathogens* 16.

Tack AJM, Thrall PH, Barrett LG, Burdon JJ, Laine AL. 2012. Variation in infectivity and aggressiveness in space and time in wild host-pathogen systems: Causes and consequences. *Journal of Evolutionary Biology* 25: 1918–1936.

Thrall PH, Burdon JJ. 2003. Evolution of Virulence in a Plant Host-Pathogen Metapopulation. *Science* 299: 1735–1737.

De Torres M, Mansfield JW, Grabov N, Brown IR, Ammouneh H, Tsiamis G, Forsyth A, Robatzek S, Grant M, Boch J. 2006. *Pseudomonas syringae* effector AvrPtoB suppresses basal defence in Arabidopsis. *Plant Journal* 47: 368–382.

Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R. 2006. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439: 344–8.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, *et al.* 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 17: 261–272.

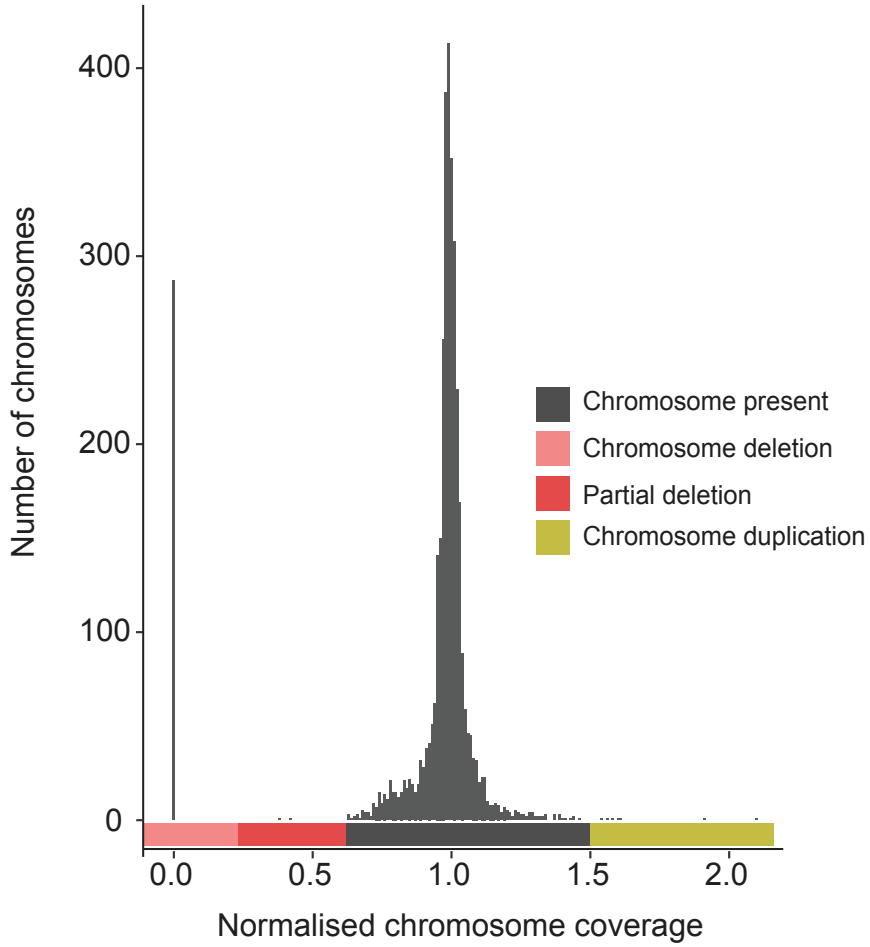
Walther BA, Ewald PW. 2004. Pathogen survival in the external environment and the evolution of virulence. *Biological Reviews of the Cambridge Philosophical Society* 79: 849–869.

Wickham H. 2016. Ggplot2 : elegant graphics for data analysis. *Springer-Verlag New York*.

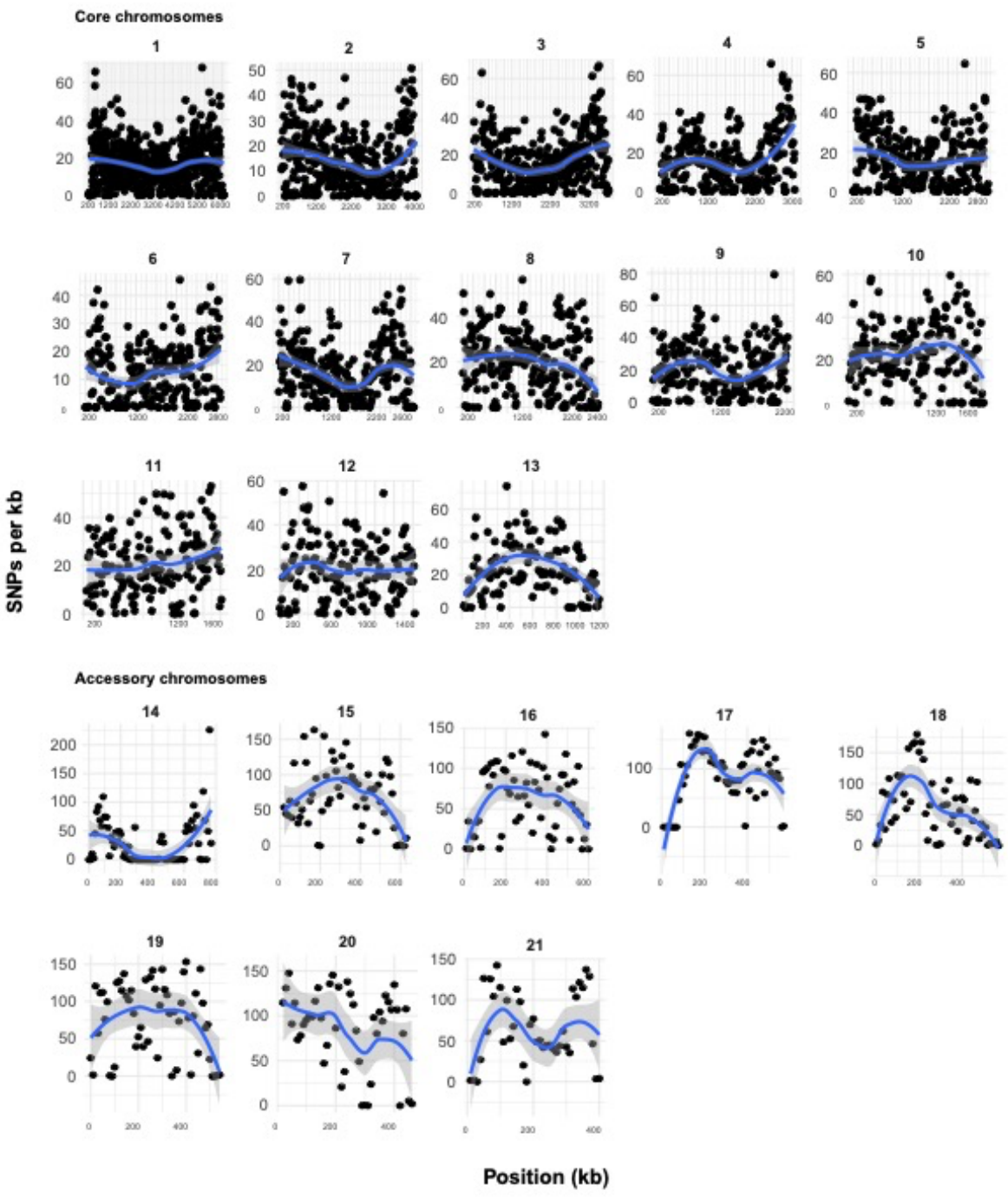
Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, Zala M, Ducasse A, Confais J, Compain J, Lapalu N, *et al.* 2017. A small secreted protein in *Zymoseptoria tritici* is responsible for avirulence on wheat cultivars carrying the *Stb6* resistance gene. *New Phytologist* 214: 619–631.

Supplementary information

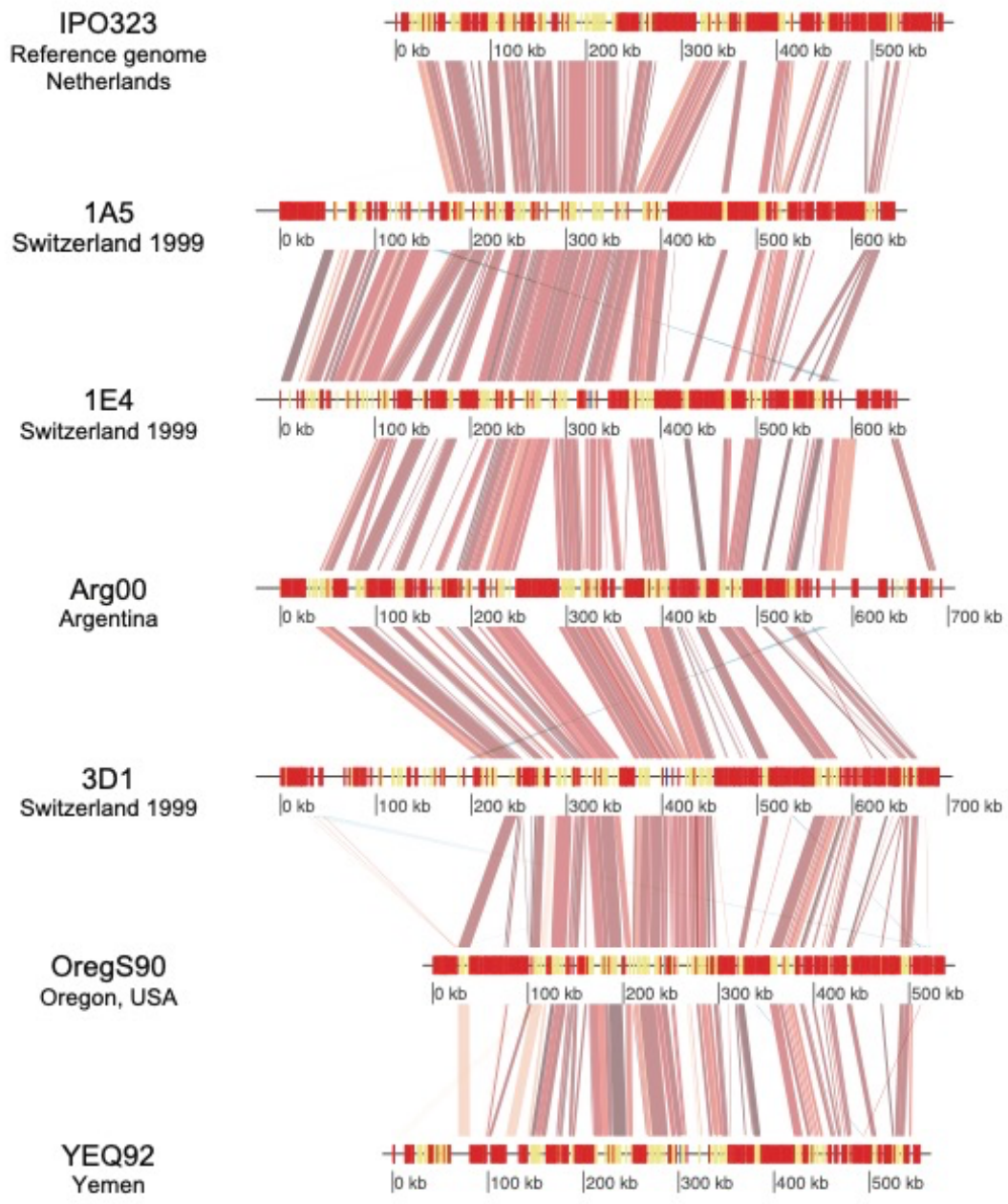
Supplementary Figures



Supplementary Figure S1: Normalized chromosome coverage of core and accessory chromosomes. The colored bars indicate normalized coverage cutoffs for defining the presence, absence, partial deletion and duplication state of individual chromosomes.



Supplementary Figure S2: Distribution of SNP density per kb across each chromosome. Non-overlapping sliding windows of 10 kb were analyzed for the total number of detected SNPs. The blue line represents a loess smooth function with standard error (grey shade).



Supplementary Figure S3: Chromosomal synteny plot of full-length chromosome 18 assembly. Based analyses by Badet et al. (2020a), seven isolates of *Zymoseptoria tritici* have fully assembled genomes and carry the accessory chromosome 18. The yellow and red segments along chromosomes represent genes and transposable elements, respectively. Red segments of variable intensity connecting chromosomes represent homologous segments identified by BLASTN (minimum identity of 90%). Blue segments represent inverted segments. See Badet et al. (2020a) for further details.

Supplementary Tables

Separate excel file available at https://doi.org/10.1101/2020.07.07.191510

Supplementary Table S1: *Zymoseptoria tritici* field collection information including collection time point (C1-3), cultivar of origin, field plot and assigned genotype.

Supplementary Table S2: Analyses of clonal groups identified in the field population. "Isolate 1" and "isolate 2" refer to pairwise comparisons of genotypes**.** Clone groups identify groups of isolates with identical genotypes.

Supplementary Table S3: Clonal diversity analyses (Simpson index) among collection time points (C1-3), cultivars and field plot.

Supplementary Table S4: Field plots number sown with specific wheat cultivars. Numbers refer to the graphic shown on Figure 1.

Supplementary Table S5: Experimental analyses of virulence on cultivar Claro. PLACL: Percent leaf area covered by lesions.

Chapter 2

Rapid sequence evolution driven by transposable elements at a virulence locus in a fungal wheat pathogen

Nikhil Kumar Singh[1], Thomas Badet[1], Leen Abraham[1], Daniel Croll[1]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland

Published in BMC Genomics: 22:393

https://doi.org/10.1186/s12864-021-07691-2

Keywords: Fungal pathogens, *Zymoseptoria tritici*, whole-genome sequencing, wheat, genome-wide association mapping, transposable elements

CHAPTER 2

Abstract

Background: Plant pathogens cause substantial crop losses in agriculture production and threaten food security. Plants evolved the ability to recognize virulence factors and pathogens have repeatedly escaped recognition due rapid evolutionary change at pathogen virulence loci (*i.e.* effector genes). The presence of transposable elements (TEs) in close physical proximity of effector genes can have important consequences for gene regulation and sequence evolution. Species-wide investigations of effector gene loci remain rare hindering our ability to predict pathogen evolvability.

Results: Here, we performed genome-wide association studies (GWAS) on a highly polymorphic mapping population of 120 isolates of *Zymoseptoria tritici*, the most damaging pathogen of wheat in Europe. We identified a major locus underlying significant variation in reproductive success of the pathogen and damage caused on the wheat cultivar Claro. The most strongly associated locus is intergenic and flanked by genes encoding a predicted effector and a serine type protease, respectively. The center of the locus contained a highly dynamic region consisting of multiple families of TEs. Based on a large global collection of assembled genomes, we show that the virulence locus has undergone substantial recent sequence evolution. Large insertion and deletion events generated length variation between the flanking genes by a factor of seven (5-35 kb). The locus showed also strong signatures of genomic defenses against TEs (*i.e.* RIP) contributing to the rapid diversification of the locus.

Conclusions: In conjunction, our work highlights the power of combining GWAS and population-scale genome analyses to investigate major effect loci in pathogens.

Keywords: Pathogen evolution, crops, genome-wide association mapping, transposable elements, genome assembly, population genomics

Background

Plant pathogens are a major threat to food security and cause annual losses of 20-30% of global harvest due to the lack of durable control strategies (Strange & Scott, 2005; Chakraborty & Newton, 2011; Savary, 2020). The emergence of new pathogens, the rise of new virulence in resident pathogens or the gain in resistance against chemical control agents create significant challenges (Strange & Scott, 2005; Subbarao *et al.*, 2015; McCann, 2020). To design effective disease control strategies, understanding the molecular interaction between plants and pathogens is critical. The virulence of plant pathogens is largely determined by their repertoire of secreted proteins known as effectors (Rovenich *et al.*, 2014; Depotter & Doehlemann, 2020). Effectors target a variety of different plant proteins and metabolic pathways to manipulate the immune response and physiological state of the host (Selin *et al.*, 2016). Plants evolved a large array of receptors often organized in networks that can directly or indirectly recognize the presence of effectors (Wu *et al.*, 2017; van der Burgh & Joosten, 2019; Depotter & Doehlemann, 2020). Detection of effectors triggers a variety of defense responses preventing the spread of pathogens across plant tissues. The identification of resistance genes encoding receptors has provided key tools for the rapid breeding of resistant crop varieties (Vleeshouwers & Oliver, 2014; Lo Presti *et al.*, 2015a). The identification of effectors in plant pathogens is challenging due to the large number of genes encoding effector-like proteins. The repertoire sizes of such candidate effectors varies between filamentous pathogens (Białas *et al.*, 2018; Depotter & Doehlemann, 2020). The potato light blight pathogen *Phytophthora infestans* has 1249 predicted effector candidates, whereas the white rust pathogen of *Arabidopsis thaliana, Albugo laibachii*, has only 143 predicted effector candidates (Mcgowan & Fitzpatrick, 2017). The frequent birth and death of genes encoding effectors is underpinning at least part of the variation in candidate effector repertoires among species and underlies also variation within the same species (Fouché *et al.*, 2018). Identifying functional effectors

providing an advantage for a pathogen on a specific host remains challenging (Selin *et al.*, 2016).

Effector gene polymorphism can be a major factor driving host-pathogen interactions (Lo Presti *et al.*, 2015b; Fouché *et al.*, 2018). The analyses of complete fungal genomes in combination with mapping analyses significantly expanded our knowledge of effectors across major filamentous pathogens. Genome-wide association study (GWAS) and analyses of progeny populations revealed three effectors of the fungal wheat pathogen *Zymoseptoria tritici* (Gohari *et al.*, 2015; Zhong *et al.*, 2017; Hartmann *et al.*, 2017; Stewart *et al.*, 2018; Meile *et al.*, 2018). The analyses of multiple completely assembled genomes revealed effector genes missing among individual isolates of the species (Plissonneau *et al.*, 2016, 2018; Badet *et al.*, 2020). Hence, pangenome analyses are crucial to establish the full extent of effector candidates within species (Badet & Croll, 2020). Such effector polymorphism is thought to be at the origin of rapid gains in virulence (Dong *et al.*, 2009, 2015; Fouché *et al.*, 2018; Asai *et al.*, 2018). Breakdown in host resistance can be observed within few years following the deployment of a crop cultivar (Cowger *et al.*, 2000; Islam *et al.*, 2016; Longya *et al.*, 2019; Cowger & Brown, 2019). Effector gene evolution can be driven by the complete deletion of coding sequence, as well as the accumulation of point and frameshift mutations (Rouxel & Balesdent, 2017; Hartmann *et al.*, 2017; Fouché *et al.*, 2018; Frantzeskakis *et al.*, 2020).

The rapid evolution of effector gene sequences is often driven by features of the chromosomal sequence in which the effector genes are embedded. Effector genes can be located on lineage-specific accessory chromosomes (Ma *et al.*, 2010; Croll & McDonald, 2012; Manning *et al.*, 2013). Such accessory chromosomes are enriched in repetitive sequences (Ma *et al.*, 2010). Effector genes located on core chromosomes are often located in the most repetitive regions of

the chromosome (Wang *et al.*, 2017; Torres *et al.*, 2020). The proximity to repetitive regions, in particular transposable elements (TEs), increases the likelihood for sequence rearrangements to occur. The localization of effectors in highly repetitive sub-telomeric regions contributed to rapid virulence evolution of the rice pathogen *Magnaporthe oryzae* (Xue *et al.*, 2012; Yoshida *et al.*, 2016). The *AVR-Pita* effector gene has been shown to undergo multiple translocations in the genome contributing to the evolution of virulence on specific hosts (Chuma *et al.*, 2011). The insertion of a Mg-SINE TE in the effector gene *AvrPi9* led to a loss-of-function mutation enabling *M. oryzae* to escape host resistance (Wu *et al.*, 2015). The transposition of TEs can disrupt coding sequences or change the regulation of effector genes (Sánchez-Vallet *et al.*, 2018; Meile *et al.*, 2018; Fouché *et al.*, 2020). Additionally, repetitive sequences can lead to higher mutation rates through a mechanism known as repeat induced point (RIP) mutation (Gladyshev, 2017; Gardiner *et al.*, 2020; Wang *et al.*, 2020). *Brassica napus* (canola) carrying the *Rlm1* resistance gene suffered a breakdown of resistance against the fungal pathogen *Leptosphaeria maculans* (Van de Wouw *et al.*, 2010). The breakdown was associated with a rise in virulence alleles at the *AvrLm1* locus (Van de Wouw *et al.*, 2010). Sequence analyses revealed that the gain in virulence was driven by RIP mutations rendering the locus non-functional. Highly similar sequences nearby effector genes can also trigger ectopic recombination and, by this, the deletion or duplication of the effector gene. Consequently, the genomic context of effector genes provides critical information about effector evolvability. Hence, within-species analyses of effector gene diversification and TE dynamics of the surrounding regions have become key tools to retrace the evolution of virulence.

The haploid ascomycete *Zymoseptoria tritici* is one of the most destructive pathogens of wheat leading to yield losses of ~ 5–30% depending on climatic conditions (Jørgensen *et al.*, 2014; Fones & Gurr, 2015). Pathogen populations across the wheat-producing areas of the world

harbor significant variation in pathogenicity and genetic diversity (Zhong *et al.*, 2017; Hartmann *et al.*, 2017; Hartmann & Croll, 2017b; Krishnan *et al.*, 2018a; Singh *et al.*, 2020). GWAS were successfully used to identify the genetic basis of virulence on two distinct wheat cultivars (Zhong *et al.*, 2017; Hartmann *et al.*, 2017). In addition, analyses of progeny populations revealed a third effector gene related to a resistance breakdown (Stewart *et al.*, 2018; Meile *et al.*, 2018). GWAS was also successfully used to map the genetic architecture of a broad range of phenotypic traits related to abiotic stress tolerance (Dutta *et al.*, 2020b). TE dynamics are playing a key role in influencing the sequence dynamics at effector gene loci (Hartmann *et al.*, 2017; Meile *et al.*, 2018; Fouché *et al.*, 2020). Gene gain and loss dynamics are accelerated in proximity to TEs (Hartmann & Croll, 2017b). TEs shape also the epigenetic landscape in proximity to effectors (Fouché *et al.*, 2020; Meile *et al.*, 2020). Phenotypic traits expressed across the life cycle of the pathogen show extensive trade-offs possibly constraining the evolution of virulence (Dutta *et al.*, 2020b,a). Identifying additional loci underlying pathogenicity on specific hosts remains a priority as for most wheat resistance (*i.e. Stb*) genes, the matching effector remains unknown (Brown *et al.*, 2015).

In this study, we aimed to identify the genetic basis of virulence on the wheat cultivar Claro using GWAS performed on a genetically highly diverse mapping population established from a single wheat field. We analyzed the expression patterns of genes in proximity to the top associated SNP, the presence of TEs and genetic variation at the locus in populations across the world to build a comprehensive picture of sequence dynamics at the newly identified virulence locus.

Results

Genome sequencing of a highly polymorphic pathogen field population

In order to build a mapping population for GWAS, we obtained total of 120 isolates of *Z. tritici* collected from a multi-year experimental wheat field in Switzerland planted with 335 wheat cultivars (Karisto *et al.*, 2017; Singh *et al.*, 2020) (Supplementary Table S1). The isolates were collected from a subset of 10 genetically different winter wheats (7-20 isolates per cultivar) from two different time points during a single growing season. We analyzed whole-genome sequencing datasets of each isolate constituting an average coverage of 21X as previously described (Singh *et al.*, 2020). After quality filtering, we obtained 788'313 high-confidence SNPs. We constructed an unrooted phylogenetic network using SplitsTree to visualize the genotypic differentiation within the population (Fig 1A). Compared to the broader field population analyzed previously, our GWAS mapping population contained 10 clonal groups comprising a total of 21 isolates (Singh *et al.*, 2020) (Supplementary Table S2). A principal component analysis confirmed the overall genetic differentiation within the population (Supplementary Fig 1). Nearly all genotypes were at similar genetic distances to each other with the exception of five genotypes with significantly larger genetic distances to the main cluster of genotypes (Singh *et al.*, 2020) (Supplementary Fig 1). Interestingly, the five isolates were all collected from cultivar CH Combin, which is susceptible to *Z. tritici* (Courvoisier *et al.*, 2016).

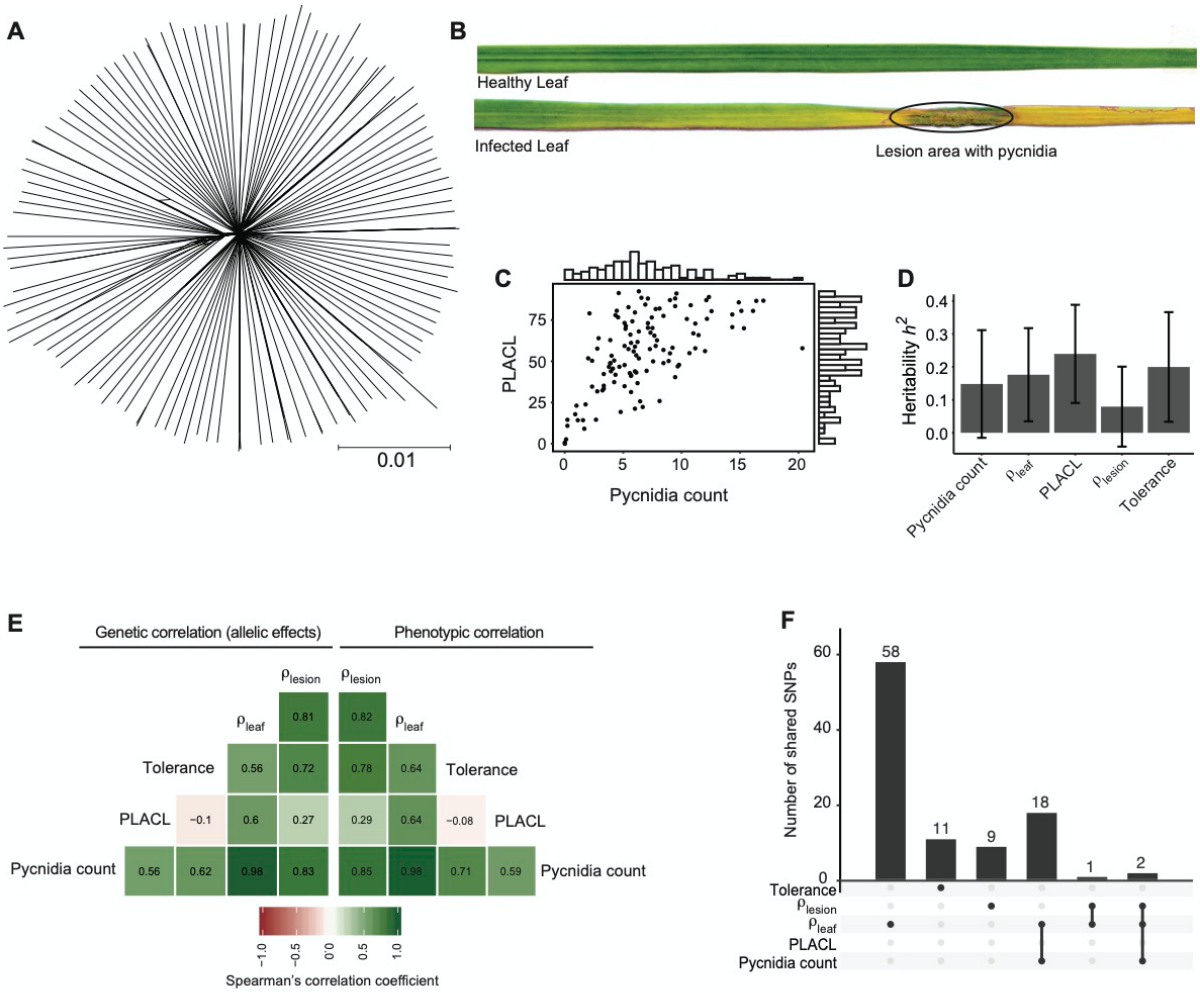


Figure 1 : Genetic and phenotypic diversity in a single field population of *Zymoseptoria tritici.* A) Phylogenetic network of 120 isolates constructed using SplitsTree B) Photographs showing the difference between a mock treated and infected leaf. C) Trait distribution of pycnidia counts in lesions and the percentage of leaf area covered by lesion (PLACL). D) SNP based heritability (h^2 SNP) of the virulence phenotypes estimated following a GREML approach. Error bars indicate standard errors. E) Mean allelic effect (*i.e.* genetic) correlation and phenotypic correlation coefficients for all measured virulence phenotypes. F) Number of significantly associated SNPs (5% FDR threshold) exclusive to an individual virulence trait or shared among traits.

Heritability and correlations among pathogenicity traits

We experimentally assessed the expression of pathogenicity traits of each individual isolate on the winter cultivar Claro using a greenhouse assay. The cultivar Claro was among the cultivars used in the multi-year experimental wheat field from which the isolates were sampled from (Karisto *et al.*, 2017). The cultivar is widely planted in Switzerland and is generally mildly

susceptible to *Z. tritici* (Courvoisier *et al.*, 2016). We obtained quantitative data on symptom development from a total of 1'800 inoculated leaves using automated image analysis (Karisto *et al.*, 2017). The image analyses pipeline was previously optimized to detect symptoms caused by *Z. tritici* under greenhouse conditions and uses a series of contrast analyses to obtain estimates of the surface covered by symptoms. For each leaf, we recorded the counts of pycnidia (structures containing asexual spores) and the percentage of leaf area covered by lesion (PLACL) (Fig 1B). We considered the pycnidia count as a proxy for reproductive success of the pathogen on the host and PLACL as an indication of host damage due to pathogen infection. From these measurements, we derived three quantitative resistance measures: ρ_{leaf} is the pycnidia count per cm^2 of leaf area, ρ_{lesion} is defined as the total number of pycnidia divided by per cm^2 lesion area, and tolerance is expressed as the pycnidia count divided by PLACL. ρ_{leaf} represents the overall reproductive success per area while ρ_{lesion} focuses on the reproductive success within the lesion area. Tolerance indicates the ability of the host to tolerate pathogen reproduction while limiting damage by lesions (Mikaberidze & McDonald, 2020). We found that the mean pycnidia count ranged from 0-20 (mean 7, median 6.3) among isolates and PLACP ranged from 2-97 % (mean 56%, median 57.7%) (Fig 1C, Supplementary table S1, Supplementary Fig 2). ρ_{leaf} values ranged from 0.04-7.2 (mean 2.4, median 2.15); ρ_{lesion} ranged from 0-13.8 (mean 3.6, median 3.3) and tolerance ranged from 0.15-0.3 (mean 0.12, median 0.17) (Supplementary table S1, Supplementary fig 2).

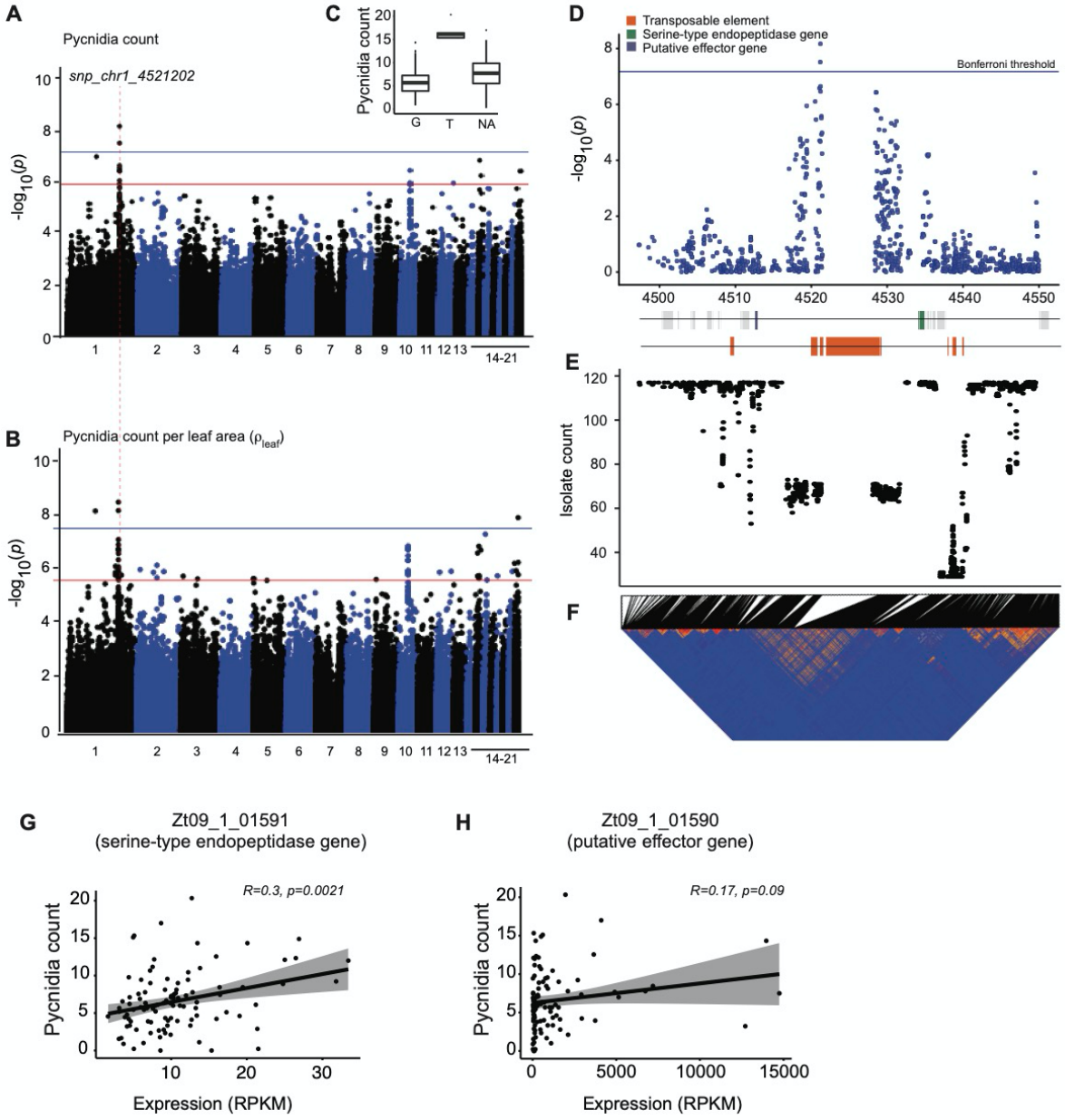


Figure 2: Genome-wide association mapping for virulence. Manhattan plots showing SNP marker association *p-values* for A) pycnidia count and B) ρ_{leaf} (pycnidia count per cm^2 of leaf area). The blue and red lines indicate the significance thresholds for Bonferroni ($\alpha = 0.05$) and false discovery rate (FDR) at 5%, respectively. The dotted line represents the most significant association on chromosome 1 (*snp_chr1_4521202*). C) Boxplot showing the pycnidia counts of isolates carrying the reference allele G or alternative allele T at the top significant SNP. D) Zommed in Manhattan plot for association *p-values* of SNPs in a ~25 kb region centered on the top SNP *snp_chr1_452120*. Horizontal lines represent the Bonferroni threshold ($\alpha = 0.05$). E) Genotyping rates of SNPs in the mapping population. F) Linkage disequilibrium r^2 heatmap. G-H) Correlation plot of pycnidia count with gene expression of the flanking effector candidate gene (*Zt09_1_01590*) and the serine-type endopeptidase gene (*Zt09_1_01591*).

We estimated SNP-based heritability (h^2_{snp}) for each trait using a genomic-relatedness-based restricted maximum-likelihood approach to partition the observed phenotypic variation (Fig 1D). The h^2_{snp} ranged from 0.08 – 0.23 among different phenotypes (Fig 1D). Heritability for pycnidia counts and PLACL was 0.17 (SE = 0.14) and 0.15 (SE = 0.16), respectively. We found the highest h^2_{snp} for ρ_{leaf} (0.24, SE=0.15) exceeding h^2_{snp} for ρ_{lesion} (0.19, SE=0.16). Pathogenicity-related traits have overlapping genetic architectures leading to phenotypic and genetic correlations (Dutta *et al.*, 2020b). To identify potential trade-offs among traits, we analyzed correlations among all pairs of traits (Fig 1E). We found overall positive phenotypic trait correlations except for PLACL and tolerance (r_p = -0.08; Fig 1E). To assess genetic correlations among traits, we performed GWAS on each trait. To avoid *p*-value inflation due to non-random degrees of relatedness among genotypes, we used a mixed linear model that included a kinship matrix. We assessed the allelic effects across all SNPs for all traits to estimate the degree of genetic correlation among trait pairs. We found the genetic correlations (r_p) to vary from -0.1 to 0.98 (Fig 1E). Pycnidia counts and ρ_{leaf} showed the highest degree of genetic correlation. PLACL and tolerance showed the lowest degree of genetic correlation. Overall, phenotypic and genetic correlations among pairs of traits were highly similar.

Major effect locus for pathogen reproduction on the cultivar Claro

We used the GWAS on each trait to identify the most significantly associated SNPs in the genome. We focused on association *p*-values passing the 5% false discovery rate threshold for all the phenotypes except for PLACL where we found no significant associations (Supplementary Fig 3). All significantly associated SNPs for pycnidia count were overlapping with significantly associated SNPs for ρ_{leaf} and ρ_{lesion} (Fig 1F). The traits $\rho_{leaf,}$ ρ_{lesion} and tolerance had 58, 9 and 11 associated SNPs, respectively, which were uniquely associated with the specific trait and not overlapping with any other trait (Fig 1F). We then focused our

investigation on the most significantly associated SNPs passing the Bonferroni threshold (α = 0.05). We found a single locus on chromosome 1 with significantly associated SNPs for pycnidia count, ρ_{leaf} and ρ_{lesion} (Fig 2A-B, Supplementary Fig 3C,E). The top SNP (chr1_4521202) showed an association of isolates carrying the non-reference allele T with higher pycnidia production compared isolates with reference allele G (Fig 2C). The non-reference allele was less frequent in the population (10%) and nearly half (48%) of all isolates were not assigned a SNP genotype at the locus.

We analyzed sequence characteristics of the chromosomal region surrounding the top locus. The SNP *chr1_4521202* is located in an intergenic region rich in TEs (Fig 2D-E). The closest identified genes include a gene encoding a putative effector (Zt09_1_01590) and a gene encoding a serine-type endopeptidase (Zt09_1_01591). The genes were at a distance of ~8 kb and ~4.5 kb, respectively, from the SNP *chr1_4521202* (Supplementary Table S3). The low genotyping rate at the SNP suggests that segmental deletions are present. The genotyping rate was 58%, which is consistent with the SNP genotyping rate for nearby SNPs (within ~5 kb; Fig 2E). We recovered no SNPs in the immediate vicinity (at around 4.25 Mb). The genotyping rate increases to close to 100% at a further distance of the top SNP (> 10 kb; Fig 2E. The segmental pattern in the reduced genotyping rate close to the most significant SNP suggests that a substantial fraction of the isolates harbor deletions. We analyzed patterns of linkage disequilibrium among pairs of SNPs (Fig 2F). We found that the decay in linkage disequilibrium generally occurred at short distance (~1-2 kb) with the exception of the region surrounding the top SNP *chr1_4521202* (Fig 2F). The increased linkage disequilibrium suggests that the physical distance among SNPs in the analyzed isolates is shorter consistent with the detection of deletions.

We analyzed transcription levels of the two closest genes using RNA-seq data generated under culture conditions simulating starvation (minimal medium) for all isolates of the GWAS panel. Both genes were conserved in all the isolates and appear transcriptionally active with variable expression levels among the isolates. The candidate effector gene was transcribed between 12-14'750 RPKM (Fig 2H, Supplementary Fig 4). The serine-type endopeptidase gene showed much lower transcription ranging from 1.6-33.4 RPKM (Fig 2G, Supplementary Fig 4). We found that transcription levels of the gene encoding the endopeptidase was positively correlated with the amount of pycnidia produced ($r = 0.3$, $p = 0.0021$). We found no significant correlation with the expression of the effector candidate gene (Fig 2H).

Transposable element dynamics and sequence rearrangements

Given the indications for segmental deletions at the virulence locus, we analyzed multiple completely assembled genomes of the species. We included genomes from isolates from Switzerland, United States, Australia and Israel covering the global distribution range of the pathogen (Badet *et al.*, 2020). The locus showed a highly variable content in TEs underlying significant length variation. The distance by the two flanking genes is 20.2 kb in the reference genome IPO323 used for mapping (Fig 3A-B). However, this distance varies from 4.8–35.3 kb between the genes depending on the genome for an average distance of ~17 kb (Fig 3B). The longest distance between genes was found in the genome of the Swiss strain CH99_1A5 and the shortest distance was found in the genome of the Israeli strain ISY92.

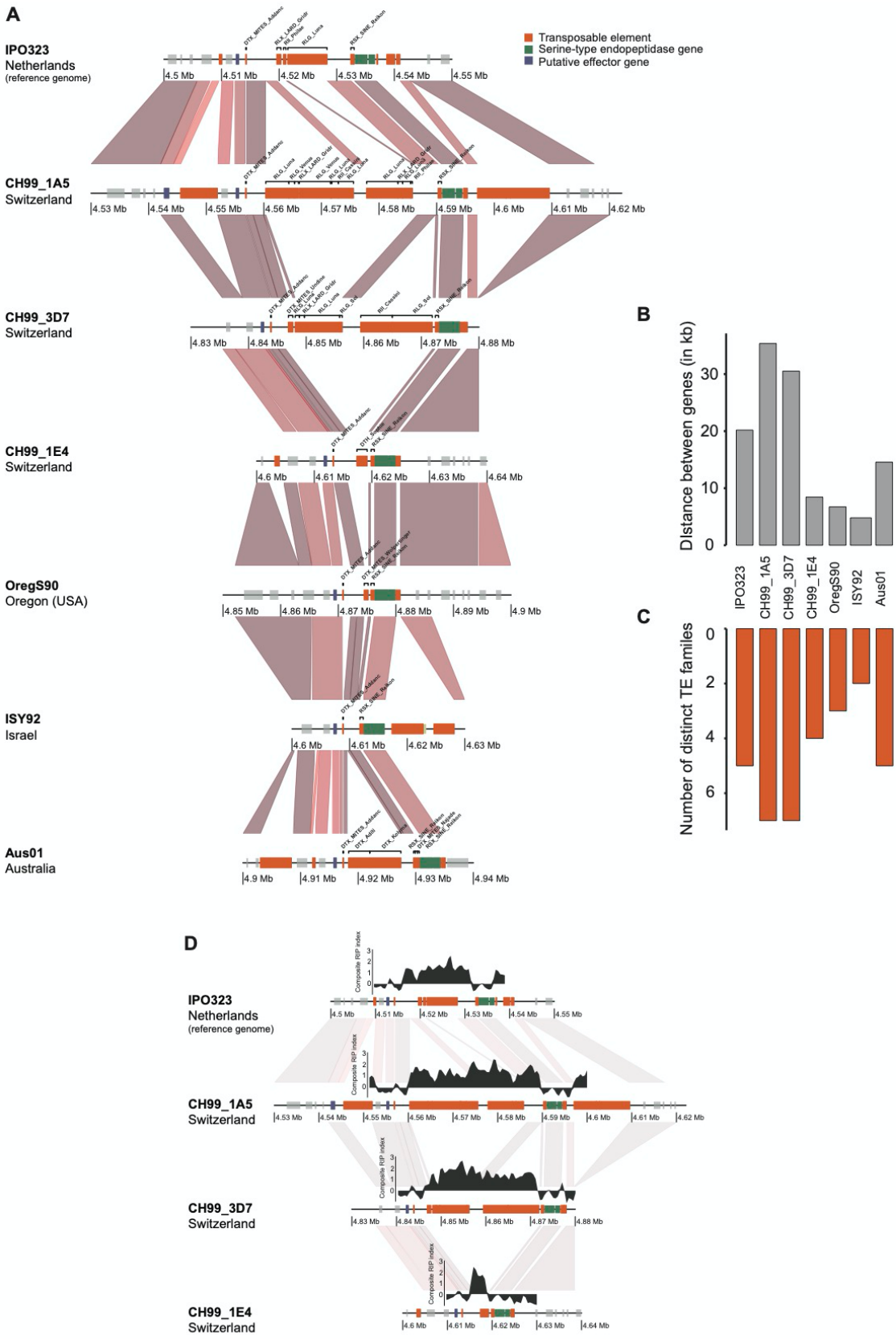


Figure 3: TE content variation at the virulence locus. A) Synteny plot of the top locus analyzed in seven completely assembled genomes. The red gradient segments represent the percentage of sequence identity from BLASTN alignments. Darker colors indicate higher

identity. B) Distance variation between the two genes surrounding the top locus (*Zt09_1_01590* and *Zt09_1_01591*). C) The number of different TE families found at least once per isolate at the top locus. D) Repeat induced point (RIP) mutation signatures in the topic locus. The Large RIP Affected Regions (LRARs) composite index was calculated using The RIPper tool (van Wyk et al., 2019).

We identified five different TE families in the reference genome IPO323 covering a segment of ~20 kb (Fig 3C). We detected additional TE families in two of the three genomes from Switzerland (CH99_1A5 and CH99_3D7). The genomes carry multiple copies of a total of seven different TE families. Meanwhile, the two genomes from Israel and the United States showed a reduction in TEs with the region carrying only single copies of two and three different TE families, respectively (Fig 3A-C). The presence of TEs in fungal genomes can trigger RIP mutations. We found consistent signatures of RIP between the two flanking genes but we found no indications for RIP leakage into the flanking genes (Fig 3D, Supplementary Fig 5).

Transposable element insertion dynamics across populations

The small set of completely assembled genomes provides only a partial view on the sequence rearrangement dynamics within the species. Hence, we generated draft genome assemblies for 432 isolates from previously analyzed field populations in the United States (n = 56 + 97), Switzerland (n = 37 + 185), Israel (n = 30) and Australia (n = 27; Supplementary table S4). The two population from United States and Switzerland were collected at an interval of 25 and 20 years, respectively, from the same field (Supplementary table S4). Illumina sequencing datasets for fungi with compact genomes produce reasonably accurate draft assemblies (Torriani *et al.*, 2011; Mohd-Assaad *et al.*, 2019; Stauber *et al.*, 2020). We used BLASTN (Altschul *et al.*, 1990) to locate the two genes *Zt09_1_01590* and *Zt09_1_01591* adjacent to the top SNP across all assemblies. We retained only draft assemblies for which both genes were located on the same scaffold. Hence, these scaffolds provide a contiguous view on the sequences located between the two adjacent genes. With this filtering step, we retained 122

isolates from all four different locations including the United States (n=49), Israel (n=17) and Switzerland (n =6 and 50) (Fig. 4A; Supplementary table S4). The distance between the two genes ranged from 5-35 kb, which is highly consistent with the gene distances observed in the completely assembled genomes (Fig 4B). The isolates from Israel and Switzerland (collection 2016) showed shorter distance ranging from 5-15 kb. The United States population and the older Switzerland population (collection 1999) showed a range of 6.8-35 kb between the genes (Fig 4B).

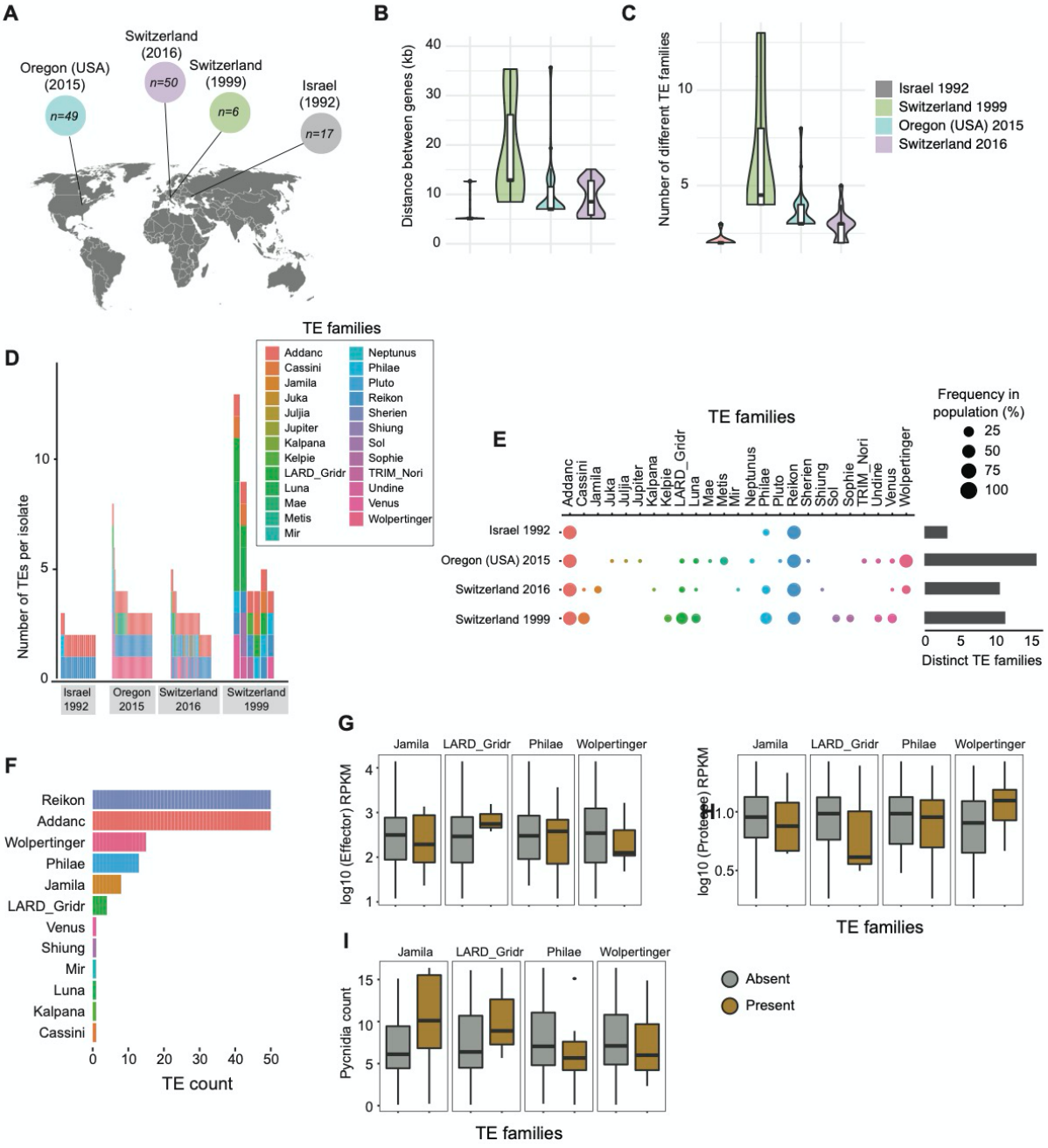


Figure 4: Analysis of transposable element dynamics across continents. A) Analysis of 122 *Zymoseptoria tritici* isolates for which a draft genome assembly produced a scaffold containing both genes *Zt09_1_01590* and *Zt09_1_01591*. B) Boxplot showing variation in the distance between the two genes per population. C) TE content variation of the sequence flanked by the two genes. D) Total TE copies in the sequence flanked by the two genes. E) Frequency of the TE families in the sequence flanked by the two genes as a percentage of the population. F) Frequency of TE families among the isolates from the GWAS mapping population ($n = 50$ with a scaffold spanning both genes) G-H) Boxplots showing the expression of the genes *Zt09_1_01590* and *Zt09_1_01591*, and pycnidia counts for isolates carrying or not specific TEs at the top locus.

We annotated the scaffolds matching the top GWAS locus using consensus sequences of known TE families. Overall, the TEs between the two adjacent genes grouped into 11 superfamilies and 25 families (Fig 4D). The two most frequent TEs included both retrotransposons and miniature-inverted repeat transposable elements. We found that genomes from the United States and the earlier Switzerland population (collection 1999) had higher TE copy numbers compared to other genomes from the other populations (2-13 TE copies; Fig 4C). The locus contains overall 16 different TE families in the United States population (Fig 4D-E). The locus contained 3, 11 and 16 different TE families in the Israel, and the Switzerland 1999 and 2016 populations, respectively (Fig 4D-E). The TEs RSX_SINE_Reikon and DTX_MITES_Addanc were found in all analyzed populations population while other TE families were segregated in populations in different proportions (Fig 4E). To test for potential associations of TE presence and pathogenicity traits, we focused on the complete scaffolds retrieved from 50 different isolates of the GWAS population. We found segregating presence-absence polymorphism for the four TE families RII_Philae and RLX_LARD_Gridr (retrotransposons), as well as DTX_MITES_Wolpertinger and DTC_Jamila (DNA transposons; Fig 4F). None of the TE presence-absence polymorphism showed a significant association with the transcription of adjacent genes (effector candidate *Zt09_1_01590*: Student's *t*-test, $p > 0.15$; serine-type endopeptidase *Zt09_1_01591*: Student's *t*-test, $p > 0.05$). We also found no significant association with the TE presence-absence polymorphism and any of the pathogenicity traits (Student's *t*-test, $p > 0.2$).

Discussion

We used whole-genome sequencing data and association mapping to unravel the genetic architecture of pathogenicity of *Z. tritici* on the wheat cultivar Claro. The identified locus is rich in TEs and is flanked by genes encoding an effector candidate and a serine type protease. We analyzed a worldwide set of populations to analyze sequence variation at the pathogenicity locus. We found significant length variation caused by the insertion of a diverse set of TEs.

Variation in pathogenicity on the wheat cultivar Claro was largely quantitative. We found that heritability was higher for pathogen virulence (damage to host) than pathogen reproduction (production of pycnidia). This is in contrast to analyses of heritability across 12 different wheat cultivars where heritability for pathogen reproduction was typically higher compared to lesion damage (Dutta *et al.*, 2020b). However, virulence and reproduction were overall positively correlated in both studies. We also found a high degree of phenotypic and genetic correlation with tolerance (*i.e.* preventing lesion damage despite high reproduction of the pathogen). Using GWAS, we identified several loci significantly associated with different pathogenicity traits. The most significant associations were found for pycnidia counts and ρ_{leaf} both related to reproductive success of the pathogen. Pathogen reproduction showed a strong single locus association while host damage (*i.e.* lesions) revealed no single gene effects. The difference between traits may be due to the fact that the genetics underlying host damage is more complex. Lesions are caused by host cell death triggered as a response to pathogen attack (Coll *et al.*, 2011; Dickman & de Figueiredo, 2013). Hence, variation in lesion development among isolates could be due to the host's ability to perceive specific molecules produced only by a subset of the isolates. Furthermore, variation in the pathogen's ability to spread across tissue and manipulate host immune responses could also lead to variation in overall lesion development. Interestingly, extensive lesion development is not necessarily related to pycnidia production

by the pathogen across cultivars (Karisto *et al.*, 2018; Mikaberidze & McDonald, 2020). This suggests that despite damage to the leaf, the host immune system can efficiently repress the pathogen from acquiring nutrients to reproduce. In contrast, the strong single locus association for pycnidia production on the cultivar Claro suggests a rather simple genetic architecture. Hence, the action of a single pathogen factor (*e.g.* an effector) may be largely sufficient to determine variation in host exploitation and reproduction.

We identified a highly polymorphic chromosomal locus associated with pathogenicity on the cultivar Claro. The most significant SNPs mapped in an intergenic region flanked by a large cluster of diverse TEs. We found no clear evidence for a coding sequence in immediate proximity of the most significantly associated SNPs. The closest genes encode functions, which may be relevant for host infection though. Serine-type endopeptidases play pivotal roles in nutrient degradation and subsequent assimilation, as well as protection from the host immune system (Muszewska *et al.*, 2017). Serine proteases can also help the pathogen to escape the host's immune system by degrading chitinases targeted at the fungal cell wall (Langner & Göhre, 2016). Furthermore, serine proteases play a role in the nutrient acquisition from plant tissue (Jashni *et al.*, 2015) and potentially during the initiation of necrosis (Palma-Guerrero *et al.*, 2016). The second gene encodes a putative effector, which is a category of genes showing frequent presence-absence polymorphism within the species (Hartmann & Croll, 2017a; Badet & Croll, 2020). Our analyses of linkage disequilibrium decay suggest that neither of the two adjacent genes play a causal role in pathogenicity on Claro. The most dramatic changes occurred due to the insertion and deletion of TEs next to the most significantly associated SNPs. The TE dynamics diversified the locus to the extent that the distance between the adjacent genes varies by a factor of seven (5-35 kb). The insertion and deletion of TEs can have both an impact on gene regulation by inducing epigenetic silencing or upregulation. Both

mechanisms are well established in *Z. tritici* and underlie variation in melanin production, virulence and fungicide resistance (Omrane *et al.*, 2017; Meile *et al.*, 2018; Krishnan *et al.*, 2018b). The locus flanked by the two genes showed strong signatures of RIP. The elevated mutation rates triggered by this genomic defense mechanism against TEs likely contributed to the rapid diversification of the locus. Yet, we could not establish any direct association between the insertion of individual TEs and the expression of pathogenicity. Targeted deletion assays focusing on individual sequence segments may provide experimental evidence for the sequence variation underlying pathogenicity on the wheat cultivar.

Conclusions

The effects of gene-TE proximity have been studied mainly in animal (Rebollo *et al.*, 2012; Cowley & Oakey, 2013) or plant models (Bennetzen & Wang, 2014). Only a handful studies are focused on fungi. Some fungal pathogens have genomes with a clearly compartmentalized architecture described by the two-speed model (Dong *et al.*, 2015). The core genome encodes all essential genes while niche- or host-specific genes (*e.g.* effectors) are typically encoded in the repeat-rich genome compartment. Such genome architectures have been identified in *Mycosphaerella fijiensis* (Santana *et al.*, 2012), *Cochliobolus heterostrophus* (Santana *et al.*, 2014), *Fusarium* species (Sperschneider *et al.*, 2015), *Leptosphaeria maculans* (Rouxel & Balesdent, 2017) and *Verticillium* species (Faino *et al.*, 2016). However, systematic investigation of TEs and co-localizing genes have rarely been extended to the within species level. Our study shows that a combination of genome-wide association mapping, complete and draft genome assemblies can provide a comprehensive insight into the evolutionary dynamics of virulence loci. Hence, even in absence of experimentally validated effectors, the evolutionary trajectory of virulence loci becomes tractable. Our approach should be broadly applicable to many fungal pathogen system.

Methods

Field collection and storage

Z. tritici isolates were collected from the Field Phenotyping Platform (FIP) site of the ETH Zürich, Switzerland (Eschikon, coordinates 47.449°N, 8.682°E). We analyzed a total of 120 isolates collected during the 2015/2016 growing season from 10 winter wheat cultivars, which are commonly grown in Switzerland (Levy et al., 2017). We analyzed isolates originating from two collection time points over the season (Table S1). Isolates from the first collection (n = 62) were collected when wheat plants were in Growth stage (GS) 41 while the second collection (n = 58) was performed when the plants were in GS 85 stage. After sampling, spores of each isolate were stored in either 50% glycerol or anhydrous silica gel at −80 °C. Additional information regarding sampling schemes and genetic diversity is available (Singh *et al.*, 2020)

Culture preparation and seedling infection assay

Isolates were revived from glycerol stock by adding 50 µl fungal stock solution to a 50 ml conical flask containing 35 ml liquid YSB (yeast-sucrose broth) medium. The inoculated flasks were incubated in the dark at 18° C and 140-180 rpm on a shaker-incubator. After 8 days of incubation, the cultures were passed through four layers of meshed cheesecloth and washed twice with sterile water to remove media traces. The filtering step also largely eliminated hyphal biomass but retained spores. The Swiss winter wheat cultivar Claro was used for virulence assays (provided by DSP Delley, Inc.). Four seeds were sown in pots with commercial compost soil in triplicates. The pots were frequently in the growth chamber. The plants were grown under controlled conditions as follows: 16/8 hours day/night periods at 18°C throughout the experiment. The growth chamber was maintained at 70% humidity. Plants were grown for three weeks before infection with *Z. tritici*. To initiate infections, washed spores were diluted to 2 x 10^5 spores/ml in 15 ml of sterile water containing 0.1% TWEEN20. For

each isolate, plants from three pots were infected using spray bottles. After spray inoculation, the plants were allowed to dry before sealing them in clear plastic bags to maintain 100% humidity for 48 hours. Plastic bags were removed after 48 hours and conditions were kept as described above.

Automated image-based evaluation of infection

Twenty-one days post inoculation (dpi), the second leaf of each plant was cut and fixed on a barcoded white paper. Leaves were scanned immediately using a flatbed scanner at 1200 dpi. The scanned images were batch-processed using a macro (Stewart *et al.*, 2016; Karisto *et al.*, 2017) based on routines implemented in the image analysis software ImageJ (Rasband, W.S., ImageJ; U. S. National Institutes of Health, http://imagej.nih.gov/ij/, 1997–2012). Briefly, the macro recorded the total leaf area, total lesion area, the number of pycnidia, mean size of pycnidia and pycnidia grey value. The percent leaf area covered by lesions (PLACL) was calculated as the ratio of the total lesion area and total leaf area (Karisto *et al.*, 2018).

Whole-genome sequencing, variant calling and RNA-seq analyses

Approximately 100 mg of lyophilized spores were used to extract high-quality genomic DNA using the Qiagen DNeasy Plant Mini Kit according to the manufacturer's protocol. We sequenced paired-end reads of 100 bp each with an insert size of ~550 bp on the Illumina HiSeq 4000 platform. Raw reads are available on the NCBI Sequence Read Archive under the BioProject PRJNA596434 (Oggenfuss *et al.*, 2020). For RNA sequencing, the same isolates were cultured in a Vogel Minimal N Medium (Vogel, 1956) where ammonium nitrate was replaced with potassium nitrate and ammonium phosphate (Metzenberg, 2003). The medium contained no sucrose and agarose in order to induce hyphal growth. Total RNA was isolated from the filtered mycelium after 10-15 days using the NucleoSpin® RNA Plant and Fungi kit.

The RNA concentration and integrity were checked using a Qubit 2.0 Fluorometer and an Agilent 4200 TapeStation System, respectively. Only high-quality RNA (RIN>8) was used to prepare TruSeq stranded mRNA libraries with a 150 bp insert size and sequenced on an Illumina HiSeq 4000 in the single-end mode for 100 bp.

Sequencing filtering and analysis

We performed sequencing quality checks using FastQC v. 0.11.9. (Andrews S., 2010) and extracted read counts. Sequencing reads were then trimmed for adapter sequences and sequencing quality using Trimmomatic v. 0.39 (Bolger et al., 2014) using the following settings: illuminaclip=TruSeq3-PE.fa:2:30:10, leading=10, trailing=10, sliding-window=5:10 and minlen=50. Trimmed sequencing reads were aligned to the reference genome IPO323 (Goodwin et al., 2011); accessible from https://fungi.ensembl.org/Zymoseptoria_tritici/Info/Index) and the mitochondrial sequence (European Nucleotide Archive EU090238.1) using Bowtie2 v. 2.4.1 (Langmead & Salzberg, 2012). Multi-sample joint variant calling was performed using the HaplotypeCaller and GenotypeGVCF tools of the GATK package v. 4.0.1.2 (McKenna et al., 2010). We retained only SNP variants (excluding indels) and proceeded to hard filtering using the GATK VariantFiltration tool based on the following cutoffs: QD < 5.0; QUAL < 1000.0; MQ < 20.0; -2 > ReadPosRankSum > 2.0; -2 > MQRankSum > 2.0; -2 > BaseQRankSum > 2.0. After filtering for locus level genotyping rate (>50%) and minor allele count (MAC) of 1 using VCFtools v. 0.1.15 (Danecek et al., 2011). Similarly, RNA sequences were checked for quality using FastQC v. 0.11.9. and trimmed with Trimmomatic v0.39 to remove adapter sequences and low-quality reads with parameters: illuminaclip:TruSeq3-SE.fa:2:30:10 leading=3, trailing=3, sliding-window=4:15 and minlen=36. Trimmed sequences were aligned to the

reference genome IPO323 using HISAT2 v. 2.1.0 (source?) with the parameter "--RNA-strandedness reverse".

Population genetic analyses

Population structure and relatedness among individuals in the mapping population may be a source of *p*-value inflation due to non-random phenotype-genotype associations (Bergelson & Roux, 2010; Korte & Farlow, 2013). To account for this, we analyzed the population structure and genetic relatedness of all isolates by performing a principal component analysis (PCA). We performed and visualized the PCA using the R packages vcfR v. 1.8.0 (Knaus & Grünwald, 2017), adegenet v. 2.1.1 (Jombart & Ahmed, 2011), and ggplot2 v. 3.1.0 (Wickham, 2016). We also generated an unrooted phylogenetic network using SplitsTree v4.14.6 (Huson, 1998). File format conversions were performed using PGDSpider v2.1.1.5 (Lischer & Excoffier, 2012). To identify groups of clonal genotypes, we calculated the pairwise genetic distances between all genotypes using the function "dist.dna" included in the R package ape v. 5.3 (Paradis & Schliep, 2019). Isolate pairs with a pairwise genetic distance below 0.01 were considered as clones for further analyses (see (Singh *et al.*, 2020) for more details). The SNP-based heritability (h^2_{snp}; equivalent to narrow-sense heritability) for each trait was estimated using the genome-wide complex trait analysis (GCTA) tool v.1.93.0 (Yang *et al.*, 2011). The h^2_{snp} was estimated using a genome-based restricted maximum likelihood (GREML) approach using the phenotypic values of each trait and considering the additive effect of all the SNPs represented by the GRM.

Genome-wide association mapping and linkage disequilibrium analyses

We performed GWAS based on mixed linear models accounting for kinship (MLM K). We first estimated relatedness among genotypes by computing a kinship matrix using the scaled

identity by state (IBS) algorithm implemented in TASSEL v. 20201110 (Bradbury *et al.*, 2007). We included the kinship matrix as a random effect in the mixed linear models for association mapping using TASSEL. We used the allelic effect output of TASSEL to compute the pairwise genetic correlation (Spearman's correlation) values using complete observations (use = pairwise.complete.obs) and visualized the values using the *ggcorr* function from the GGally R package (source). Similarly, we used the Spearman's correlation to compute pairwise phenotypic trait correlations. Association mapping outcomes were visualized using the R package *qqman* v 0.1.4 (Turner, 2014). We considered associations to be significant when *p*-values were smaller than the Bonferroni threshold at $\alpha = 0.05$ ($p < 1.1e\text{-}07$). False discovery rate (FDR) thresholds of 5% were determined using the *p.adjust* function in the stat package in R. We explored the genomic regions containing significantly associated loci using the "closest" command in bedtools v. 2.29.0 (Quinlan & Hall, 2010). Regions in the genome spanning the most significant associations were investigated for linkage disequilibrium patterns. We calculated the linkage disequilibrium r^2 between marker pairs using the "option–hap-r2" in VCFtools v. 0.1.15 (Danecek *et al.*, 2011) with "--ld-window-bp" of 10000. A heatmap was generated based on the r^2 values with the R package *LDheatmap* v 0.99-7 (Shin *et al.*, 2006).

De novo genome assemblies, TE annotation, synteny analyses

We analyzed the locus surrounding the genes *Zt09_1_01590* and *Zt09_1_01591* in multiple completely assembled genomes of isolates collected in Switzerland, the United States, Australia and Israel covering the global distribution range of the pathogen (Badet et al., 2020). For synteny plots, the available repeat-masked chromosome-scale assemblies were analyzed using pairwise BLASTN. Information on BLAST hits among homologous chromosomes was visualized in R using the *genoplotR* package (Guy *et al.*, 2010). We analyzed signatures of

repeat induced point mutations (RIP) using The RIPper online tool available at https://theripper.hawk.rocks/ (van Wyk *et al.*, 2019).

To analyze sequence polymorphism at the locus, we used draft genome assemblies of 432 isolates from previously analyzed field populations in the United States, Switzerland, Israel and Australia (Hartmann *et al.*, 2017; Oggenfuss *et al.*, 2020). Illumina short read data was obtained from the NCBI Sequence Read Archive under the BioProject PRJNA327615 (Hartmann *et al.*, 2017) and PRJNA596434 (Oggenfuss *et al.*, 2020). We used SPAdes version 3.14.0 to produce draft assemblies for each isolate (Bankevich *et al.*, 2012). We ran the tool with the following settings: -k 21,33,55,75,95 --careful. De novo assemblies were annotated for TEs using the TE consensus sequences (https://github.com/crolllab/datasets) generated for the species (Badet *et al.*, 2020). Consensus sequences were previously manually curated and renamed based on the three-letter classification system (Wicker *et al.*, 2007; Bao *et al.*, 2015). The curated consensus sequences were used for annotation of each individual de novo assembly using RepeatMasker version 4.0.8 with a cut-off value set to 250 (Smit & Hubley, 2015), ignoring simple repeats and low complexity regions. Further filtering of the TE annotation included: (1) removal of element annotations shorter than 100 bp, (2) merging of identical adjacent TE families overlapping by more than 100 bp, (3) renaming of overlapping TE families overlapping by more than 100 bp as nested insertions, and (4) grouping of interrupted elements separated by less than 200 bp into a single element using a minimal distance between start and end positions.

Declarations

Ethics approval and consent to participate: not applicable

Consent for publication: not applicable

***Availability of data and materials*:** Illumina short reads were retrieved from the NCBI Sequence Read Archive (BioProject accessions PRJNA327615, PRJNA596434 and PRJNA650267) accessible from https://www.ncbi.nlm.nih.gov/sra. All other data are reported in Supplementary Information.

Competing interests: The authors declare that they have no competing interests.

Funding: The research was funded by a Swiss National Science Foundation grant to DC (number 173265).

Authors' contributions: NKS and DC conceived the study, NKS and TB performed analyses, LA provided datasets, NKS and DC wrote the manuscript with input from all co-authors.

Acknowledgements: We are grateful for advice and critical feedback on a previous version of this manuscript from Ursula Oggenfuss and Guido Puccetti. Seeds for the experiments were kindly provided by DSP Delley Inc.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

Asai S, Furzer OJ, Cevik V, Kim DS, Ishaque N, Goritschnig S, Staskawicz BJ, Shirasu K, Jones JDG. 2018. A downy mildew effector evades recognition by polymorphism of expression and subcellular localization. *Nature Communications* 9: 5192.

Badet T, Croll D. 2020. The rise and fall of genes: origins and functions of plant pathogen pangenomes. *Current Opinion in Plant Biology* 56: 65–73.

Badet T, Oggenfuss U, Abraham L, McDonald BA, Croll D. 2020. A 19-isolate reference-quality global pangenome for the fungal wheat pathogen Zymoseptoria tritici. *BMC Biology* 18: 12.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, *et al.* 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* 19: 455.

Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* 6.

Bennetzen JL, Wang H. 2014. The Contributions of Transposable Elements to the Structure, Function, and Evolution of Plant Genomes. *Annual Review of Plant Biology* 65: 505–530.

Bergelson J, Roux F. 2010. Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. *Nature Reviews Genetics* 11: 867–879.

Białas A, Zess EK, De La Concepcion JC, Franceschetti M, Pennington HG, Yoshida K, Upson JL, Chanclud E, Wu CH, Langner T, *et al.* 2018. Lessons in effector and NLR biology of plant-microbe systems. *Molecular Plant-Microbe Interactions* 31: 34–45.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30: 2114–20.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633–2635.

Brown JKM, Chartrain L, Lasserre-Zuber P, Saintenac C. 2015. Genetics of resistance to Zymoseptoria tritici and applications to wheat breeding. *Fungal Genetics and Biology* 79: 33–41.

van der Burgh AM, Joosten MHAJ. 2019. Plant Immunity: Thinking Outside and Inside the Box. *Trends in Plant Science* 24: 587–601.

Chakraborty S, Newton AC. 2011. Climate change, plant diseases and food security: an overview. *Plant Pathology* 60: 2–14.

Chuma I, Isobe C, Hotta Y, Ibaragi K, Futamata N, Kusaba M, Yoshida K, Terauchi R, Fujita Y, Nakayashiki H, *et al.* 2011. Multiple Translocation of the AVR-Pita Effector Gene among Chromosomes of the Rice Blast Fungus Magnaporthe oryzae and Related Species (B Tyler, Ed.). *PLoS Pathogens* 7: e1002147.

Coll NS, Epple P, Dangl JL. 2011. Programmed cell death in the plant immune system. *Cell Death and Differentiation* 18: 1247.

Courvoisier N, Häner LL, Bertossa M, Thévoz E, Anders M, Stoll P, Weisflog T, Dugon J, Graf B, Hofer M. 2016. *céréales-variétés 2.21 Blé d'automne Juin 2016*.

Cowger C, Brown JKM. 2019. Durability of Quantitative Resistance in Crops: Greater Than We Know? *Annual Review of Phytopathology* 57: 253–277.

Cowger C, Hoffer ME, Mundt CC. 2000. Specific adaptation by Mycosphaerella graminicola

to a resistant wheat cultivar. *Plant Pathology* 49: 445–451.
Cowley M, Oakey RJ. 2013. Transposable Elements Re-Wire and Fine-Tune the Transcriptome. *PLoS Genetics* 9: e1003234.
Croll D, McDonald BA. 2012. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathogens* 8.
Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.
Depotter JRL, Doehlemann G. 2020. Target the core: durable plant resistance against filamentous plant pathogens through effector recognition. *Pest Management Science* 76: 426–431.
Dickman MB, de Figueiredo P. 2013. Death Be Not Proud—Cell Death Control in Plant Fungal Interactions. *PLoS Pathogens* 9: e1003542.
Dong S, Qutob D, Tedman-Jones J, Kuflu K, Wang Y, Tyler BM, Gijzen M. 2009. The Phytophthora sojae Avirulence Locus Avr3c Encodes a Multi-Copy RXLR Effector with Sequence Polymorphisms among Pathogen Strains. *PLoS ONE* 4: e5556.
Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: Waltz with plants. *Current Opinion in Genetics and Development* 35: 57–65.
Dutta A, Croll D, McDonald BA, Barrett LG. 2020a. Maintenance of variation in virulence and reproduction in populations of an agricultural plant pathogen. *Evolutionary Applications*: eva.13117.
Dutta A, Hartmann FE, Francisco CS, McDonald BA, Croll D. 2020b. Mapping the adaptive landscape of a major agricultural pathogen reveals evolutionary constraints across heterogeneous environments. *bioRxiv*: 2020.07.30.229708.
Faino L, Seidl MF, Shi-Kunne X, Pauper M, Berg GCM van den, Wittenberg AHJ, Thomma BPHJ. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Research* 26: 1091–1100.
Fones H, Gurr S. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology* 79: 3–7.
Fouché S, Badet T, Oggenfuss U, Plissonneau C, Francisco CS, Croll D. 2020. Stress-Driven Transposable Element De-repression Dynamics and Virulence Evolution in a Fungal Pathogen (I Arkhipova, Ed.). *Molecular Biology and Evolution* 37: 221–239.
Fouché S, Mence Plissonneau C, Croll D. 2018. The birth and death of effectors in rapidly evolving filamentous pathogen genomes. *Curr Op Micro* 46:34-42.
Frantzeskakis L, Di Pietro A, Rep M, Schirawski J, Wu C, Panstruga R. 2020. Rapid evolution in plant–microbe interactions – a molecular genomics perspective. *New Phytologist* 225: 1134–1142.
Gardiner DM, Rusu A, Barrett L, Hunter GC, Kazan K. 2020. Can natural gene drives be part of future fungal pathogen control strategies in plants? *New Phytologist* 228: 1431–1439.
Gladyshev E. 2017. Repeat-Induced Point Mutation and Other Genome Defense Mechanisms in Fungi. In: The Fungal Kingdom. Washington, DC, USA: ASM Press, 687–699.
Gohari AM, Ware SB, Wittenberg AHJ, Mehrabi R, M'Barek S Ben, Verstappen ECP, Lee TAJ van der, Robert O, Schouten HJ, Wit PPJGM de, *et al.* 2015. Effector discovery in the fungal wheat pathogen Zymoseptoria tritici. *Molecular Plant Pathology* 16: 931.
Goodwin SB, M'Barek S Ben, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, van der Lee TAJ, Grimwood J, Aerts A, *et al.* 2011. Finished genome of the fungal wheat pathogen Mycosphaerella graminicola reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genetics* 7: e1002070.
Guy L, Roat Kultima J, Andersson SGE. 2010. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* 26: 2334–2335.

Hartmann FE, Croll D. 2017a. Distinct trajectories of massive recent gene gains and losses in populations of a microbial eukaryotic pathogen. *Molecular Biology and Evolution* 34: 2808–2822.

Hartmann FE, Croll D. 2017b. Distinct trajectories of massive recent gene gains and losses in populations of a microbial eukaryotic pathogen. *Molecular Biology and Evolution* 34: 2808–2822.

Hartmann FE, Sánchez-Vallet A, McDonald BA, Croll D. 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME Journal* 11: 1189–1204.

Huson DH. 1998. SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics* 14: 68–73.

Islam MT, Croll D, Gladieux P, Soanes DM, Persoons A, Bhattacharjee P, Hossain MS, Gupta DR, Rahman MM, Mahboob MG, *et al.* 2016. Emergence of wheat blast in Bangladesh was caused by a South American lineage of Magnaporthe oryzae. *BMC Biology* 14: 84.

Jashni MK, Dols IHM, Iida Y, Boeren S, Beenen HG, Mehrabi R, Collemare J, de Wit PJGM. 2015. Synergistic Action of a Metalloprotease and a Serine Protease from *Fusarium oxysporum* f. sp. *lycopersici* Cleaves Chitin-Binding Tomato Chitinases, Reduces Their Antifungal Activity, and Enhances Fungal Virulence. *Molecular Plant-Microbe Interactions®* 28: 996–1008.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.

Jørgensen LN, Hovmøller MS, Hansen JG, Lassen P, Clark B, Bayles R, Rodemann B, Flath K, Jahn M, Goral T, *et al.* 2014. IPM Strategies and Their Dilemmas Including an Introduction to www.eurowheat.org. *Journal of Integrative Agriculture* 13: 265–281.

Karisto P, Hund A, Yu K, Anderegg J, Walter A, Mascher F, McDonald BA, Mikaberidze A. 2017. Ranking quantitative resistance to Septoria tritici blotch in elite wheat cultivars using automated image analysis. *Phytopathology* 108:568-581

Karisto P, Hund A, Yu K, Anderegg J, Walter A, Mascher F, McDonald BA, Mikaberidze A. 2018. Ranking quantitative resistance to septoria tritici blotch in elite wheat cultivars using automated image analysis. *Phytopathology* 108: 568–581.

Knaus BJ, Grünwald NJ. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Molecular Ecology Resources. 17:44–53.

Korte A, Farlow A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* 9: 29.

Krishnan P, Ma X, McDonald BA, Brunner PC. 2018a. Widespread signatures of selection for secreted peptidases in a fungal plant pathogen. *BMC evolutionary biology* 18: 7.

Krishnan P, Meile L, Plissonneau C, Ma X, Hartmann FE, Croll D, McDonald BA, Sánchez-Vallet A. 2018b. Transposable element insertions shape gene regulation and melanin production in a fungal pathogen of wheat. *BMC Biology* 16: 78.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* 9: 357–9.

Langner T, Göhre V. 2016. Fungal chitinases: function, regulation, and potential roles in plant/pathogen interactions. *Current Genetics* 62: 243–254.

Levy L, Courvoisier N, Rechsteiner S, Herrera J, Brabant C, Hund A, Weissflog T, Dierauer H, Pellet D. 2017. Winterweizen: Bilanz aus 15 Jahren Sortenprüfung unter extensiven Anbaubedingungen. *Agrarforschung Schweiz* 8: 300–309.

Lischer HEL, Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28: 298–299.

Longya A, Chaipanya C, Franceschetti M, Maidment JHR, Banfield MJ, Jantasuriyarat C.

2019. Gene Duplication and Mutation in the Emergence of a Novel Aggressive Allele of the *AVR-Pik* Effector in the Rice Blast Fungus. *Molecular Plant-Microbe Interactions®* 32: 740–749.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, *et al.* 2010. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature* 464: 367–373.

Manning VA, Pandelova I, Dhillon B, Wilhelm LJ, Goodwin SB, Berlin AM, Figueroa M, Freitag M, Hane JK, Henrissat B, *et al.* 2013. Comparative Genomics of a Plant-Pathogenic Fungus, Pyrenophora tritici-repentis, Reveals Transduplication and the Impact of Repeat Elements on Pathogenicity and Population Divergence. *G3: Genes, Genomes, Genetics* 3: 41–63.

McCann HC. 2020. Skirmish or war: the emergence of agricultural plant pathogens. *Current Opinion in Plant Biology* 56: 147–152.

Mcgowan J, Fitzpatrick DA. 2017. Genomic, Network, and Phylogenetic Analysis of the Oomycete Effector Arsenal. *mSphere* 2:e00408-17

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al.* 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20: 1297–303.

Meile L, Croll D, Brunner PC, Plissonneau C, Hartmann FE, McDonald BA, Sánchez-Vallet A. 2018. A fungal avirulence factor encoded in a highly plastic genomic region triggers partial resistance to septoria tritici blotch. *New Phytologist* 219: 1048–1061.

Meile L, Peter J, Puccetti G, Alassimone J, McDonald BA, Sánchez-Vallet A. 2020. Chromatin Dynamics Contribute to the Spatiotemporal Expression Pattern of Virulence Genes in a Fungal Plant Pathogen. *mBio* 11:e02343-20

Metzenberg RL. 2003. Vogel's Medium N salts: avoiding the need for ammonium nitrate. *Fungal Genetics Reports* 50: 14–14.

Mikaberidze A, McDonald BA. 2020. A tradeoff between tolerance and resistance to a major fungal pathogen in elite wheat cultivars. *New Phytologist* 226: 879–890.

Mohd-Assaad N, McDonald BA, Croll D. 2019. The emergence of the multi-species NIP1 effector in *Rhynchosporium* was accompanied by high rates of gene duplications and losses. *Environmental Microbiology* 21: 2677–2695.

Muszewska A, Stepniewska-Dziubinska MM, Steczkiewicz K, Pawlowska J, Dziedzic A, Ginalski K. 2017. Fungal lifestyle reflected in serine protease repertoire. *Scientific Reports* 7: 9147.

Oggenfuss U, Badet T, Wicker T, Hartmann FE, Singh NK, Abraham LN, Karisto P, Vonlanthen T, Mundt CC, McDonald BA, *et al.* 2020. A population-level invasion by transposable elements in a fungal pathogen. *bioRxiv*: 2020.02.11.944652.

Omrane S, Audéon C, Ignace A, Duplaix C, Aouini L, Kema G, Walker A-S, Fillinger S. 2017. Plasticity of the MFS1 Promoter Leads to Multidrug Resistance in the Wheat Pathogen Zymoseptoria tritici. *mSphere* 2:e00393-17.

Palma-Guerrero J, Torriani SFF, Zala M, Carter D, Courbot M, Rudd JJ, McDonald BA, Croll D. 2016. Comparative transcriptomic analyses of *Z ymoseptoria tritici* strains show complex lifestyle transitions and intraspecific variability in transcription profiles. *Molecular Plant Pathology* 17: 845–859.

Paradis E, Schliep K. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35: 526–528.

Plissonneau C, Hartmann FE, Croll D. 2018. Pangenome analyses of the wheat pathogen Zymoseptoria tritici reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology* 16.

Plissonneau C, Stürchler A, Croll D, Taylor JW. 2016. The Evolution of Orphan Regions in Genomes of a Fungal Pathogen of Wheat. *mBio* 7(5):e01231-16

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. 2015a. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66: 513–545.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. 2015b. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66: 513–545.

Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842.

Rebollo R, Romanish MT, Mager DL. 2012. Transposable Elements: An Abundant and Natural Source of Regulatory Sequences for Host Genes. *Annual Review of Genetics* 46: 21–42.

Rouxel T, Balesdent M-H. 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of Brassica crops, *Leptosphaeria maculans*. *New Phytologist* 214: 526–532.

Rovenich H, Boshoven JC, Thomma BP. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology* 20: 96–103.

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The genome biology of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology* 56: 21–40.

Santana MF, Silva JC, Batista AD, Ribeiro LE, da Silva GF, de Araújo EF, de Queiroz M V. 2012. Abundance, distribution and potential impact of transposable elements in the genome of Mycosphaerella fijiensis. *BMC Genomics* 13: 720.

Santana MF, Silva JC, Mizubuti ES, Araújo EF, Condon BJ, Turgeon B, Queiroz M V. 2014. Characterization and potential evolutionary impact of transposable elements in the genome of Cochliobolus heterostrophus. *BMC Genomics* 15: 536.

Savary S. 2020. Plant health and food security. *Journal of Plant Pathology* 102: 605–607.

Selin C, Kievit TR de, Belmonte MF, Fernando WGD. 2016. Elucidating the Role of Effectors in Plant-Fungal Interactions: Progress and Challenges. *Frontiers in Microbiology* 7: 600

Shin JH, Blay S, McNeney B, Graham J. 2006. LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *Journal of Statistical Software* 16: 1–9.

Singh NK, Chanclud E, Croll D. 2020. Population-level deep sequencing reveals the interplay of clonal and sexual reproduction in the fungal wheat pathogen Zymoseptoria tritici. *bioRxiv*: 2020.07.07.191510.

Smit A, Hubley R. 2015. RepeatModeler Open-1.0. http://www.repeatmasker.org

Sperschneider J, Gardiner DM, Thatcher LF, Lyons R, Singh KB, Manners JM, Taylor JM. 2015. Genome-Wide Analysis in Three Fusarium Pathogens Identifies Rapidly Evolving Chromosomes and Genes Associated with Pathogenicity. *Genome Biology and Evolution* 7: 1613–1627.

Stauber L, Prospero S, Croll D. 2020. Comparative Genomics Analyses of Lifestyle Transitions at the Origin of an Invasive Fungal Pathogen in the Genus Cryphonectria. *mSphere* 5:e00737-20

Stewart E l., Croll D, Lendenmann MH, Sanchez-Vallet A, Hartmann FE, Palma-Guerrero J, Ma X, McDonald BA. 2018. Quantitative trait locus mapping reveals complex genetic architecture of quantitative virulence in the wheat pathogen *Zymoseptoria tritici*. *Molecular Plant Pathology* 19: 201–216.

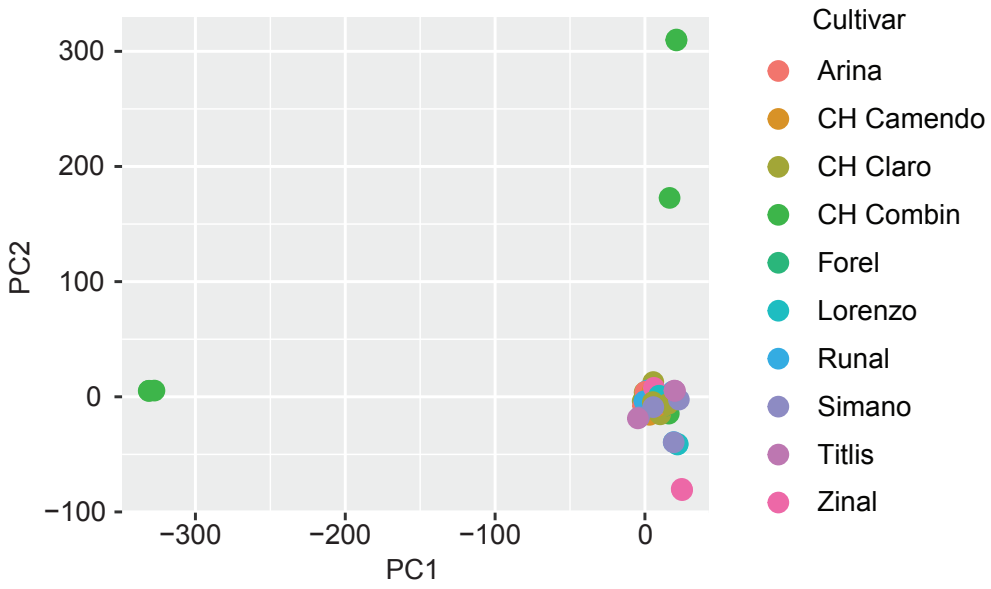
Stewart EL, Hagerty CH, Mikaberidze A, Mundt C, Zhong Z, McDonald BA. 2016. An improved method for measuring quantitative resistance to the wheat pathogen

Zymoseptoria tritici using high throughput automated image analysis. *Phytopathology* 106: 782–788.
Strange RN, Scott PR. 2005. Plant disease: A threat to global food security. *Annual Review of Phytopathology* 43: 83–116.
Subbarao K V, Sundin GW, Klosterman SJ. 2015. Focus Issue Articles on Emerging and Re-Emerging Plant Diseases. *Phytopathology* 105:852-854
Torres DE, Oggenfuss U, Croll D, Seidl MF. 2020. Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fungal Biology Reviews* 34: 136–143.
Torriani SFF, Stukenbrock EH, Brunner PC, McDonald BA, Croll D. 2011. Evidence for Extensive Recent Intron Transposition in Closely Related Fungi. *Current Biology* 21: 2017–2022.
Turner SD. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*: 005165.
Vleeshouwers VGAA, Oliver RP. 2014. Effectors as Tools in Disease Resistance Breeding Against Biotrophic, Hemibiotrophic, and Necrotrophic Plant Pathogens. *Molecular Plant-Microbe Interactions®* 27: 196–206.
Vogel HJ. 1956. A Convenient Growth Medium for Neurospora crassa. *Microbial Genetics Bulletin* 13: 42–47.
Wang Q, Jiang C, Wang C, Chen C, Xu J-R, Liu H. 2017. Characterization of the Two-Speed Subgenomes of Fusarium graminearum Reveals the Fast-Speed Subgenome Specialized for Adaption and Infection. *Frontiers in Plant Science* 8:140
Wang L, Sun Y, Sun X, Yu L, Xue L, He Z, Huang J, Tian D, Hurst LD, Yang S. 2020. Repeat-induced point mutation in Neurospora crassa causes the highest known mutation rate and mutational burden of any cellular life. *Genome Biology* 21: 142.
Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, *et al.* 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* 8: 973–982.
Wickham H. 2016. Ggplot2 : elegant graphics for data analysis. *Springer-Verlag New York*.
Van de Wouw AP, Cozijnsen AJ, Hane JK, Brunner PC, McDonald BA, Oliver RP, Howlett BJ. 2010. Evolution of Linked Avirulence Effectors in Leptosphaeria maculans Is Affected by Genomic Environment and Exposure to Resistance Genes in Host Plants (A Sil, Ed.). *PLoS Pathogens* 6: e1001180.
Wu C-H, Abd-El-Haliem A, Bozkurt TO, Belhaj K, Terauchi R, Vossen JH, Kamoun S. 2017. NLR network mediates immunity to diverse plant pathogens. *Proceedings of the National Academy of Sciences* 114: 8113–8118.
Wu J, Kou Y, Bao J, Li Y, Tang M, Zhu X, Ponaya A, Xiao G, Li J, Li C, *et al.* 2015. Comparative genomics identifies the *Magnaporthe oryzae* avirulence effector *AvrPi9* that triggers *Pi9* -mediated blast resistance in rice. *New Phytologist* 206: 1463–1475.
van Wyk S, Harrison CH, Wingfield BD, De Vos L, van der Merwe NA, Steenkamp ET. 2019. The RIPper, a web-based tool for genome-wide quantification of Repeat-Induced Point (RIP) mutations. *PeerJ* 7: e7447.
Xue M, Yang J, Li Z, Hu S, Yao N, Dean RA, Zhao W, Shen M, Zhang H, Li C, *et al.* 2012. Comparative Analysis of the Genomes of Two Field Isolates of the Rice Blast Fungus Magnaporthe oryzae. *PLoS Genetics* 8: e1002869.
Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. *American Journal of Human Genetics* 88: 76.
Yoshida K, Saunders DGO, Mitsuoka C, Natsume S, Kosugi S, Saitoh H, Inoue Y, Chuma I, Tosa Y, Cano LM, *et al.* 2016. Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable

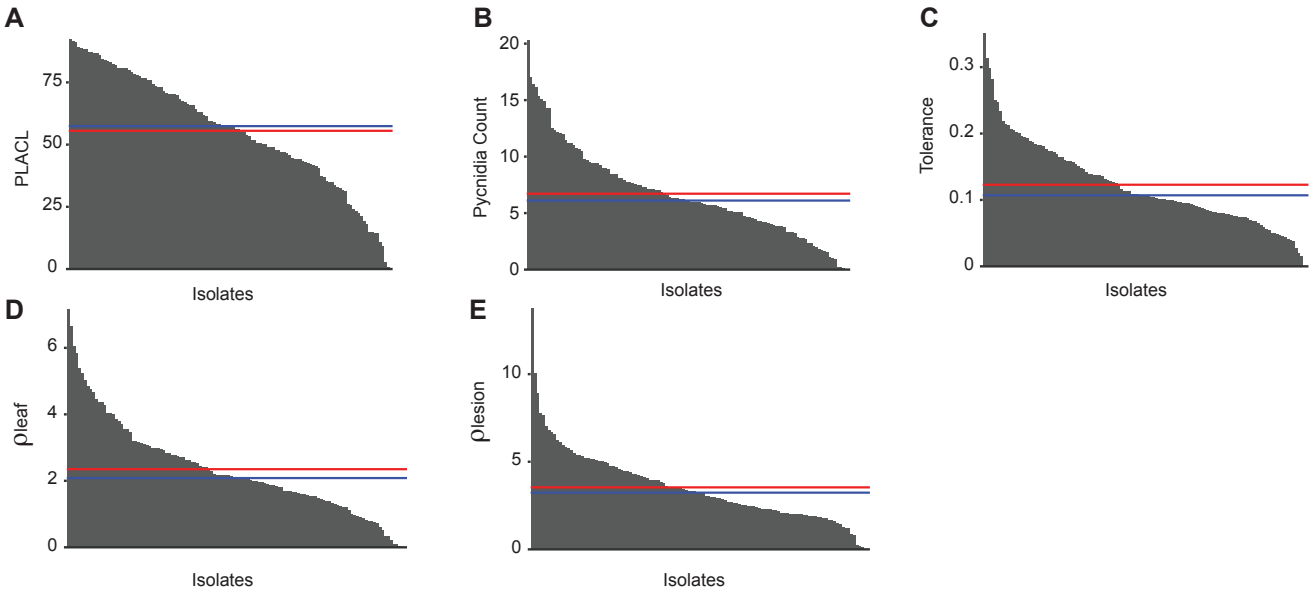
elements. *BMC genomics* 17: 370.
Zhong Z, Marcel TC, Hartmann FE, Ma X, Plissonneau C, Zala M, Ducasse A, Confais J, Compain J, Lapalu N, *et al.* 2017. A small secreted protein in Zymoseptoria tritici is responsible for avirulence on wheat cultivars carrying the Stb6 resistance gene. *New Phytologist* 214: 619–631.

Supplementary Information

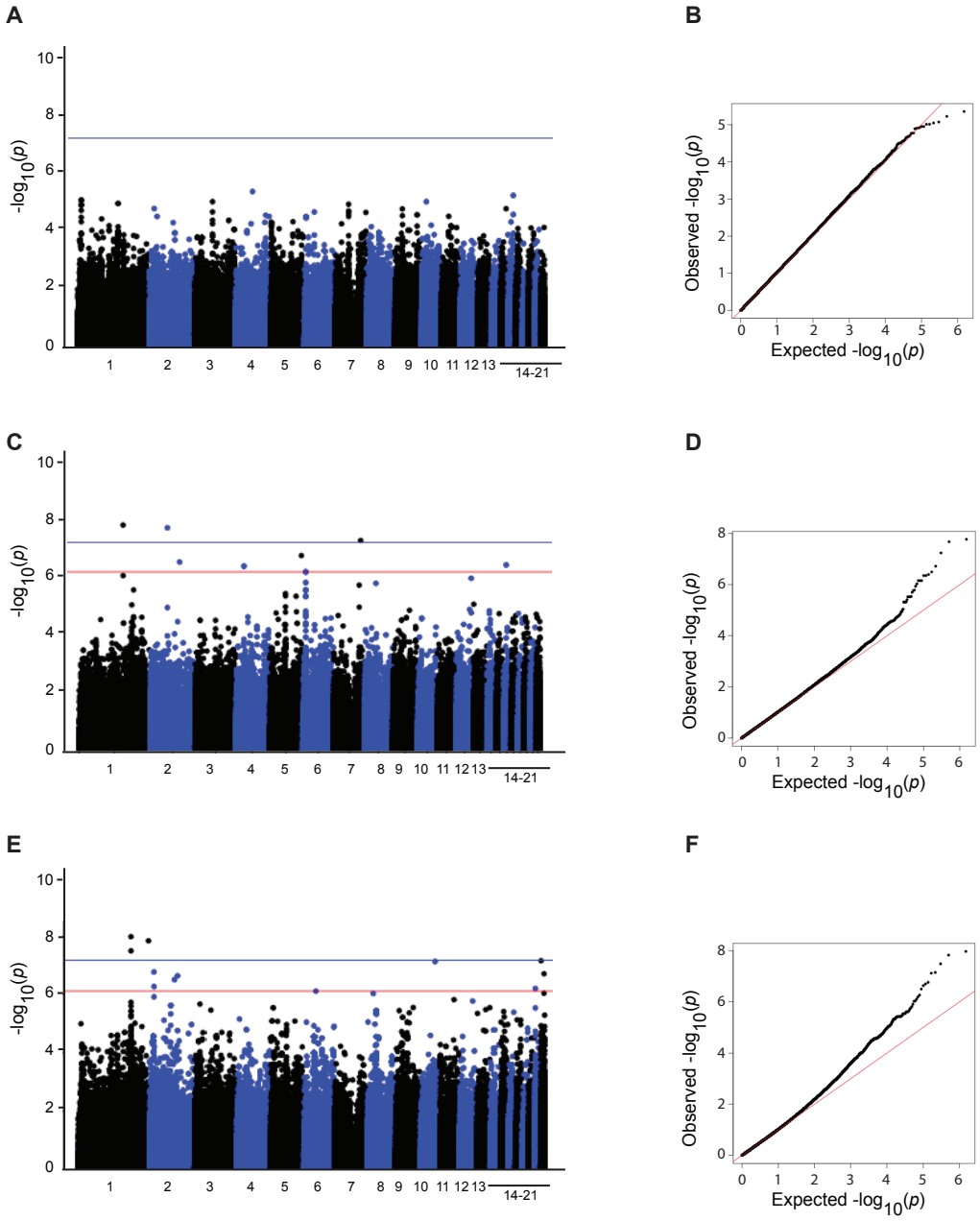
Supplementary Figures



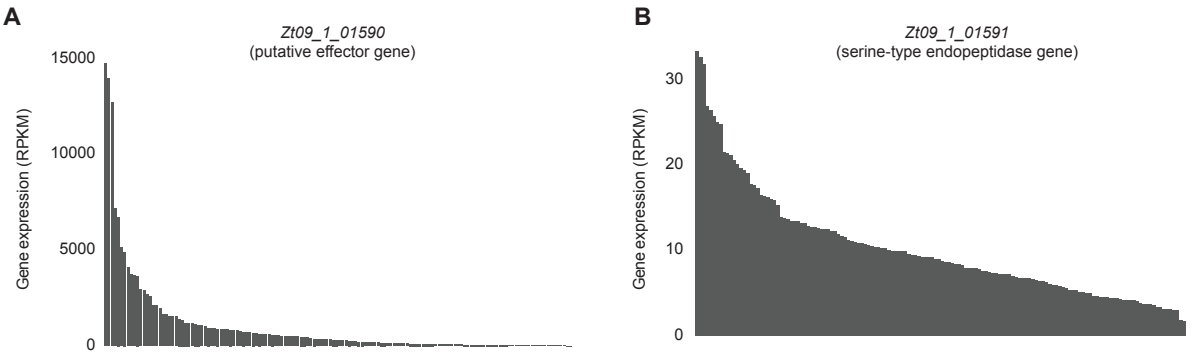
Supplementary Figure 1 : The first two principal components (PC) from a PC analysis of 788'313 genome-wide SNPs. Isolates are color-coded by the cultivar of the origin.



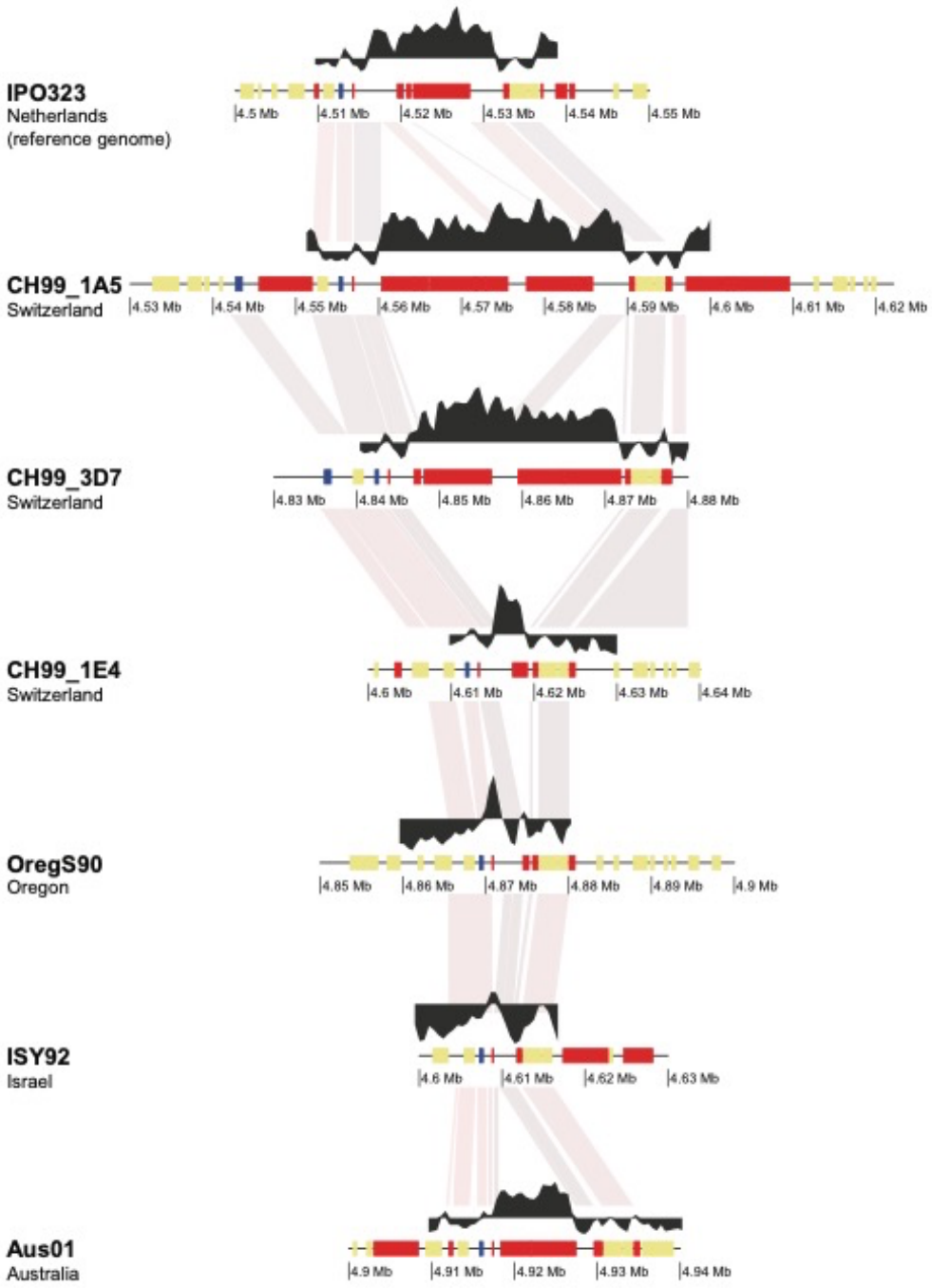
Supplementary Figure 2 : Phenotypic trait values. Red and blue lines represent the mean and median, respectively. A) The percentage of leaf area covered by lesion (PLACL). B) pycnidia count. C) Tolerance expressed as the pycnidia count divided by PLACL. D) ρ_{leaf} is the pycnidia count per cm^2 of leaf area and E) ρ_{lesion} is defined as the total number of pycnidia divided by per cm^2 lesion area.



Supplementary Figure 3 : Manhattan and QQ-plot representing of the genome-wide association mapping analyses. The blue line corresponds to the Bonferroni threshod (alpha = 0.05) and the red line corresponds to the 5% FDR. A-B) Percentage of leaf area covered by lesions (PLACL), C-D) tolerance E-F) ρ_{lesion}



Supplementary Figure 4 : Gene expression in reads per kilobase of transcript, per million mapped reads (RPKM) for genes closest to the top-associated SNP. A) Putative effector gene (*Zt09_1_01590*) and B) serine-type endopeptidase gene (Zt09_1_01591).



Supplementary Figure 5 : The Large RIP Affected Regions (LRARs) composite index was calculated using The RIPer tool (van Wyk et al., 2019) and shown in black. The region flanked by the genes *Zt09_1_01590* and *Zt09_1_01591* is shown in complete genome assemblies of seven isolates from global population.

Supplementary Tables

separate excel file available at https://doi.org/10.1101/2021.02.16.431386

Supplementary Table S1: Phenotypic trait values used for GWAS. The percentage of leaf area covered by lesion (PLACL); ρ_{leaf} is the pycnidia count per cm^2 of leaf area; ρ_{lesion} is defined as the total number of pycnidia divided by per cm^2 lesion area. Tolerance is expressed as the pycnidia count divided by PLACL

Supplementary Table S2: Groups of clonal genotypes identified in the GWAS population with information about the collection time point and cultivar of origin. The clonal genotype columns provides a unique identifier. See Singh *et al.* (2020) for more detailed analyses.

Supplementary Table S3: List of significantly associated SNPs above 5% FDR for pycnidia counts .

Supplementary Table S4: Number of isolates from populations on different continents analyzed for transposable element content. Total assembled genomes and total of isolates per population where a scaffold was retrieved containing both genes *Zt09_1_01590* and *Zt09_1_01591*.

Chapter 3

Genome-wide association mapping to identify genes underlying population-level metabolome diversity in a fungal crop pathogen

Nikhil Kumar Singh[1], Gaétan Glauser[2], Daniel Croll[1,*]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland
[2] Neuchâtel Platform of Analytical Chemistry, University of Neuchâtel, 2000 Neuchâtel, Switzerland
[*]Author for correspondence: daniel.croll@unine.ch

Keywords: Fungal pathogens, *Zymoseptoria tritici*, whole-genome sequencing, ultra-high performance liquid chromatography, metabolite genome-wide association mapping, secondary metabolites

CHAPTER 3

Abstract

Fungi produce a wide range of specialized metabolites (SMs) involved in biotic interactions. Pathways for the production of SMs are often encoded in clusters of tightly arranged genes. Such gene clusters can undergo horizontal gene transfers between species and rapid evolutionary change within species. The acquisition, rearrangement and deletion of gene clusters can generate significant metabolome diversity. However, the genetic basis underlying variation in SM production remains poorly understood. Here, we analyzed metabolite production of a large population of the fungal pathogen of wheat, *Zymoseptoria tritici*. The pathogen causes major yield losses and likely utilizes SM to manipulate the host and to interact with the leaf microbiome during infection. The species shows substantial variation in gene clusters encoding SMs. We cultured 102 isolates obtained from a single wheat field under standardized conditions to extract metabolites produced by the cells. Then, we performed untargeted ultra-high performance liquid chromatography to profile the diversity in the produced metabolites. We found that the majority of the detected metabolites are unlikely to be shared or produced in similar quantities across the population suggesting substantial metabolome diversity. Integrating whole-genome sequencing data for all isolates, we performed metabolite genome-wide association mapping to identify loci underlying variation in metabolite production. Associated genes encode three main functional categories including transporters, enzymes related to metabolic pathways and regulators of cellular structures. Gene ontology enrichment analyses revealed an over-representation of endosome functions, which are typically involved in cellular homeostasis of metabolites. Our study highlights the power of association mapping to identify candidate loci underlying metabolite diversity. Integrating metabolome screens should be feasible for a range of different plant pathogens and help prioritize molecular studies.

Introduction

Fungi are capable of synthesizing a wide range of compounds including amino acids, pigments, antibiotics and toxins (Vining, 1990; Hoffmeister & Keller, 2007; Bouws *et al.*, 2008). Such metabolites are typically classified as primary and specialized (or secondary) metabolites. Primary metabolites are essential for growth, development and reproduction and tend to be conserved across the phylogeny of fungi. Specialized metabolites (SM) confer benefits in specific ecological niches but are not essential for cell survival (Keller *et al.*, 2005). The production of primary metabolites is encoded by a vast array of enzymes involved in metabolic reactions. In contrast, the production of SMs in fungi is often encoded by a small set of enzymes establishing a minimal metabolic pathway. The underlying genes tend to be arranged in close proximity forming a biosynthetic gene cluster (BCGs) including tight regulatory control (Keller, 2015). SMs play an important role in shaping species interactions with the microbiome. For example, the production of the antibacterial bikaverin is induced specifically upon contact with the bacterium *Ralstonia solanacearum* in conserved in distant fungal pathogens (Spraker *et al.*, 2018). Besides, SMs are essential for regulating fungal pathogenicity by acting as virulence factors against animal and plant hosts (Fox & Howlett, 2008). In the case of *Fusarium graminearum*, the BGC encoding the pathway for the production of the mycotoxin trichothecene is specifically upregulated during colonization of wheat (Lysøe *et al.*, 2011). SM can also act as messengers in inter and intra-species communications between fungi (Yim *et al.*, 2007; Netzker *et al.*, 2015). The high degree of niche specificity of SMs generates diversity in BGCs among closely related species and even within some species (Keller, 2015, 2019). BGCs are also among the most mobile elements in fungal genomes with extensive evidence for horizontal gene transfer and turnover within species (Lind *et al.*, 2015).

Large-scale fungal genome analyses have revolutionized the discovery of genes underlying the specialized metabolism and rapidly evolving BGCs transferred among species (Keller, 2019; Gil-Serna *et al.*, 2020; Naranjo-Ortiz & Gabaldón, 2020). Genetic changes underlying the production of SM can be pinned down to single nucleotide variation as shown for fumonisin production in *Fusarium* fungi (Proctor *et al.*, 2008). The evolution of BGCs often involves larger changes including the gain and loss of individual genes or entire clusters (Carbone *et al.*, 2007; Tralamazza *et al.*, 2019). BGCs are often located in more repetitive regions of the genome including subtelomeres (Palmer & Keller, 2010). Furthermore, BGCs in some groups of fungi are mainly regulated by epigenetic changes including chromatin modification (Collemare & Seidl, 2019). The fact that BGCs encode entire pathways for the production of SM, horizontal gene transfers effective means for species to acquire novel ecological functions (Cornell *et al.*, 2007; Marcet-Houben & Gabaldón, 2010; Wisecaver *et al.*, 2014). Horizontal gene transfers and rearrangements have led to substantial differences within and among closely related species in the content of BGCs and the potential to produce SMs. Among *Aspergillus* fungi, BGCs contribute to shared virulence profiles based on gliotoxin and fumitremorgin production and species-specific virulence *e.g.* due to fumagillin (Steenwyk *et al.*, 2020). Genetic differences in qualitative and quantitative variation of SM production are hence likely under strong selection. Variation in metabolite production among conspecific isolates can be used for association mapping. Identifying associations between genetic variants in populations with relative levels of metabolite production is performed as metabolite genome-wide association studies (mGWAS) with a primary focus on the plant model *Arabidopsis thaliana* (Strauch *et al.*, 2015). Recent applications on rice cultivars identified genes underlying flavonoid production (Peng *et al.*, 2017) and contributions to the phenol-amides metabolic pathway (Dong *et al.*, 2015). Metabolite variation was often mapped to a small number of major effect loci (Chen *et al.*, 2014; Jacobowitz & Weng, 2020). The intra-specific

variation in metabolite profiles, high-quality genomic resources and experimental tractability make fungi attractive models for genome-wide association mapping approaches.

The ascomycete *Zymoseptoria tritici* is a highly polymorphic fungal pathogen of wheat and shows a marked variability in BGCs among members of the species (Plissonneau *et al.*, 2016, 2018; Badet *et al.*, 2020). The pathogen causes yield losses of ~ 5–30% depending on environmental conditions (Jørgensen *et al.*, 2014; Fones & Gurr, 2015). Populations sampled across the world show evidence for the gain and loss of genes constituting BGCs (Hartmann & Croll, 2017). Comparisons of complete genome assemblies confirmed that BGCs are partly or entirely missing in some isolates (Badet *et al.*, 2020) raising questions about the functional relevance of the encoded SMs. However, the role of SM in the lifecycle of *Z. tritici* remains poorly understood. A metabolome analysis of the wheat infection process showed that lipid metabolism is the initial energy source during leaf colonization prior to the induction of host cell death *(*Rudd *et al.*, 2015). The pathogen upregulates BGCs mostly during the same transition to feeding from dead plant material *(i.e.* necrotrophic lifestyle) ~10 days after infection (Rudd *et al.*, 2015; Palma-Guerrero *et al.*, 2016). Precursors of melanin were also found in the metabolite profile at this stage (Rudd *et al.*, 2015). Melanin, which plays an important role in virulence, ultraviolet protection and anti-microbial resistance in fungi, is one of the best studied SM of *Z. tritici.* The locus underlying variation in melanin production was first identified using quantitative trait mapping and then confirmed to be a polyketide synthase gene cluster (Lendenmann *et al.*, 2014; Krishnan *et al.*, 2018). Among individual variation in melanin accumulation is governed by the insertion of a transposable element, which impacts the regulation of the BGC (Krishnan *et al.*, 2018). Beyond melanin production, the metabolomic diversity within the species and the underlying genetic basis remains unknown.

Here, we take advantage of genome-wide association (GWA) mapping using the production of individual metabolites under standardized conditions as trait values. GWA mapping in *Z. tritici* has been used successfully to identify the genetic basis underlying virulence on different wheat cultivars, resistance to fungicides and temperature adaptation (Hartmann *et al.*, 2017, 2020; Singh *et al.*, 2021; Dutta *et al.*, 2021). We focused on a single wheat field to establish a panel of 102 isolates showing considerable genetic diversity (Singh *et al.*, 2020). We performed untargeted ultra-high performance liquid chromatography tandem mass spectrometry on individual fungal cultures growing under sterile conditions. We found considerable variation in metabolite profiles with the majority of metabolites showing abundance variation among isolates. GWA mapping revealed significantly associated loci in proximity to genes encoding functions related to endosomal functions and transporters. Taken together, our study provides a significant resource to unravel polymorphism underlying metabolome diversity within a species.

Results

Species-wide polymorphism in specialized metabolite gene clusters

A pangenome based on 19 completely assembled genomes of *Z. tritici* isolates collected from six continents and 13 different countries is available for the species (Badet *et al.*, 2020). We retrieved available gene annotation data and identified 29-33 BGCs per genome (Fig 1A). Non-ribosomal peptide synthetase (NRPS) and type 1-polyketide synthase are the most abundant gene clusters. Approximately 72% of the predicted core biosynthetic genes are conserved among isolates and ~24% are accessory (present in 2-18 isolates) (Fig 1B). We also retrieved a singleton core biosynthetic gene found only in an isolate sampled in Tunisia. We found similar proportions for additional biosynthetic genes with ~60% and 30% core and accessory

genes, respectively (Fig 1B). Regulatory genes of gene clusters were all conserved, but only 90% of the transporter genes were conserved (Fig 1B). BGCs show polymorphism also at the level of single fields. Focusing on BGCs encoded in the genomes of four isolates collected in the same year from nearby wheat field in central Europe, seven clusters showed presence/absence variation (Badet *et al.*, 2020) (Fig 1C).

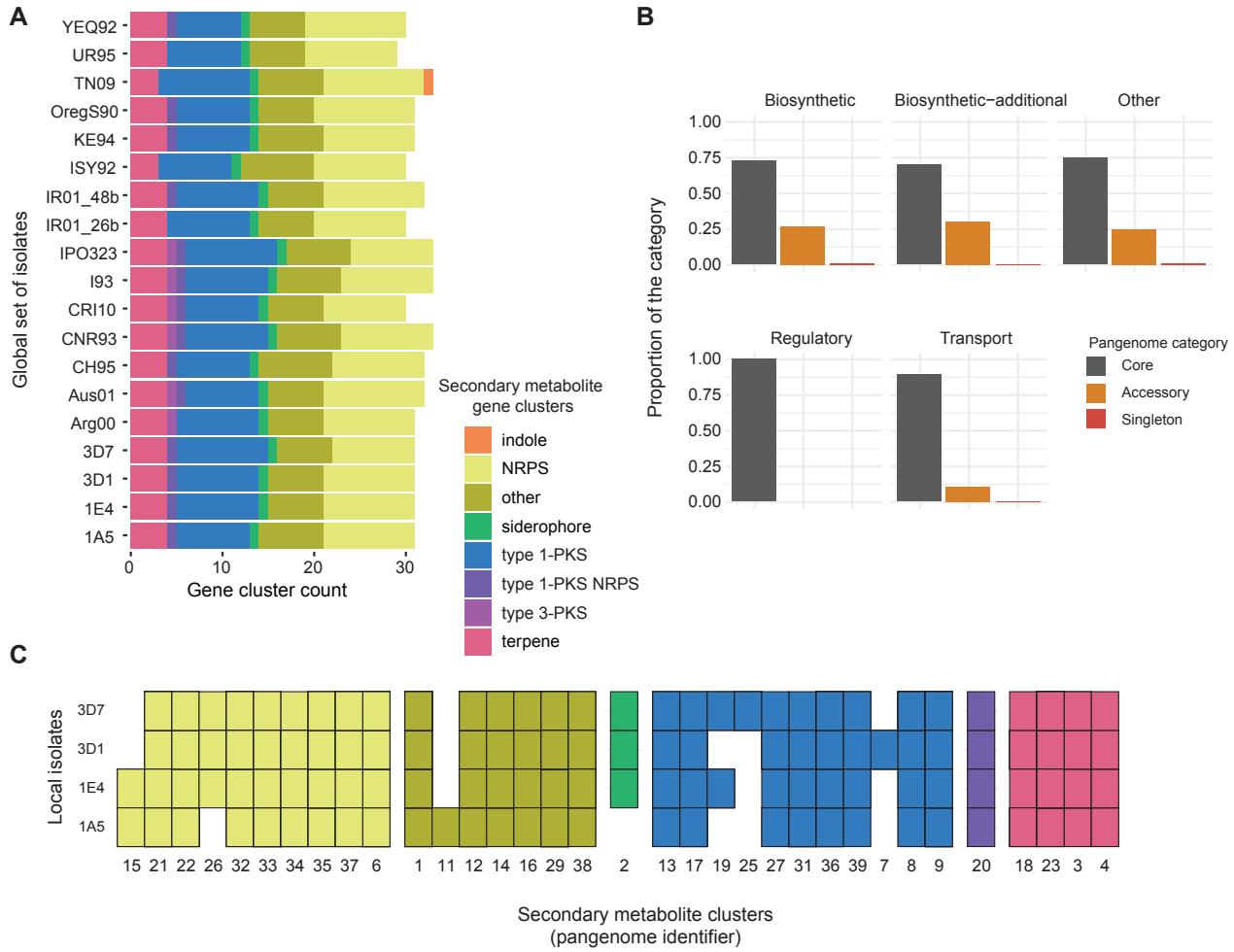


Figure 1: Genetic diversity of specialized metabolite gene clusters in *Zymoseptoria tritici*. A) Detected gene clusters in 19 genomes constituting the species pangenome (Badet *et al.*, 2020). Colors indicate different categories of specialized metabolite gene clusters depending on the metabolite compound or the core biosynthetic gene: indoles, non-ribosomal peptide synthetase (NRPS), siderophores, polyketide synthase (PKS) and terpenes. B) The proportions of core, accessory and singletons specialized metabolite gene clusters in the species pangenome. C) Presence-absence variation of 39 gene clusters in genomes of four isolates collected in Switzerland.

Genetic diversity in a single-field mapping population

To identify whether individuals differ in the production of metabolites and show heritable genetic variation, we performed mGWAS. We analyzed whole genome sequences of 102 isolates isolated from a single wheat field. The mapping population was previously shown to be well-suited for GWAS (Singh *et al.*, 2021). The isolates were collected from 11 genetically different winter wheat cultivars (1-9 isolates per cultivar) from three collection time points (Fig 2A, Supplementary Table S1). At the first collection time point, the wheat seedling was at growth stage (GS) 41 where flag leaves start extending. At the second (GS 75) and third collection (GS 85) time point, plants were fully developed and grains were reaching maturity (Meier, 2001). The average Illumina sequencing coverage was 21X as previously described (Singh *et al.*, 2020). After quality filtering, we retained 788,312 high-confidence single nucleotide polymorphisms (SNPs). We constructed an unrooted phylogenetic network using Splitstree and we found isolates to be at similar genetic distance (Fig 2B). The mapping population also included eight clonal groups for a total of 19 isolates. We used relatedness as a random factor in association mapping to account for the uneven relatedness. A principal component analysis (Fig 2C) identified a small group of outlier isolates. No genetic structure was found among collection time points or cultivars (Singh *et al.*, 2020).

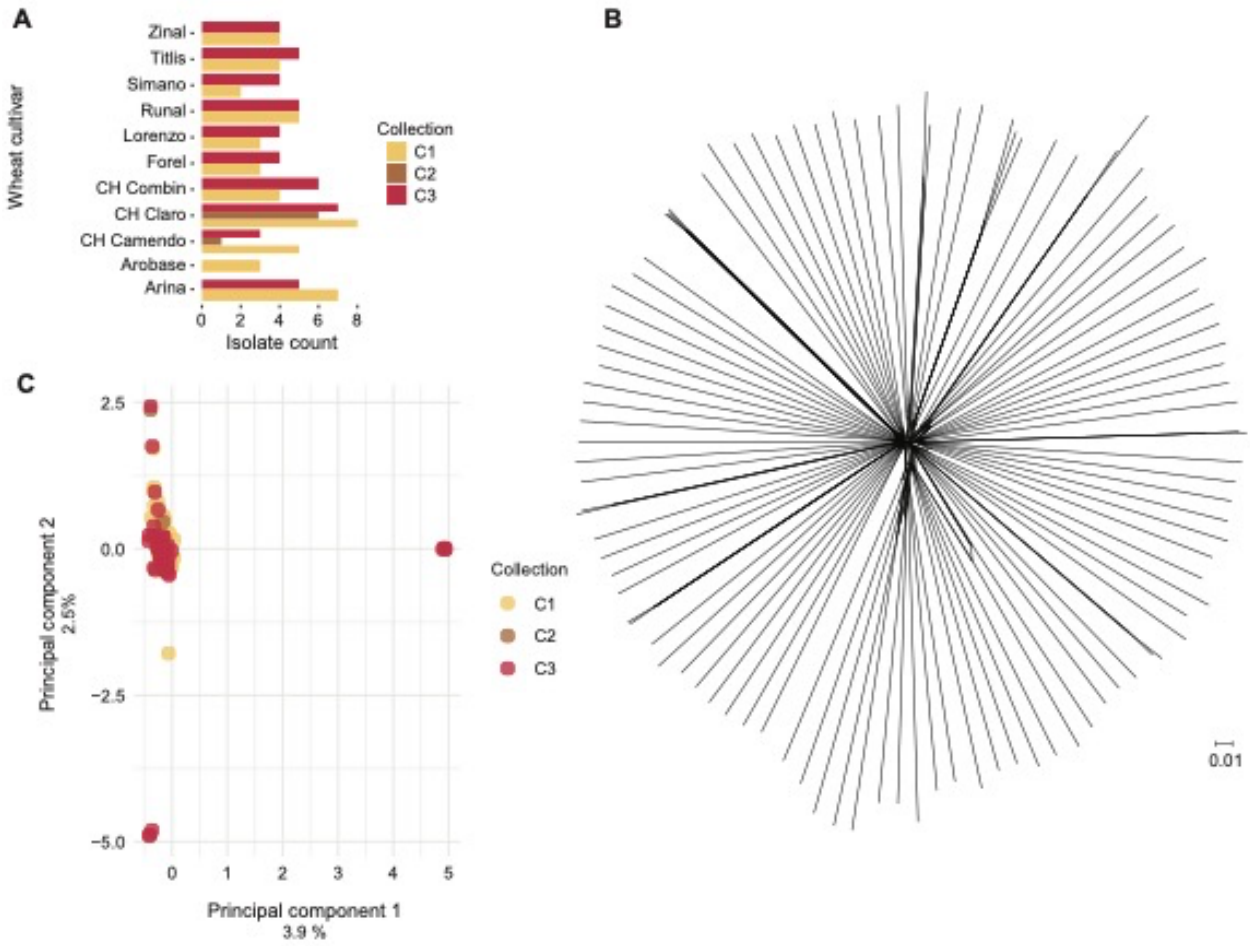


Figure 2: Whole-genome sequence analyses of 102 *Zymoseptoria tritici* isolates collected from a single field. A) Number of isolates collected from each of the eleven cultivars at each collection time point (C1-3; early, middle and late in the season; see also (Singh *et al.*, 2020). B) A SplitsTree phylogenetic network constructed from genome-wide single nucleotide polymorphism (SNP) data. C) The first two principal components (PC) from a PC analysis of genome-wide SNPs. Isolates are color coded by the collection time point.

Untargeted metabolite profiling at the population level

We performed an untargeted metabolite profiling of all 102 isolates using ultra-high performance liquid chromatography tandem mass chromatography (UHPLC-QTOF-MS). The analyzed metabolites were extracted from blastospores after washing. Hence, the process should largely recover intracellular or not yet secreted metabolites. We focused on retention times excluding heavy molecules (*e.g.* phospholipids). We retained 2633 metabolite marker peaks (Supplementary table S2). Using relative abundance across all peaks, we performed a principal component analysis (Fig 3A). Interestingly, the metabolite profiles of different

isolates clustered based on their collection time point (Fig 3A). This is in contrast to the genome-wide differentiation at neutral markers (Fig 2C). We identified the top metabolite markers contributing to the clustering of the isolate metabolite profiles (Fig 3B). The top two markers had mass-to-charge (m/z) ratios of 183.99 and 344.243, respectively. None of the two markers showed reliable hits in databases. The marker *219_1345* was predicted to be dinitrophenol. UHPLC-QTOF-MS–based profiling often detects fragmented peak profiles for individual metabolites. For 79 out of 2633 metabolite markers, a signal was detected in every analyzed isolates. For the remaining 2554 (~96%) metabolite markers, we detected no reliable signal in one or more isolates, which may be due to secondary peaks or represent true differences in metabolic diversity. We performed a down sampling analysis to identify the proportion of metabolite markers detected in isolates. We found that approximately equal proportions of metabolite markers detected in 75%, 50% and 25% of the total isolates (Fig 2C).

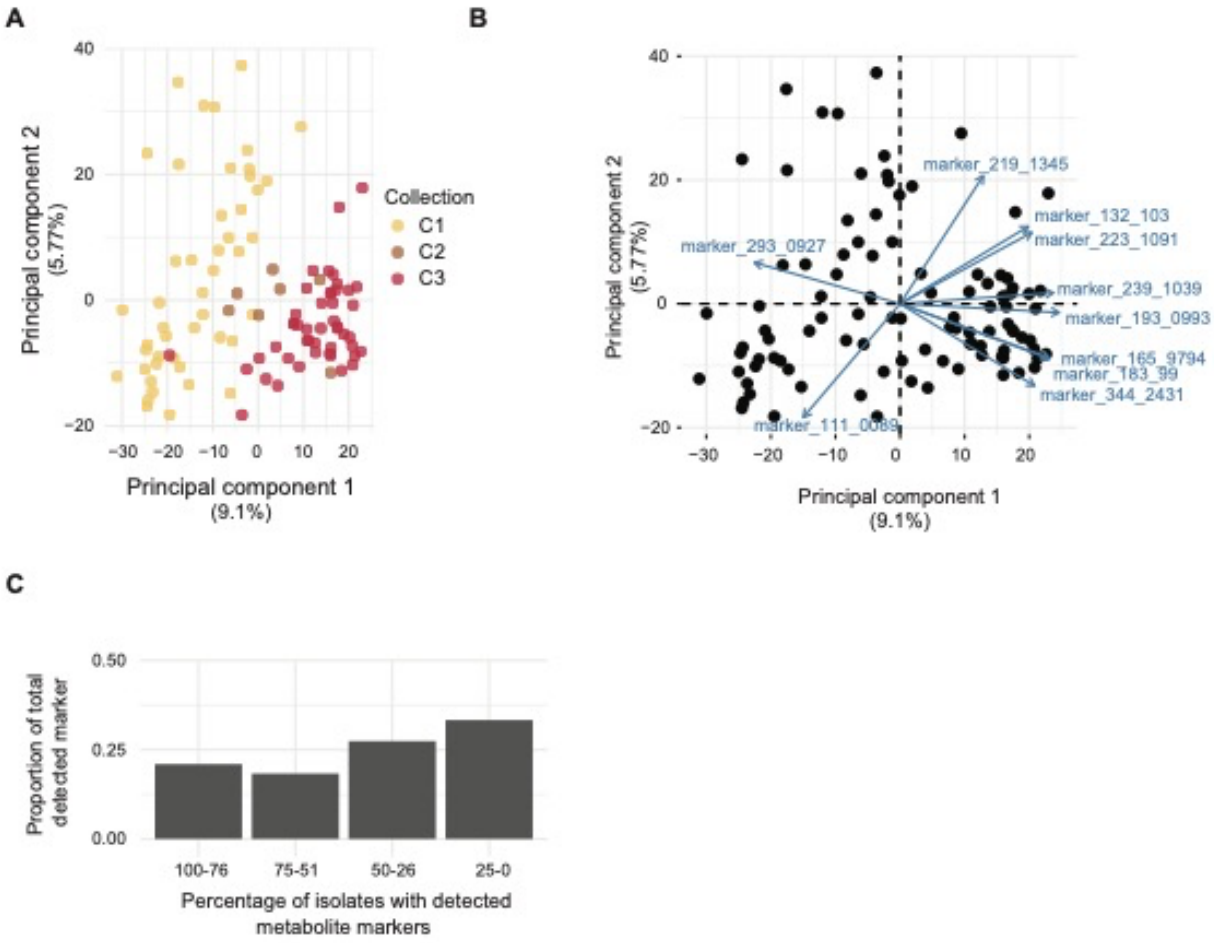


Figure 3: Untargeted metabolite profiling of 102 isolates using ultra-high performance liquid chromatography tandem mass chromatography (UHPLC-QTOF-MS). A) The first two principal components (PC) from a PC analysis of 2633 metabolite markers. Isolates are colour coded by the collection time point. B) Bi-plot of 2633 metabolite markers showing the contribution of the ten metabolite markers contribution most to differentiation of metabolite profiles. Metabolite markers are labelled using their respective *m/z* values. C) Down sampling analysis to identify the proportion of 2633 metabolite markers detected in mapping population

Metabolite-GWAS based on variation in metabolite abundance

Variation in relative abundance among isolates may have a genetic basis (Fig 4A). We performed mGWAS on relative metabolite marker intensities based on mixed linear models taking genetic relatedness into account. We first evaluated top association *p*-values for 2633 metabolite markers. We evaluate evidence for *p*-value inflation and high coefficients of variation were associated with excessively low *p*-values compared to expectations given the size of the mapping population and previous phenotype-genotype association mapping (Fig

4B-C, Supplementary Fig 1) (Hartmann *et al.*, 2017; Dutta *et al.*, 2021). To reduce the risk of *p*-value inflation, we removed all metabolite markers with a relative maximum intensity <100 (Fig 4D). The reduced dataset included 62 metabolite markers and association *p*-value ranging from $1.28e^{-74}$ – $1.89e^{-06}$ (Fig 4E, Supplementary Table S3). Among the retained metabolite markers, the maximum relative intensity and standard deviation among isolates are positively correlated ($r = 0.96$, $p < 2.2e^{-16}$; Fig 4F).

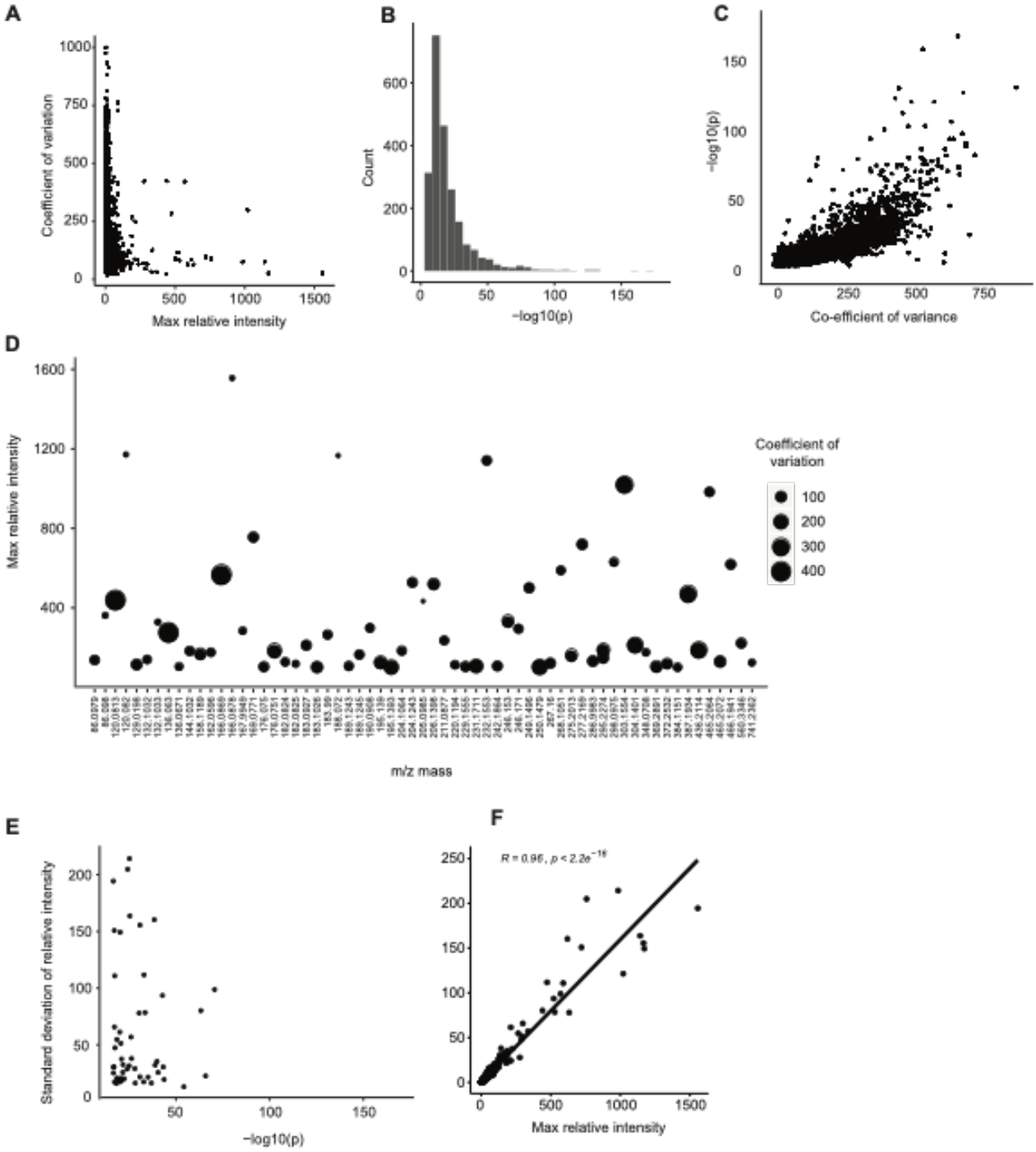


Figure 4: Metabolite abundance variation and metabolite genome wide association studies (mGWAS). A) Maximum relative intensity of 2633 metabolite markers compared to the co-efficient of variation among isolates. B) Histogram showing the distribution of −log10(*p*) values of the most significantly associated single nucleotide polymorphisms (SNP) for each independent mGWAS. C) Distribution of −log10(*p*) values of top associated SNPs against the co-efficient of variation of relative intensities. D) Metabolite markers filtered for a relative intensity >100 (n = 62) with the dot size representing the co-efficient of variation. E) Standard deviation of relative intensities compared to the −log10(*p*) values of the top associated SNP for each mGWAS. F) Correlation of standard deviation of relative intensities and maximum relative intensities.

For each of the 62 high abundance metabolites, we analyzed the closest gene to most significantly associated SNP (Supplementary Table S3; Fig 5A). We found very few associated SNPs on accessory chromosomes (Fig 5A). Linkage disequilibrium on chromosomes decays to $r^2 < 0.2$ within 500-1500 bp (Singh *et al.*, 2020). This range matches also the average gene distance in the *Z. tritici* genome (Goodwin *et al.*, 2011). We found that the most significantly associated SNP for each metabolite marker was at a distance of 0-19.5 kb to the closest gene (mean = 5 kb, median = 3 kb; Fig 5B). For the remainder, we focused only on the most significantly associated SNPs for each metabolite marker falling within 2 kb of a gene (Table 1).

Table 1: List of significantly associated single nucleotide polymorphisms (SNPs) from the metabolite genome-wide association study. Only SNPs within 2 kb of a coding sequence are reported. Markers of interest were characterized using the mass spectrometry reference spectra database MassBank (https://massbank.eu).

Mass-to-charge ratio (m/z)	Putative metabolite	Gene_id	chromosome	SNP position	*p* value	Distance to gene	Predicted function encoded by gene
136.063		Zt09_model_17_00044	17	261587	1.28E-74	0	
436.2114	Artocaprin	Zt09_model_11_00005	11	19156	2.50E-29	0	Permease for cytosine/purines, uracil, thiamine, allantoin
206.1398	Thioctic acid	Zt09_model_9_00007	9	36871	2.49E-19	0	Uncharacterized alpha/beta hydrolase domain (DUF2235)
466.1941	T-2 mycotoxin	Zt09_model_8_00667	8	1983993	1.98E-17	0	Transcription elongation factor Elf1 like
158.1189	Pelargonic acid	Zt09_model_12_00099	12	348419	3.98E-11	0	Dynamin family, Dynamin central region
167.9949	Phosphoenolpyruvic acid	Zt09_model_2_00928	2	2774876	2.74E-08	0	Amidase
295.2274	Butralin	Zt09_model_2_00083	2	320647	5.81E-08	0	Vacuolar protein sorting-associated protein 35
182.0824		Zt09_model_2_00083	2	320647	5.81E-08	0	Vacuolar protein sorting-associated protein 35
211.0877		Zt09_model_2_01158	2	3497499	2.10E-07	0	
205.0985	Phosphocreatine	Zt09_model_9_00314	9	1023909	1.89E-06	0	Enoyl-CoA hydratase/isomerase, Metallopeptidase family M24, Putative cyclase
129.0198	indoleacetic acid	Zt09_model_4_00688	4	2244297	9.25E-07	102	Vacuolar protein sorting 36 NZF-N zinc-finger domain,
136.0671	Pipecolic acid	Zt09_model_4_00192	4	726826	6.15E-07	156	Telomere stability and silencing, Hsp20/alpha crystallin family
387.1934		Zt09_model_10_00138	10	392483	1.22E-07	171	

231.1711	Hydroxytamoxifen	Zt09_model_1_01422	1	4069626	4.90E-15	635	Major Facilitator Superfamily
116.0543	aminophenazone	Zt09_model_4_00599	4	2017113	7.06E-09	692	Flavin-binding monooxygenase-like
232.1553		Zt09_model_13_00009	13	58225	5.51E-16	709	
190.0908	acetylcytisine	Zt09_model_11_00328	11	1013791	9.17E-12	955	linker histone H1 and H5 family
204.1064		Zt09_model_7_00038	7	138324	3.50E-08	975	
176.075	Tryptophan	Zt09_model_8_00674	8	2003171	1.48E-11	1091	Zinc finger, C2H2 type
384.1151	Allantoic acid	Zt09_model_7_00575	7	1864154	8.66E-10	1102	Glycosyl hydrolase family 36 N-terminal domain,
86.0979		Zt09_model_21_00036	21	235144	2.17E-09	1496	
169.0771	Piperazine	Zt09_model_3_00317	3	1146972	3.62E-12	1549	PP-loop family, Zinc-ribbon
268.1051	Methylhistidine	Zt09_model_4_00031	4	113041	2.94E-11	1705	Glycerophosphoryl diester phosphodiesterase family
136.063		Zt09_model_19_00073	19	443326	3.16E-08	1854	

We used metabolite marker signature look-ups in databases to identify putative metabolite structures based on *m*/*z* values and profiles (Table 1). We found seven SNPs within the coding sequences of genes. One of these seven SNPs (*chr8_ 1983993*) falls within a gene encoding a homolog to an enzyme involved in T2 mycotoxin production found in *Fusarium* species (Table 1). The SNP *chr11_ 19156* was located in the coding sequence of a permease related to nucleotide transport. Other SNPs in coding sequences were *chr9_ 36871, chr12_348419, chr2_ 2774876, chr2_ 320647* and *chr9_ 1023909,* which encode a hydrolase domain protein, a dynamin-like protein, an amidase, a vacuolar protein sorting protein and a metallopeptidase, respectively (Table 1). Furthermore, we found SNPs nearby genes encoding a major facilitator super family up-regulated during late infection stages (*i.e.* necrotrophic stage) and encoding a flavin-dependent monooxygenase. We also found SNPs nearby genes encoding hydrolases and histone linker proteins, which maintain chromatin structure (Table 1). We used gene ontology (GO) term enrichment analyses of the protein functions encoded nearby the most significantly associated SNPs for each metabolite marker. We found the strongest enrichment for endosome transport functions (Fig 5D). Endosomes play key roles in cell homeostasis and metabolite regulation.

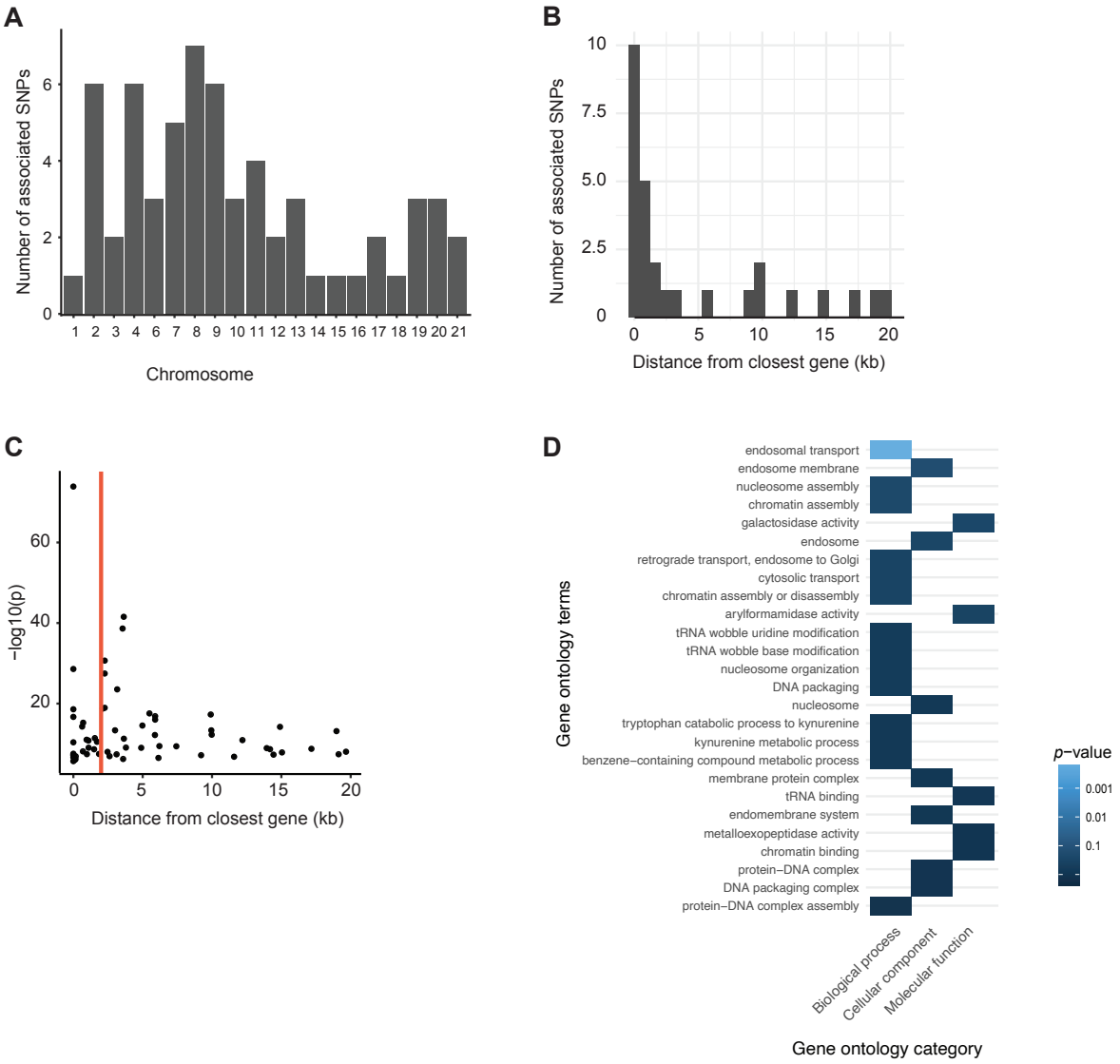


Figure 5: Metabolite genome wide association studies (mGWAS) and candidate gene functions. A) Number of the most significantly associated single nucleotide polymorphisms (SNP) across chromosomes for each of the metabolite markers with a relative intensity >100 (n = 62). B) Distance of the most significantly associated SNP for each metabolite marker to the closest gene. C) −log10(p) values of most significantly associated SNPs compared to the distance of the closest gene. The red line indicates the threshold of 2 kb used for investigations of the encoded gene functions. D) Gene ontology term enrichment analysis of functions encoded by the closest genes.

Discussion

We used untargeted metabolite profile and genome-wide association mapping to identify candidate loci underlying the production of metabolites. We found that a single, highly diverse field population of *Z. tritici* likely harbors substantial variation in the production of individual compounds. We performed association mapping based on standardized metabolite peak intensities as a phenotypic trait and identified significantly associated loci in close proximity to genes encoding membrane transporters and hydrolases. Functions related to endosomes were enriched suggesting an involvement of cellular homeostasis in determining the levels of the quantified metabolites.

The study species harbors a significant diversity in BGC with likely consequences for the production of SM. Consistent with the high degrees of genetic diversity detected at the local scale (Hartmann & Croll, 2017), BGC show presence/absence variation even among isolates from the same field. We found that the local isolates were missing a cluster from every major category of BGC with the exception of the rarer type 3-PKS and terpene types. To what extent the dispensability of individual BGCs relates to ecological functions and niche adaptation remains unknown. An understanding of the metabolites produced by the encoded pathways will help clarify the evolutionary significance of individual BGC polymorphisms. Standing variation for the production of SM elevates the evolutionary potential of *Z. tritici*, because populations could respond rapidly to selection for or against the production of specific metabolites. Intra-species variation in BGCs and SM production have been observed in a range of fungi including the rice blast pathogen *Magnaporthe oryzae*, where a gain of virulence was linked to the duplication of a hybrid PKS-NRPS cluster (Zhong *et al.*, 2020). *Aspergillus* species show also significant intra- and interspecific variation in BGCs and the potential to produce specific SM with consequences for niche adaptation (Theobald *et al.*). Members of the

F. graminearum species complex carries clearly distinct sets of BGCs depending on the evolutionary conservation of the cluster (Tralamazza *et al.*, 2019). In most species, including *Z. tritici*, it remains however unclear whether BGCs presence/absence variation is directly associated to niche and host adaptation.

The production of many metabolites, including SMs, is tightly regulated over the life cycle stages of fungal pathogens (Calvo *et al.*, 2002; Keller, 2019). Using untargeted metabolite screening, we identified highly variable metabolite profiles among isolates of a single field population. Given our approach, the detected metabolites are produced under nutrient-rich culture conditions. To assess metabolite production of *Z. tritici* in more complex environments (*e.g.* on the host) will require targeted approaches and high sensitivity. Focused analyses on identifying individual chemical structures are required to resolve the significance of metabolite peak profiles. Nevertheless, our analyses have established that metabolite production is a highly variable trait within pathogen populations. Given the standardized conditions, we expected significant genetic contributions to the observed metabolome diversity. The analyzed population is genetically highly diverse with a small degree of genetic substructure (Singh *et al.*, 2020) and the metabolome differentiation largely reflects this genetic diversity. Surprisingly, we found that the metabolite profiles clustered according to the isolate collection time point over the growing season (*i.e.* C1-C3). This is in contrast to the genetic differentiation of the same isolates where no clustering is evident based on collection time points (Singh *et al.*, 2020). The metabolome profile clustering is also robust to the relative abundance levels, as a more selective filtering did not meaningfully impact the outcome. A possible explanation for the consistent shifts of the metabolome over the course of the growing season is selection for isolates expressing a specific set of metabolites. Temporal shifts in phenotypic traits were also apparent for pathogen reproduction (*i.e.* pycnidia formation) on wheat leaves where isolates

collected later in the season showed higher reproductive output (Singh *et al.*, 2020, 2021). Shifts in aggressiveness were also previously observed in multi-year studies of *Z. tritici* over years, cultivars and leaves on individual plants (Suffert *et al.*, 2018). Identifying the chemical structure of the compounds contributing most strongly to shifts in the metabolome over the course of the growing season will provide insights into the adaptive nature of such changes.

Variation in metabolite abundance among individuals can be used to identify candidate genes involved in differential regulation. Based on mGWAS, we identified genetic polymorphisms close to genes encoding a broad range of functions. However, we have also encountered substantial *p*-value inflation for association mapping on low intensity peaks highlighting the challenge to improve the signal-to-noise ratio in large metabolomics datasets without manually curated markers. Promising candidate genes encode transporter proteins including a permease related to purine–pyrimidine transport. Such permeases play an important role in nucleic acid biosynthesis and other roles related to cellular nutrition, communication, and defense mechanisms in filamentous fungi (Prather *et al.*, 2005). Major facilitator superfamily (MFS) transporters are known to underlie fungicide and stress tolerance (Roohparvar *et al.*, 2007; Omrane *et al.*, 2017; Zhong *et al.*, 2020). MFS transporters are also involved in the secretion of phytotoxins during infection and, hence, can contribute to virulence (Zwiers *et al.*, 2003; Vela-Corcía *et al.*, 2019). Interestingly we found an association for variation in a putative amino-phenazone, which can act as an anti-bacterial compound (Asiri & Khan, 2010). The associated SNP was close to a gene encoding a flavin-containing mono-oxygenases. In *Aspergillus*, flavin-containing monooxygenases contribute to the biosynthesis of an antimitotic tetrapeptide ustiloxin B (Ye *et al.*, 2016). Ustiloxins are toxic cyclic peptides first identified in the rice pathogen *Ustilaginoidea virens* (Tsukui *et al.*, 2015). Gene ontology enrichment

analyses revealed an over-representation of endosome functions, which are typically involved in cellular homeostasis of metabolites.

In conclusion, our study highlights the power of association mapping to identify candidate loci underlying metabolite diversity in fungi. Previously, mGWAS was limited to plant models, but our findings illustrate how applications to plant pathogenic fungi can efficiently produce candidate lists. A key requirement is access to large, experimentally tractable isolate collections. High degrees of recombination and rapid decay of linkage disequilibrium are further requirements for the successful application of association mapping approaches (Sánchez-Vallet *et al.*, 2018). Establishing metabolome-wide profiles for a range of pathogens under variable conditions will complement analyses focusing on proteaceous effectors.

Material and Methods

Field collection and storage

We collected *Z. tritici* isolates from the Field Phenotyping Platform (FIP) site of the ETH Zürich, Switzerland (Eschikon, coordinates 47.449°N, 8.682°E) (Kirchgessner *et al.*, 2017). We analyzed a total of 102 isolates collected during the 2016 growing season from 11 winter wheat cultivars, which are commonly grown in Switzerland (Levy et al., 2017). We analyzed isolates originating from three collection time points over the season (Supplementary Table S1). The first collection (n = 48) was established in May when wheat plants were in growth stage (GS) 41. The second (n = 7) and third collection (n = 47) was established when the plants were in GS 75 and GS 85, respectively. After sampling, spores of each isolate were stored in either 50% glycerol or anhydrous silica gel at −80 °C. Additional information regarding the sampling scheme and genetic diversity of the collection is described in Singh *et al.* (2020)

Culture preparation and metabolite extraction

Isolates were revived from glycerol stock by adding 50 µl fungal stock solution to a 50 ml conical flask containing 35 ml liquid YSB (yeast-sucrose broth) medium. The inoculated flasks were incubated in the dark at 18° C and 160 rpm on a shaker-incubator. After 8 days of incubation, the cultures were passed through four layers of meshed cheesecloth and washed thrice with sterile milli-Q water to remove media traces. Spores were then lyophilised and metabolites extracted by resuspending the spores (~80 mg) in 1 ml of HPLC-grade methanol. The extract was centrifuged at 15,000 rpm for 5 minutes to pellet down debris before retrieving the supernatant. The last step was repeated until a clear supernatant was recovered.

Whole-genome sequencing and variant calling

Approximately 100 mg of lyophilized spores were used to extract high-quality genomic DNA with the Qiagen DNeasy Plant Mini Kit according to the manufacturer's protocol. We sequenced paired-end reads of 100 bp each with an insert size of ~550 bp on the Illumina HiSeq 4000 platform. Raw reads are available on the NCBI Sequence Read Archive under the BioProject PRJNA596434 (Oggenfuss *et al.*, 2020). Illumina sequences were quality checksed using FastQC v. 0.11.9. (Andrews, 2010). Sequencing reads were then screened for adapter sequences and quality trimmed using Trimmomatic v. 0.39 (Bolger et al., 2014) using the following settings: illuminaclip=TruSeq3-PE.fa:2:30:10, leading=10, trailing=10, sliding-window=5:10 and minlen=50. Trimmed sequencing reads were aligned to the reference genome IPO323 (Goodwin et al., 2011) available from Ensembl Fungi (https://fungi.ensembl.org/Zymoseptoria_tritici/Info/Index) and the mitochondrial genome (European Nucleotide Archive accession EU090238.1) using Bowtie2 v. 2.4.1 (Langmead & Salzberg, 2012). Multi-sample joint variant calling was performed using the HaplotypeCaller and GenotypeGVCF tools of the GATK package v. 4.0.1.2 (McKenna et al., 2010). We retained only SNP variants (excluding indels) and proceeded to hard filtering using the GATK VariantFiltration tool based on the following cutoffs: QD < 5.0; QUAL < 1000.0; MQ < 20.0; -2 > ReadPosRankSum > 2.0; -2 > MQRankSum > 2.0; -2 > BaseQRankSum > 2.0. After filtering for locus level genotyping rate (>50%) and minor allele count (MAC) of 1 using VCFtools v. 0.1.15 (Danecek et al., 2011). Additional sequencing and variant call statistics are available from Singh *et al.* (2000).

Untargeted metabolite profiling using UPLC-Q-TOF-MS

Metabolome analyses were carried out by UHPLC-QTOF-MS using an Acquity UPLC coupled to a Synapt G2 QTOF mass spectrometer (Waters). An Acquity UPLC HSS T3 column

(100x2.1mm, 1.8 µm; Waters) was employed at a flow rate of 500 µl/min and maintained at a temperature of 40°C. The following gradient with 0.05% formic acid in water as mobile phase A and 0.05% formic acid in acetonitrile as mobile phase B was applied: 0-100 % B in 10 min, holding at 100% B for 2.0 min, re-equilibration at 0% B for 3.0 min. The injection volume was 2.5 µl. The QTOF was operated in electrospray negative mode using the so-called MSE acquisition mode which alternates between two acquisitions functions, one at low and another at high fragmentation energies. Mass spectrometric parameters were as follows: mass range 50-1200 Da, scan time 0.2 s, source temperature 120°C, capillary voltage -2.5 kV, cone voltage -25V, desolvation gas flow and temperature 900 L/h and 400°C, respectively, cone gas flow 20 L/h, collision energy 4 eV (low energy acquisition function) or 15-50 eV (high energy acquisition function). A 500 ng/ml solution of the synthetic peptide leucine-enkephaline in water:acetonitrile:formic acid (50:50:0.1) was infused constantly into the mass spectrometer as internal reference to ensure accurate mass measurements (<2ppm). Data was recorded by Masslynx v.4.1. Marker detection was performed using Markerlynx XS (Waters) with the following parameters: initial and final retention time 1.5 and 10.0 min, mass range 85-1200 Da, mass window 0.02 Da, retention time window 0.08 min, intensity threshold 500 counts, automatic peak width and peak-to-peak baseline noise calculation, deisotoping applied. Data was mean-centered and Pareto scaled before applying multivariate analysis. Markers of interest were characterized using the mass spectrometry reference spectra database MassBank (https://massbank.eu). We selected the metabolite hit which was +/- 0.05 m/z value to the query metabolite m/z value

Genome-wide association mapping and linkage disequilibrium analyses

We performed GWAS based on mixed linear models accounting for degrees of kinship among genotypes (MLM+K). The kinship matrix was computed using the scaled identity-by-state

(IBS) algorithm implemented in TASSEL v. 20201114 (Bradbury *et al.*, 2007). We included the kinship matrix as a random effect in the mixed linear models for association mapping using TASSEL. Accounting for kinship performs sufficiently well to control for genetic substructure in the mapping population (Singh *et al.* 2021). Untransformed relative abundance values for each peak were used as trait values for association mapping. Outcomes were visualized using the R package *qqman* v. 0.1.4 (Turner, 2014). We filtered association *p*-values based on the Bonferroni threshold at $\alpha = 0.05$. The closest gene for each associated SNP was identified using the "closest" command in bedtools v. 2.29.0 (Quinlan & Hall, 2010).

Acknowledgements

We would like to thank Ophélie Gning for providing technical support in the metabolite sample preparation and Sabina Tralamazza for helpful feedback on a previous version of the manuscript. The research was funded by a Swiss National Science Foundation grant to DC (number 173265).

References

Andrews S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data.

Asiri AM, Khan SA. 2010. Synthesis and anti-bacterial activities of some novel schiff bases derived from aminophenazone. *Molecules* 15: 6850–6858.

Badet T, Oggenfuss U, Abraham L, McDonald BA, Croll D. 2020. A 19-isolate reference-quality global pangenome for the fungal wheat pathogen Zymoseptoria tritici. *BMC Biology* 18: 12.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30: 2114–20.

Bouws H, Wattenberg A, Zorn H. 2008. Fungal secretomes—nature's toolbox for white biotechnology. *Applied Microbiology and Biotechnology* 80: 381.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23: 2633–2635.

Calvo AM, Wilson RA, Bok JW, Keller NP. 2002. Relationship between Secondary Metabolism and Fungal Development. *Microbiology and Molecular Biology Reviews* 66: 447–459.

Carbone I, Ramirez-Prado JH, Jakobek JL, Horn BW. 2007. Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. *BMC Evolutionary Biology* 7: 111.

Chanclud E, Morel JB. 2016. Plant hormones: a fungal point of view. *Molecular plant pathology* 17: 1289–1297.

Chen W, Gao Y, Xie W, Gong L, Lu K, Wang W, Li Y, Liu X, Zhang H, Dong H, *et al.* 2014. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. *Nature Genetics* 46: 714–721.

Collemare J, Seidl MF. 2019. Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: Is the picture complete? *FEMS Microbiology Reviews* 43: 591–607.

Cornell MJ, Alam I, Soanes DM, Wong HM, Hedeler C, Paton NW, Rattray M, Hubbard SJ, Talbot NJ, Oliver SG. 2007. Comparative genome analysis across a kingdom of eukaryotic organisms: specialization and diversification in the fungi. *Genome research* 17: 1809–22.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Dong X, Gao Y, Chen W, Wang W, Gong L, Liu X, Luo J. 2015. Spatiotemporal Distribution of Phenolamides and the Genetics of Natural Variation of Hydroxycinnamoyl Spermidine in Rice. *Molecular Plant* 8: 111–121.

Dutta A, Hartmann FE, Francisco CS, McDonald BA, Croll D. 2021. Mapping the adaptive landscape of a major agricultural pathogen reveals evolutionary constraints across heterogeneous environments. *The ISME Journal*: 1–18.

Fones H, Gurr S. 2015. The impact of Septoria tritici Blotch disease on wheat: An EU perspective. *Fungal Genetics and Biology* 79: 3–7.

Fox EM, Howlett BJ. 2008. Secondary metabolism: regulation and role in fungal biology. *Current Opinion in Microbiology* 11: 481–487.

Gil-Serna J, Vázquez C, Patiño B. 2020. Genetic regulation of aflatoxin, ochratoxin A, trichothecene, and fumonisin biosynthesis: A review. *International Microbiology* 23: 89–96.

Goodwin SB, M'Barek S Ben, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, van der Lee TAJ, Grimwood J, Aerts A, *et al.* 2011. Finished genome of the fungal wheat pathogen Mycosphaerella graminicola reveals dispensome structure, chromosome plasticity, and stealth pathogenesis (HS Malik, Ed.). *PLoS Genetics* 7: e1002070.

Hartmann FE, Croll D. 2017. Distinct trajectories of massive recent gene gains and losses in populations of a microbial eukaryotic pathogen. *Molecular Biology and Evolution* 34: 2808–2822.

Hartmann FE, Sánchez-Vallet A, McDonald BA, Croll D. 2017. A fungal wheat pathogen evolved host specialization by extensive chromosomal rearrangements. *The ISME Journal* 11: 1189–1204.

Hartmann FE, Vonlanthen T, Singh NK, McDonald M, Milgate A, Croll D. 2020. The complex genomic basis of rapid convergent adaptation to pesticides across continents in a fungal plant pathogen. *Molecular Ecology*: 1–16.

Hoffmeister D, Keller NP. 2007. Natural products of filamentous fungi: Enzymes, genes, and their regulation. *Natural Product Reports* 24: 393–416.

Jacobowitz JR, Weng JK. 2020. Exploring Uncharted Territories of Plant Specialized Metabolism in the Postgenomic Era. *Annual Review of Plant Biology* 71: 631–658.

Jørgensen LN, Hovmøller MS, Hansen JG, Lassen P, Clark B, Bayles R, Rodemann B, Flath K, Jahn M, Goral T, *et al.* 2014. IPM Strategies and Their Dilemmas Including an Introduction to www.eurowheat.org. *Journal of Integrative Agriculture* 13: 265–281.

Keller NP. 2015. Translating biosynthetic gene clusters into fungal armor and weaponry. *Nature Chemical Biology* 11: 671–677.

Keller NP. 2019. Fungal secondary metabolism: regulation, function and drug discovery. *Nature Reviews Microbiology* 17: 167–180.

Keller NP, Turner G, Bennett JW. 2005. Fungal secondary metabolism — from biochemistry to genomics. *Nature Reviews Microbiology* 3: 937–947.

Kirchgessner N, Liebisch F, Yu K, Pfeifer J, Friedli M, Hund A, Walter A. 2017. The ETH field phenotyping platform FIP: A cable-suspended multi-sensor system. *Functional Plant Biology* 44: 154–168.

Krishnan P, Meile L, Plissonneau C, Ma X, Hartmann FE, Croll D, McDonald BA, Sánchez-Vallet A. 2018. Transposable element insertions shape gene regulation and melanin production in a fungal pathogen of wheat. *BMC Biology* 16: 78.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* 9: 357–9.

Lendenmann MH, Croll D, Stewart EL, McDonald BA. 2014. Quantitative trait locus mapping of melanization in the plant pathogenic fungus Zymoseptoria tritici. *G3 (Bethesda, Md.)* 4: 2519–33.

Lind AL, Wisecaver JH, Smith TD, Feng X, Calvo AM, Rokas A. 2015. Examining the Evolution of the Regulatory Circuit Controlling Secondary Metabolism and Development in the Fungal Genus Aspergillus (G Butler, Ed.). *PLOS Genetics* 11: e1005096.

Lysøe E, Seong K-Y, Kistler HC. 2011. The Transcriptome of *Fusarium graminearum* During the Infection of Wheat. *Molecular Plant-Microbe Interactions®* 24: 995–1000.

Marcet-Houben M, Gabaldón T. 2010. Acquisition of prokaryotic genes by fungal genomes. *Trends in genetics : TIG* 26: 5–8.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al.* 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20: 1297–303.

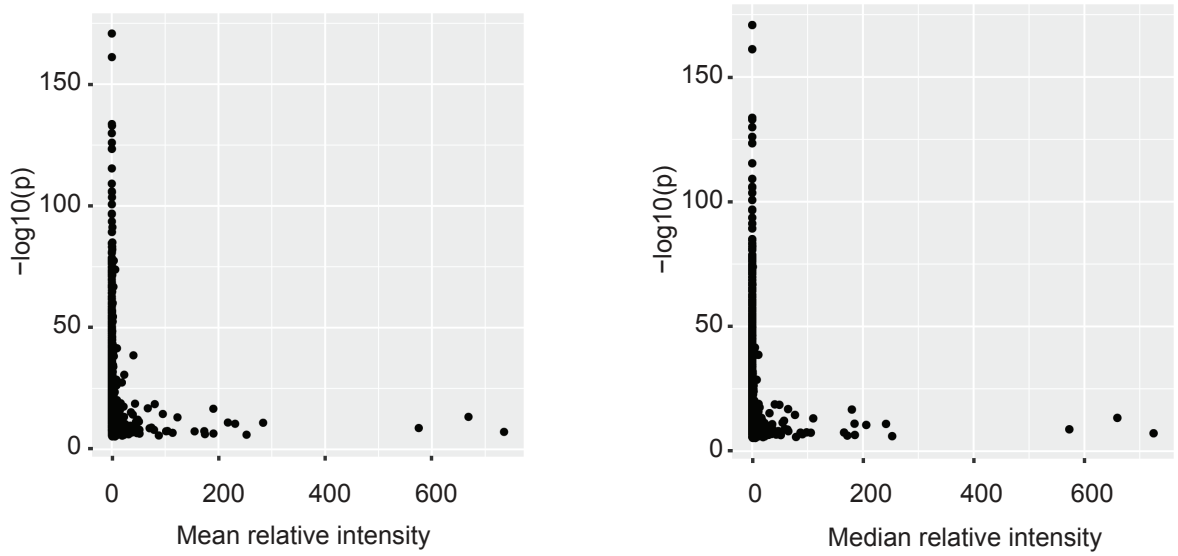
Meier U. 2001. Growth stages of mono-and dicotyledonous plants BBCH Monograph Edited

by Uwe Meier Federal Biological Research Centre for Agriculture and Forestry.
Naranjo-Ortiz MA, Gabaldón T. 2020. Fungal evolution: cellular, genomic and metabolic complexity. *Biological Reviews* 95: brv.12605.
Netzker T, Fischer J, Weber J, Mattern DJ, König CC, Valiante V, Schroeckh V, Brakhage AA. 2015. Microbial communication leading to the activation of silent fungal secondary metabolite gene clusters. *Frontiers in Microbiology* 6: 299.
Oggenfuss U, Badet T, Wicker T, Hartmann FE, Singh NK, Abraham LN, Karisto P, Vonlanthen T, Mundt CC, McDonald BA, *et al.* 2020. A population-level invasion by transposable elements in a fungal pathogen. *bioRxiv*: 2020.02.11.944652.
Omrane S, Audéon C, Ignace A, Duplaix C, Aouini L, Kema G, Walker A-S, Fillinger S. 2017. Plasticity of the MFS1 Promoter Leads to Multidrug Resistance in the Wheat Pathogen Zymoseptoria tritici. *mSphere* 2.
Palma-Guerrero J, Torriani SFF, Zala M, Carter D, Courbot M, Rudd JJ, McDonald BA, Croll D. 2016. Comparative transcriptomic analyses of Zymoseptoria tritici strains show complex lifestyle transitions and intraspecific variability in transcription profiles. *Molecular plant pathology* 17: 845–859.
Palmer JM, Keller NP. 2010. Secondary metabolism in fungi: Does chromosomal location matter? *Current Opinion in Microbiology* 13: 431–436.
Peng M, Shahzad R, Gul A, Subthain H, Shen S, Lei L, Zheng Z, Zhou J, Lu D, Wang S, *et al.* 2017. Differentially evolved glucosyltransferases determine natural variation of rice flavone accumulation and UV-tolerance. *Nature Communications* 8: 1975.
Plissonneau C, Hartmann FE, Croll D. 2018. Pangenome analyses of the wheat pathogen Zymoseptoria tritici reveal the structural basis of a highly plastic eukaryotic genome. *BMC Biology* 16.
Plissonneau C, Stürchler A, Croll D, Taylor JW. 2016. The Evolution of Orphan Regions in Genomes of a Fungal Pathogen of Wheat.
Prather D, Krogan NJ, Emili A, Greenblatt JF, Winston F. 2005. Identification and Characterization of Elf1, a Conserved Transcription Elongation Factor in Saccharomyces cerevisiae. *Molecular and Cellular Biology* 25: 10122–10135.
Proctor RH, Busman M, Seo J-A, Lee YW, Plattner RD. 2008. A fumonisin biosynthetic gene cluster in Fusarium oxysporum strain O-1890 and the genetic basis for B versus C fumonisin production. *Fungal Genetics and Biology* 45: 1016–1026.
Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842.
Roohparvar R, De Waard MA, Kema GHJ, Zwiers L-H. 2007. MgMfs1, a major facilitator superfamily transporter from the fungal wheat pathogen Mycosphaerella graminicola, is a strong protectant against natural toxic compounds and fungicides. *Fungal Genetics and Biology* 44: 378–388.
Rudd JJ, Kanyuka K, Hassani-Pak K, Derbyshire M, Andongabo A, Devonshire J, Lysenko A, Saqi M, Desai NM, Powers SJ, *et al.* 2015. Transcriptome and Metabolite Profiling of the Infection Cycle of *Zymoseptoria tritici* on Wheat Reveals a Biphasic Interaction with Plant Immunity Involving Differential Pathogen Chromosomal Contributions and a Variation on the Hemibiotrophic Lifest. *Plant Physiology* 167: 1158–1185.
Sánchez-Vallet A, Hartmann FE, Marcel TC, Croll D. 2018. Nature's genetic screens: Using genome-wide association studies for effector discovery. *Molecular Plant Pathology* 19: 3–6.
Singh NK, Badet T, Abraham L, Croll D. 2021. Rapid sequence evolution driven by transposable elements at a virulence locus in a fungal wheat pathogen.
Singh NK, Chanclud E, Croll D. 2020. Population-level deep sequencing reveals the interplay of clonal and sexual reproduction in the fungal wheat pathogen Zymoseptoria tritici.

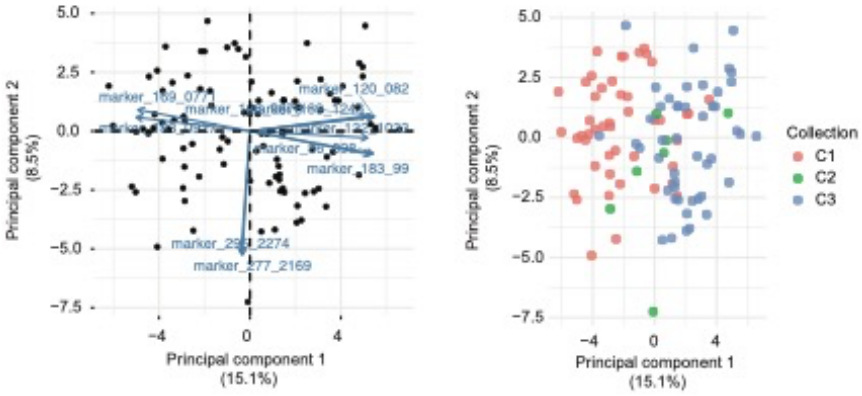
bioRxiv: 2020.07.07.191510.
Spraker JE, Wiemann P, Baccile JA, Venkatesh N, Schumacher J, Schroeder FC, Sanchez LM, Keller NP. 2018. Conserved Responses in a War of Small Molecules between a Plant-Pathogenic Bacterium and Fungi. *mBio* 9.
Steenwyk JL, Mead ME, Knowles SL, Raja HA, Roberts CD, Bader O, Houbraken J, Goldman GH, Oberlies NH, Rokas A. 2020. Variation among biosynthetic gene clusters, secondary metabolite profiles, and cards of virulence across aspergillus species. *Genetics* 216: 481–497.
Strauch RC, Svedin E, Dilkes B, Chapple C, Li X. 2015. Discovery of a novel amino acid racemase through exploration of natural variation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* 112: 11726–11731.
Suffert F, Goyeau H, Sache I, Carpentier F, Gélisse S, Morais D, Delestre G. 2018. Epidemiological trade-off between intra- and interannual scales in the evolution of aggressiveness in a local plant pathogen population. *Evolutionary Applications* 11: 768–780.
Theobald S, Vesth TC, Rendsvig JK, Fog Nielsen K, Riley R, Magalhães De Abreu L, Salamov A, Frisvad JC, Larsen TO, Andersen MR, *et al.* Uncovering secondary metabolite evolution and biosynthesis using gene cluster networks and genetic dereplication OPEN.
Tralamazza SM, Rocha LO, Oggenfuss U, Corrêa B, Croll D. 2019. Complex Evolutionary Origins of Specialized Metabolite Gene Cluster Diversity among the Plant Pathogenic Fungi of the Fusarium graminearum Species Complex (L Rose, Ed.). *Genome Biology and Evolution* 11: 3106–3122.
Tsukui T, Nagano N, Umemura M, Kumagai T, Terai G, Machida M, Asai K. 2015. Ustiloxins, fungal cyclic peptides, are ribosomally synthesized in Ustilaginoidea virens. *Bioinformatics* 31: 981–985.
Turner SD. 2014. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv*: 005165.
Vela-Corcía D, Aditya Srivastava D, Dafa-Berger A, Rotem N, Barda O, Levy M. 2019. MFS transporter from Botrytis cinerea provides tolerance to glucosinolate-breakdown products and is required for pathogenicity. *Nature Communications* 10: 2886.
Vining LC. 1990. Functions of Secondary Metabolites. *Annual Review of Microbiology* 44: 395–427.
Wisecaver JH, Slot JC, Rokas A. 2014. The Evolution of Fungal Metabolic Pathways (JE Stajich, Ed.). *PLoS Genetics* 10: e1004816.
Ye Y, Minami A, Igarashi Y, Izumikawa M, Umemura M, Nagano N, Machida M, Kawahara T, Shin-ya K, Gomi K, *et al.* 2016. Unveiling the Biosynthetic Pathway of the Ribosomally Synthesized and Post-translationally Modified Peptide Ustiloxin B in Filamentous Fungi. *Angewandte Chemie - International Edition* 55: 8072–8075.
Yim G, Huimi Wang H, Davies FRS J. 2007. Antibiotics as signalling molecules. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362: 1195–1200.
Zhong Z, McDonald B, Palma-Guerrero J. 2020. Genetic architecture of oxidative stress tolerance in the fungal wheat pathogen Zymoseptoria tritici. *bioRxiv*.
Zwiers LH, Stergiopoulos I, Gielkens MMC, Goodall SD, De Waard MA. 2003. ABC transporters of the wheat pathogen Mycosphaerella graminicola function as protectants against biotic and xenobiotic toxic compounds. *Molecular Genetics and Genomics* 269: 499–507.

Supplementary Information

Supplementary Figures



Supplementary Figure S1: Distribution of −log10(p) values of the most significantly associated SNP for each of the 2633 metabolite markers against the mean and median relative intensity.



Supplementary Figure S2: Biplot and the first two principal components (PC) from a PC analysis of 62 metabolite markers retained after filtering for a maximum intensity >100.

Chapter 4

Tackling microbial threats in agriculture with integrative imaging and computational approaches

Nikhil Kumar Singh[1], Anik Dutta[1,2], Guido Puccetti[1,3], Daniel Croll[1]

[1] Laboratory of Evolutionary Genetics, Institute of Biology, University of Neuchâtel, CH-2000, Neuchâtel, Switzerland

[2] Plant Pathology, Institute of Integrative Biology, ETH Zurich, CH-8092 Zurich, Switzerland

[3] Syngenta Crop Protection AG, CH-4332 Stein, Switzerland

Published in Computational and Structural Biotechnology Journal: Volume 19 (2021)

https://doi.org/10.1016/j.csbj.2020.12.018

CHAPTER 4

Abstract

Pathogens and pests are one of the major threats to agricultural productivity worldwide. For decades, targeted resistance breeding was used to create crop cultivars that resist pathogens and environmental stress while retaining yields. The often decade-long process of crossing, selection, and field trials to create a new cultivar is challenged by the rapid rise of pathogens overcoming resistance. Similarly, antimicrobial compounds can rapidly lose efficacy due to resistance evolution. Here, we review three major areas where computational, imaging and experimental approaches are revolutionizing the management of pathogen damage on crops. Recognizing and scoring plant diseases have dramatically improved through high-throughput imaging techniques applicable both under well-controlled greenhouse conditions and directly in the field. However, computer vision of complex disease phenotypes will require significant improvements. In parallel, experimental setups similar to high-throughput drug discovery screens make it possible to screen thousands of pathogen strains for variation in resistance and other relevant phenotypic traits. Confocal microscopy and fluorescence can capture rich phenotypic information across pathogen genotypes. Through genome-wide association mapping approaches, phenotypic data helps to unravel the genetic architecture of stress- and virulence-related traits accelerating resistance breeding. Finally, joint, large-scale screenings of trait variation in crops and pathogens can yield fundamental insights into how pathogens face trade-offs in the adaptation to resistant crop varieties. We discuss how future implementations of such innovative approaches in breeding and pathogen screening can lead to more durable disease control.

Keywords: Agriculture, plant pathogens, high-throughput phenotyping, image analysis, genome-wide association mapping, sustainability

Introduction

Feeding the world population requires a stable production of safe food, which is threatened by factors such as climate change, land degradation and diseases. Pathogens and pests cause significant reductions in agricultural productivity worldwide, and control strategies remain ineffective for many pathogens (Strange and Scott 2005; Chakraborty and Newton 2011; Islam et al. 2016). Around 20-30% of the global harvest is lost to plant diseases with additional post-harvest losses of up to 12% (Fao 2017). The two main approaches to protect crops from diseases are the deployment of resistant cultivars often carrying specific loci conferring resistance (i.e. R genes) (Petit-Houdenot and Fudal 2017) and the application of chemical compounds (e.g. fungicides, insecticides) (D.-C. He, Zhan, and Xie 2016; Sundin et al. 2016; Rohr et al. 2017). Targeted resistance breeding has been practiced for decades to create cultivars that resist pathogens but retain desirable agricultural traits. Yet even modern molecular breeding techniques often involve multiple years of crossing, selection, and testing to create a new cultivar (Ahmar et al. 2020). At the same time, plant pathogens often rapidly overcome resistance gene mediated immunity defeating expensive breeding efforts (Palloix, Ayme, and Moury 2009; Mutka et al. 2016). Similarly, the deployment of synthetic chemicals is often followed by the rise of resistant pathogen populations (Kitchen et al. 2016; Van de Wouw et al. 2017). Tackling food security risks posed by pathogens will require bringing together fundamental insights into the biology of diseases, advancements in high-throughput approaches and innovations in computational tools.

Sustainable agricultural management practices critically rely on an understanding of the biology and evolutionary potential of the major crop pathogens. Emerging or re-emerging pathogens can only be contained by investigating their origins and migration routes, as well as their potential to counter-adapt to deployed control measures. Advancements in genomics,

transcriptomics and proteomics have been instrumental to elucidate these questions and unravelled a broad range of molecular mechanisms governing host-pathogen interactions leading to more effective control strategies (Mwadzingeni et al. 2016; Berkman et al. 2012; Andargie and Li 2016; Rustagi et al. 2018). Gene editing and breeding focused on the exploitation of natural genetic variability provide critical resources for introducing novel alleles into crop improvement efforts (Varshney, Hoisington, and Tyagi 2006; Majid et al. 2017). Genetic crop improvement requires though large-scale screening of thousands of lines grown under different environmental conditions (Shakoor, Lee, and Mockler 2017; Bevan et al. 2017). On the pathogen side, the rapid identification of emerging resistance against chemicals or virulence on previously resistant crops is of utmost importance (Corredor-Moreno and Saunders 2020; Islam et al. 2016). Genotyping of plants and pathogens has reached impressive throughput at low cost, yet equivalent improvements in high-throughput screening of phenotypic information are largely lagging behind (Mutka and Bart 2014; Shakoor, Lee, and Mockler 2017). This in turn has created a bottleneck for speed breeding efforts and automated monitoring of the plant health status during agricultural production. To highlight recent developments that have the potential to remedy these shortcomings, we first discuss crucial components of host-pathogen interactions that can be targeted. We then review recent developments in disease-focused plant phenomics enabled by high-throughput phenotyping platforms. Finally, we highlight advances in the screening of pathogen populations to detect early signs of resistance evolution or increased virulence. We argue that combining the above-mentioned approaches has yielded impressive insights into the genetic architecture of crop diseases. Finally, we discuss how innovative approaches can lead to more durable disease control.

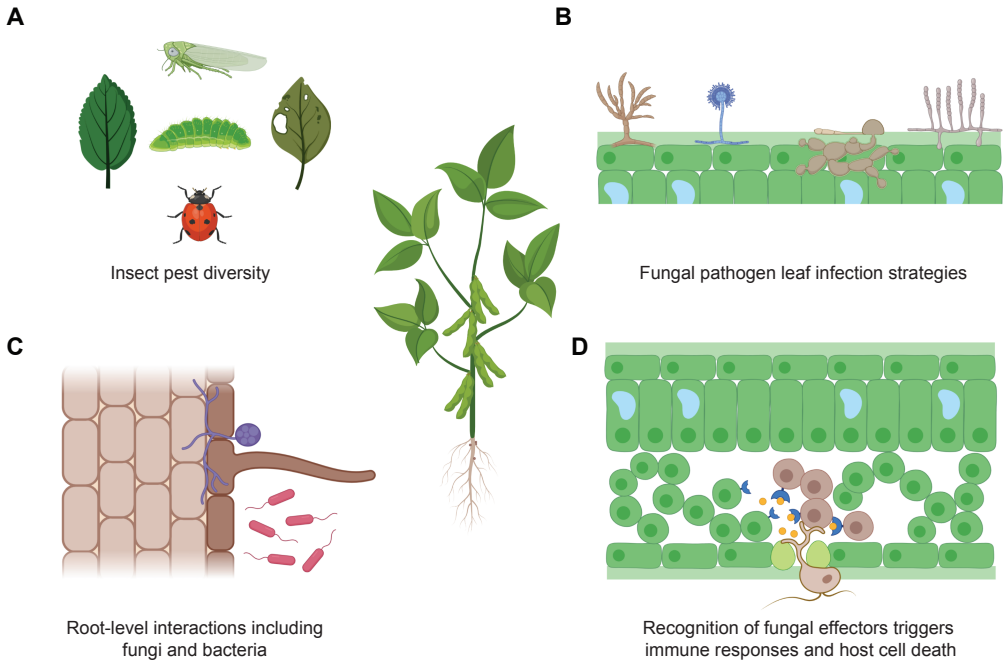


Figure 1: The biology of plant-pathogens interactions. A) A wide range of insects feed on leaves as herbivores. B) Longitudinal root section and rhizosphere. Underground interactions can have significant effects on the stem phenotype and on the entire plant. Filamentous fungi and bacteria can have beneficial effects on the plant by facilitating the acquisition of phosphorus and nitrogen or detrimental effects when confronted by root pathogens. C) Dorsiventral section of a leaf colonized by different pathogenic filamentous fungi. From the left: fungi of the *Fusarium*, *Aspergillus*, *Magnaporthe* and *Podosphaera* genus. D) Filamentous fungi and oomycetes causing leaf infections. Filamentous pathogens can penetrate into the mesophyll through stomata. Once inside the mesophyll, pathogens colonize the entire tissue within days or weeks. Depending on the lifestyle (biotrophic or necrotrophic), pathogens secrete small proteins (called effectors) to manipulate plant cells. Plants detecting effectors (blue receptors) can mount a hypersensitive response (HR) leading to an autophagy-like cell death prevents the spread of the pathogen.

Knowing the enemy: How pathogens interact with plants to cause disease

Major crop diseases are caused by pathogens including fungi, bacteria and viruses (Fig 1). In addition, insect pests also cause substantial losses in agricultural production (Horrocks, Ward, and Suckling 2020; Wilkinson et al. 2019). Spores of fungi and bacteria can be transmitted by wind, rain splash and animals including humans and insects (Leach and Cowen 2014; Perilla-Henao and Casteel 2016). Some pathogens penetrate plant tissues (Fig 1B) with the help of vectors and then are surrounded by cytoplasm, cell membrane, or cell walls (J. D. G. Jones and

Dangl 2006; Abdulkhair and Alghuthaymi 2016). In other cases, the pathogen makes contact with the external surface of the plant, and then deploys a penetration mechanism which can include highly specialized structures (*e.g.* appressoria) (Köhl, Kolnaar, and Ravensberg 2019). After invasion of the host tissue, the success of a pathogen on a specific host is largely explained by an interaction of gene products encoded in the host and pathogen genomes (Crute 1994; Alizon et al. 2009). Resistance (R) genes in plants encode proteins to directly or indirectly detect the presence of pathogen effectors (also known as avirulence factors or Avr) and trigger strong immune responses (Fig 1B-D). Pathogens can escape recognition through sequence diversification or deletion of effector genes (Schmidt et al. 2016). Despite the importance of single loci controlling the infection outcome, pathogenicity can have a complex and largely quantitative genetic basis (Lambrechts 2010; Brown et al. 2015; J. Zhang et al. 2015; Chakravarti and Turner 2016; Hartmann et al. 2017; Bettgenhaeuser et al. 2018). To successfully infect plants, pathogens also need to tolerate a series of abiotic stress factors (Fig 1C). Since pathogen species have an optimal range of environmental conditions (*i.e.* temperature, pH, humidity) to thrive, changes in the environment can alter the pathogen's ability to cause damage (van der Does and Rep 2007). Environmental factors such as annual temperature fluctuations, the quantity and pattern of precipitation, levels of CO_2 and ozone can affect plant disease severity (Chakraborty and Newton 2011; Suffert, Ravigné, and Sachec 2015). Furthermore, the efficacy of pesticides can be significantly altered by weather conditions (*e.g.* wash-off of fungicides following strong rainfall). Hence, the challenge to contain pathogen damage in agriculture is to predict the emergence of virulent strains and the rise of fungicide resistance.

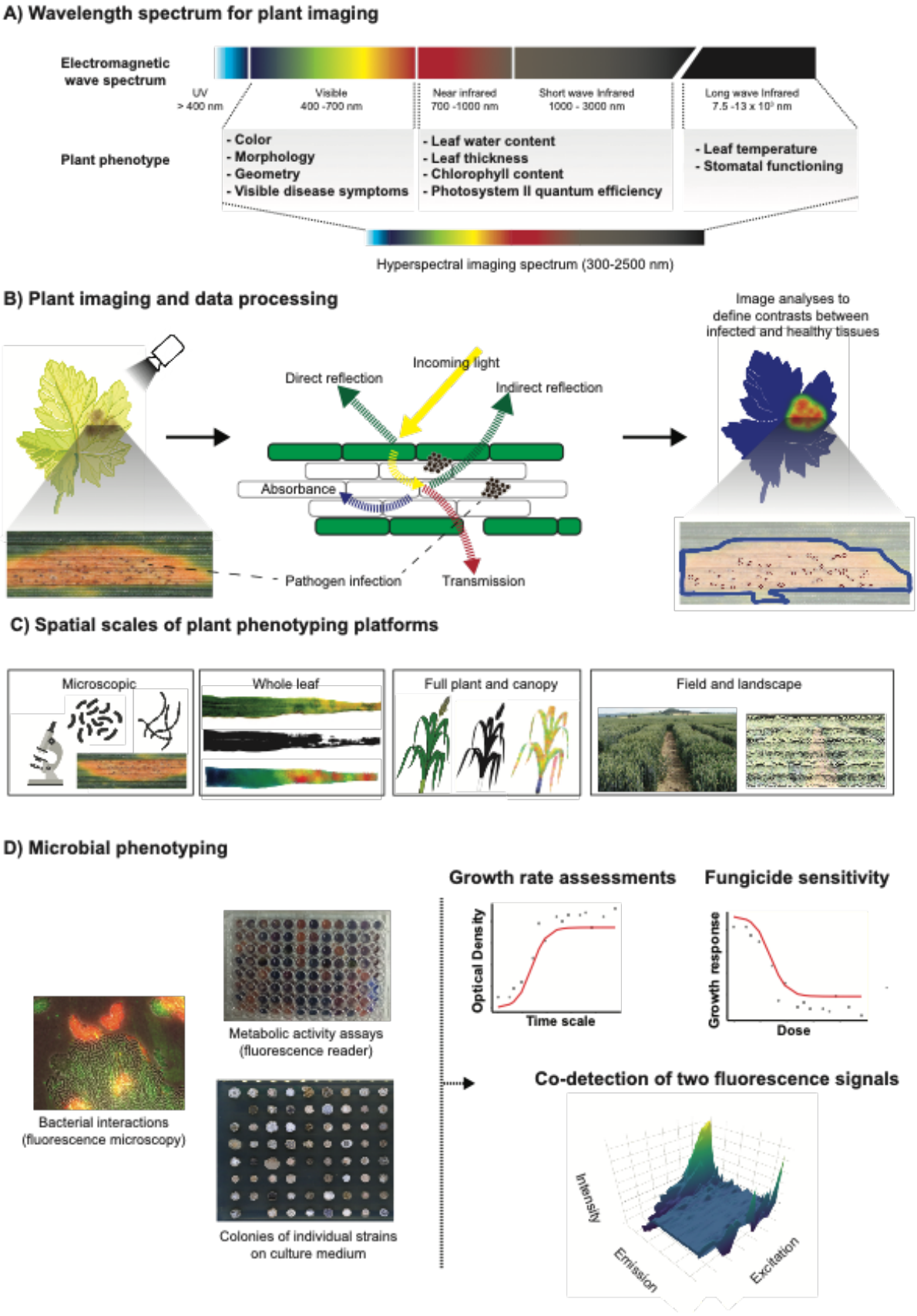


Figure 2: High-throughput phenotyping platforms for plants and pathogens. A) Light in the visible spectrum can be used to detect changes in color and morphology of infected plant tissue. Infrared and short-wave infrared enable to record changes in water content, leaf

thickness and photosynthetic efficiency. Long-wave infrared allows assessments of plant surface temperatures. Hyperspectral sensors capture multiple images across the range of 300-2500 nm. B) Imaging systems assess absorption, transmission, or reflectance characteristics of the incident electromagnetic radiation interacting with the plant surface. Diseased plant tissue often differs in reflectance compared to healthy tissue. Image analysis algorithms define contrasts between diseased and non-diseased leaf areas. C) Disease phenotyping approaches from the leaf to the field scale. D) Screening of pathogen populations can be performed in liquid cultures or on solid media. The most common experiments monitor growth rates by assessing culture densities over time or estimate the dose-response curves when exposed to antimicrobial compounds. Co-cultures of multiple microbes may be analyzed using two distinct emission/excitation pairs specific for each the co-cultured species.

Given the possibly severe consequences, the early detection of resistance breakdowns in crops or loss of sensitivity to pesticides is critical. Detection early in the growing season and identifying previously uncharacterized pathogens remain major challenges. Classic plant disease diagnostics usually relies on visual symptom scoring by trained individuals categorizing disease severity on linear scales (J. Wang et al. 2019). However, classic diagnostics is obviously only possible once symptoms have appeared (Bock et al. 2010). Some pathogens (*i.e.* biotrophs) can feed on plant nutrients for a long period without causing apparent infection symptoms (Spanu and Panstruga 2017). Pathogens killing plant cells to obtain nutrients are more obvious to detect (Shaw et al. 2016). Adding to the complexity of pathogen detection, many pathogens can undergo temporal shifts in infection lifestyles (Vleeshouwers and Oliver 2014). Plant infections often occur in patches with large areas of the field free of disease at an early stage of infestation. This is due to the short-distance dispersal of pathogens around the original disease foci (McCartney and Fitt 1998; Bravo et al. 2003). Hence, efficient, large-scale disease scoring is necessary to detect infections at early stages before the symptoms even become visible. Remote imaging and advanced data analysis can be used to identify disease foci and inform smart applications of chemical control agents (Fig 2). The agricultural environment presents many complexities that are often poorly captured in greenhouse experiments. In fields, plants face fluctuating environmental conditions and are competing for

light, water and nutrients. Hence, high-throughput phenotyping and analytical tools are largely subdivided into field-applicable approaches and laboratory/greenhouse setups.

Disease symptoms may include any range of changes in the color, shape or functioning of the plant as it responds to a pathogen and can be visualized at specific wavelengths (Fig 2A). Depending on the pathogen, the disease symptoms can range from leaf spots, chlorosis, necrosis, wilting or overgrowth (Kado 2016; Jain et al. 2019). However, plant stress beyond infections can activate protective mechanisms leading to suboptimal growth, chlorophyll loss or changes in surface temperatures (Kuska et al. 2015; Baker 2008). Such changes produce detectable shifts in the spectral signature compared to a healthy plant and can be measured by different methods (Fig 2A-B, Table 1) (Martinelli et al. 2015). Imaging systems rely on the quantification of absorption, transmission, or reflectance characteristics of the electromagnetic radiation interacting with the plant surface (Fig 2A-B) (Mulla 2013). We provide a set of specific disease-related phenotypes, which can be captured by different camera sensors across the wavelength spectrum in Table 1. For example, in the shortwave infrared spectrum (1000-3000 nm) water and biomolecules show characteristic reflectance patterns (Fig 2A). A shortwave infrared sensor can detect an infected region based on differences in water content due to the disease (Fig 2B). Computational algorithms can then highlight contrasts between infected and healthy areas to estimate disease severity at the leaf level (Fig 2B). The various algorithms available for plant phenotype assessments have recently been reviewed in more detail (Z. Li et al. 2020). Disease phenotype scoring is also largely dependent on the experimental setup. Plant disease phenotyping under greenhouse or growth chamber conditions largely eliminates fluctuations due to the environment and experiments can easily be replicated for the same genotypes. Such end-point experiments also allow invasive assessments of pathogen colonization on the surface and inside plant tissues using cross-section imaging of

entire organs (*e.g.* leaves, Fig 2C), wide-field and confocal microscopy (Fig 2C), as well as various polymerase chain reaction (PCR) and sequencing-based methods to estimate abundance. Such well-controlled measurements from mostly greenhouse experiments have been used to perform QTL mapping and genome-wide association studies (GWAS) to detect markers associated with resistance and inform breeding programs.

Table 1: Disease assessments enabled by high throughput sensors

Image sensor type	Examples of measurable phenotypes	Disease / pathogens	Reference
RGB	● Color ● Morphology ● Biomass ● Physiology ● Disease symptoms ● Germination rate	● Potato late blight ● Citrus canker ● Cercospora leaf spot ● Sugarbeet rust ● Anthracnose ● Septoria tritici blotch	(Tisné et al. 2013; X. Zhang, Hause, and Borevitz 2012; Campbell et al. 2017; S. L. Kim et al. 2020; Raza et al. 2015; Sugiura et al. 2016; Bock et al. 2008; D. Chen et al. 2014; Neumann et al. 2014; Wijekoon, Goodwin, and Hsiang 2008; Stewart and Mcdonald 2014)
Multispectral and hyperspectral sensors	● Nutrient status, ● Water content ● Senescence ● Photosynthetic efficiency	● Laurel wilt ● Powdery mildew ● Cercospora leaf spot ● Rusts ● Fusarium Head blight	(Abdulridha, Ehsani, and De Castro 2016; L. He et al. 2020; Wahabzada et al. 2015; Leucker et al. 2016; A-K Mahlein et al. 2010; Elke Bauriegel, Giebel, and Herppich 2011; Thomas et al. 2018; Rumpf et al. 2010)
Thermal sensors	● Transpiration ● Heat stress ● Senescence ● Leaf/canopy temperature	● Downy mildew ● Wilt ● Scab disease	(Oerke et al. 2006; Oerke, Fröhling, and Steiner 2011; Calderón, Navas-Cortés, and Zarco-Tejada 2015)
Infrared sensors	● Phenotypes measured with thermal and hyperspectral sensors ● Leaf water content ● Photosynthetic efficiency	● Powdery mildew ● Net blotch ● Brown rust ● Fusarium Head blight	(Kuska et al. 2015; Wahabzada et al. 2015; Apan et al. 2004; E. Bauriegel et al. 2011; Shahin and Symons 2012; Spinelli, Noferini, and Costa 2004)

High-throughput plant health assessment under field conditions

Technologies suitable for screening plant diseases even in early infection stages have become widely deployed (Shakoor, Lee, and Mockler 2017). Phenotyping platforms integrating such technology allow both proximal and remote monitoring of single plant leaves (or other organs), individual plants as well as entire fields (Fig 2C). Capturing plant phenotypes under field conditions consists either of measurements from space or the air with cameras mounted on satellites (WorldView-3, www.digitalglobe.com), unmanned aerial vehicles (UAVs), parachutes, blimps, manned or unmanned (*i.e.* drones) rotocopters or fixed-wing systems (Sankaran et al. 2015; Shakoor, Lee, and Mockler 2017; Chawade et al. 2019). Satellite imaging can provide multi-spectral images with a resolution ranging from meters to hundreds of meters (Chawade et al. 2019). However, there are major limitations due to weather conditions, the frequency of image capture, and overall costs. Satellite imaging is also most useful for regional or continental scale assessments of vegetation cover rather than crop breeding trials in individual fields. Recent improvements in high-resolution satellite imagery, e.g. GF-1 from China or European satellites such as SPOT, can provide time-resolved phenotyping of individual fields at the meter scale (Zhou et al. 2017; M. C. Hansen and Loveland 2012). Such resolution is able to capture transitions in reflectance due to senescence or pathogen outbreaks. Unmanned aerial vehicles (UAV) have become popular in recent years because of the ease of deployment and possibilities to mount high-resolution image capture systems (Sugiura et al. 2016; Sankaran et al. 2015). A range of sensors can be installed on UAVs depending on the payload capacity and type of data required. These sensors can be based on spectral interactions such as visible light imaging (RGB) (Camargo and Smith 2009; Guo, Fukatsu, and Ninomiya 2015; Klodt et al. 2015), infrared and hyperspectral imaging (M. and Kulkarni 2015; Arens et al. 2016; D. Zhang et al. 2018; Thomas et al. 2018), thermal infrared (H. G. Jones et al. 2009; Costa, Grant, and Chaves 2013) and 3D laser (Koenig et al. 2015;

Müller-Linow et al. 2015). A hexacopter mounted with RGB cameras enabled wheat plant height measurements matching well manual and ground-based LiDAR sensor measurements (Madec et al. 2017). Monitoring the progression of late blight disease in a potato field was possible using an aerial RGB camera (Sugiura et al. 2016). Thermal infrared sensors have also been successfully used to quantify plant responses towards drought stress in Black Poplar (Ludovisi et al. 2017), ground cover in sorghum, canopy temperature in sugarcane and crop lodging in wheat (Chapman et al. 2014). A major limitation of aerial camera based phenotyping is errors introduced while generating orthomosaics, which are composite images of stitched and geometrically corrected photographs (Maleki et al. 2020; Chawade et al. 2019). Inaccuracies in orthomosaics can be introduced by low camera sensor resolution, altitude and how the camera positioning was determined during capture (Liu et al. 2018). However, improvements in rectification algorithms and drone technologies have reduced orthomosaic errors significantly (T. Li et al. 2020). Orthomosaic inaccuracies can also be remedied by supplementing imagery captured by proximal phenotyping platforms (Xu, Li, and Paterson 2019; Kefauver et al. 2017).

Proximal plant phenotyping

Proximal phenotyping is mostly deployed in greenhouse experiments or in well-controlled field experiments. Sensors covering a broad range of spectra can be mounted on stationary platforms or, for outdoor use, on suspension cables, robots and tractors (Fahlgren, Gehan, and Baxter 2015; Shakoor, Lee, and Mockler 2017; Mutka and Bart 2014; Young, Kayacan, and Peschel 2019; Walter et al. 2019; Salas Fernandez et al. 2017; Kirchgessner et al. 2017). A fixed framework over a large field has advantages over both vehicle- and UAV-based phenotyping and is not limited by sensor payloads and battery capacity (Yang et al. 2020). For example, the Field Phenotyping Platform (FIP) established at the ETH Zurich uses a suspended cable setup

to move sensors over a large experimental field site (Kirchgessner et al. 2017). Similar setups exist at the University of Nebraska to phenotype maize and soybean cultures (Bai et al. 2019). These proximal phenotyping platforms benefit from having multiple mounted sensors capturing phenotypic trait information over a wide range of wavelengths. Imaging based on fluorescence is often used for proximal phenotyping to quantify phenotypic changes related to photosynthesis, nitrogen content and diseases like Septoria tritici blotch (Fig 2A-B) (Odilbekov et al. 2019; McAusland et al. 2019; J. Wang et al. 2019; S. L. Kim et al. 2020). Proximal phenotyping platforms can measure canopy temperature, nitrogen content, as well as phenotypic traits including leaf area and plant height (Yang et al. 2014). Such efficient, whole-plant phenotyping helped accelerate the selection of drought-resistant rice plants (S. L. Kim et al. 2020). The platform was designed for imaging in the visible spectrum to record morphological features. The system also included infrared and near-infrared imaging to quantify water content, as well as temperature and fluorescence measures to quantify photosynthesis efficiency (S. L. Kim et al. 2020). The platform could accurately distinguish tolerant and susceptible plants and thus enabling rapid selection for speed breeding. Downstream applications included functional genetic studies for drought tolerance and GWAS applications, which in turn can feed into marker-assisted breeding efforts. In a different application, chlorophyll fluorescence and thermal (infrared) phenotyping of ~300 Iranian wheat cultivars combined with GWAS revealed adaptive alleles for drought stress useful for marker-assisted selection (Rahimi et al. 2019).

Current plant phenotyping technologies and image processing algorithms struggle to reliably differentiate disease symptoms originating from multiple or unknown pathogens (Anne-Katrin Mahlein et al. 2012; Rumpf et al. 2010). This is partially due to the complexity of fungal and bacterial disease symptoms on crops, but is also due to the fact that disease symptoms during

early infection stages are hardly diagnostic for specific pathogens. Nutrient or water deficiencies (in the absence of pathogens) can produce symptoms that are difficult to differentiate from disease phenotypes. Molecular detection methods (*i.e.* qPCR) are often required to clearly establish what is likely to cause a disease. Interesting developments to overcome these limitations are *e.g.* the differential reflectance spectra (400–1050 nm) of sugar beet leaves, which helps distinguish three different fungal pathogens including *Cercospora beticola*, *Erysiphe betae*, and *Uromyces betae* (A-K Mahlein et al. 2010). Combined thermal and visible light image data was fed to a machine-learning algorithm to differentiate infections by the tomato powdery mildew fungus *Oidium neolycopersici (Raza et al. 2015)*. One major limitation of plant disease high-throughput phenotyping is the timeframe of disease detection. Imaging which relies on reduced photosynthesis as a proxy for disease progression can only detect symptoms once the pathogen invades the plant tissue and reduces chlorophyll activity (Fig. 1, 2). Hyperspectral or multispectral sensors can distinguish diseased leaves from healthy ones much earlier than imaging in the visible spectrum only. Abdulridha et al. (Abdulridha et al. 2020) were able to detect and distinguish the onset of target spot caused by the fungus *Corynespora cassicola* and bacterial spot caused by *Xanthomonas perforans* on tomato in the asymptomatic phase. The multi-spectral imaging approach was successful both under laboratory and field conditions. Such robust phenotyping methods could make it possible to detect infection foci in the field. Emerging infections could then be isolated and treated individually by the deployment of specific fungicides.

Plant phenotyping technologies have generated data to populate public image databases (summarized in (Cobb et al. 2013)). Such data can be used for machine learning and shows promising results in predicting disease onsets and identifying the pathogen species (Raza et al. 2015; Ashourloo et al. 2016; Mohanty, Hughes, and Salathé 2016). Using a public dataset of

54,306 images of healthy and diseased plant leaves, a deep learning network could identify 26 different diseases with 99.35% accuracy (Mohanty, Hughes, and Salathé 2016). The advancement of machine learning applied to the problem of plant phenotyping has been limited mostly by two factors: i) the limitation of publicly available image databases of crop disease phenotypes under a variety of environmental conditions and ii) poor annotation of image data and the captured symptoms (Jayme G. A. Barbedo 2018; Arsenovic et al. 2019). The lack of robust training datasets has forced researchers to use suboptimal databases leading to ambiguity in distinguishing disease phenotypes from senescence or other environmental factors (Jayme Garcia Arnal Barbedo 2016; Jayme G. A. Barbedo 2018; Mohanty, Hughes, and Salathé 2016). Thus, despite the enormous potential of machine learning for plant disease scoring, the current implementations will need to be improved with large reference datasets.

Pathogen screenings to improve the durability of plant resistance

Plant-microbe interactions are often subject to rapid co-evolutionary dynamics altering the genetic make-up of both the host and the pathogen (Frantzeskakis et al. 2020a). Evolving pathogens make breeding for resistance obviously challenging. In addition, outbreaks of new pathogens have significantly increased (Fisher et al. 2018) including *e.g.* recent outbreaks of *Magnaporthe oryzae* and *Puccinia graminis* f. sp. *tritici* in Bangladesh and Italy, respectively, endanger wheat production (Bhattacharya 2017; Inoue et al. 2017; Fisher et al. 2018; Torriani et al. 2015). Rapid change in crop pathogen populations is most evident through the rise of fungicide resistance (McDonald and Stukenbrock 2016; Blake et al. 2018) but many other phenotypes show high variability as well (Möller et al. 2018; Schröter and Dersch 2019). Some pathogens gained virulence traits by acquiring genetic material through horizontal transfer. A striking example includes the gain of virulence on wheat by transferring a key toxin-encoding gene between fungal species (Friesen et al. 2006). Among the most diverse pathogens is the

cosmopolitan fungus *Zymoseptoria tritici* attacking wheat. Cultivar resistance has been largely broken down and some new virulence arose within the span of a few years (Cowger, Hoffer, and Mundt 2000; Brown et al. 2015). *Z. tritici* populations are genetically and phenotypically highly diverse with intra-field diversity approaching levels of diversity at the continental or global scale (Croll and McDonald 2017; Zhan, Pettway, and McDonald 2003; Singh, Chanclud, and Croll, n.d.). To counter highly diverse and rapidly evolving pathogens, breeding programs often attempt to prevent resistance breakdowns by combining ('pyramiding') several resistance genes (Saintenac et al. 2018; Brown et al. 2015).

Despite the challenges associated with rapidly evolving pathogen species, heritable trait variation can be exploited to identify mechanisms underlying virulence (Hartmann et al. 2017; Martin et al. 2020). A number of major crop pathogens are readily culturable under sterile laboratory conditions allowing reproducible assessments of trait expression. The clonal propagation in sterile medium allows the efficient replication of phenotypes expressed by different genotypes. A major application of high-throughput pathogen screenings are measurements of fungicide sensitivity, which is relevant for the early detection of resistance mutations at the regional or continental scale (Berman and Krysan 2020). More broadly, stress response assessments of fungal and bacterial pathogens can provide important clues about the expression of virulence factors, multi-drug resistance, biofilm formation, and antimicrobial resistance (Saintenac et al. 2018; Nagar et al. 2016). This is because stress induced in *in vitro* setups can share similarities with infection stress conditions. During infection stages, pathogens have to cope with various host defense responses including nutrient deprivation, pH variation, *etc*. Hence, *in vitro* conditions can be useful triggers to express virulence-related proteins. Beyond measuring growth rates as a proxy for the physiological state, phenotypic screens can be extended to measurements of spore shape heterogeneity (Holland et al. 2014),

cell viability, as the percentage live cells, or cell vitality, defined as the physiological capabilities of a cell (Kwolek-Mirek and Zadrag-Tecza 2014; Guyot et al. 2015) and can be assessed with flow cytometry and fluorescence readouts. More detailed analyses focus on variation in the infection life cycle on the host (Haueisen et al. 2019). Finally, assessing temperature adaptation can provide important clues about the adaptive potential and the ability of pathogens to cope with future climates (Haueisen et al. 2019; F. Chen et al. 2017). However, manual handling procedures will need to be replaced by robotization to improve throughput.

High-throughput phenotyping platforms for microbial organisms have dramatically advanced in automating cell culturing, liquid handling, spotting on culture medium as dense arrays and integration of fluorescence measurements and microscopy (Levin-Reisman, Fridman, and Balaban 2014; Cox et al. 2009; Cairns et al. 2015; Fernandes et al. 2009). Such technologies were often first applied for active compound screens in drug discovery programs of the pharmaceutical industry (Szymański, Markowicz, and Mikiciuk-Olasik 2012). In the last years, several pipelines have been developed for high-throughput *in vitro* screening to investigate drug tolerance genes and to elucidate aspects of pathogenicity (Cairns et al. 2015; Fernandes et al. 2009). High-throughput pipelines benefit from robotics, such as Rotor-HDA, which can duplicate and back up large strain libraries in solid or liquid media in the format of 96-6144 isolates per plate. Growth and fungicide susceptibility can be assessed using optical density (OD), colony size, or the fluorescence of liquid cultures (Figure 1D) (Monteiro et al. 2012). Moreover, culture colors are powerful proxies for the production of melanin or other secondary metabolites (Asker 2018). Other high-throughput platforms incorporate multiplexed microfluidic cell culture, automated programmable fluid handling for cell perturbation, quantitative time-lapse microscopy, and computational analysis of time-lapse movies (A. S. Hansen, Hao, and O'Shea 2015). Such platforms are most useful for perturbation experiments

(*e.g.* osmotic shock or exposure to a drug) with a fluorescent readout sensitive to changes in gene expression or subcellular localization. The main goal of the technology is to accelerate drug discovery by screening large antimicrobial compound libraries at a rate of thousands of compounds per week. Efforts are as well made in the development of high-throughput approaches to generate mutated versions of drug targets (Cairns et al. 2015). High-throughput screening methods are further used to characterize metabolic, pharmacokinetic and toxicological data about new drugs and reduce the costs of antimicrobial compound development (Fernandes et al. 2009; Clark et al. 2010). During the development of a new pesticide, field trials can be extremely costly. Therefore, high-throughput *in vitro* screening of diverse pathogen population scans is very informative about possible standing resistance, which can prevent rapid efficacy failures under field conditions (Steinhauer et al. 2019).

The wealth of information on phenotypic trait variation in pathogen species combined with low-cost sequencing can be exploited for GWAS (Sánchez-Vallet et al. 2018). Since the debut of the technique in 2005, GWAS revealed hundreds of thousands of single nucleotide polymorphisms (SNPs) and structural variants associated with thousands of different phenotypes mainly focused on humans (Sánchez-Vallet et al. 2018; Aranzana et al. 2005; Hindorff et al. 2009). Plant pathogens populations often harbor both virulent and avirulent strains. Hence, with reliable and robust phenotyping, GWAS analyses can identify genes underlying specific gains in virulence. Mapping populations of ~100 strains were sufficient to identify key virulence factors (Zhong et al. 2017; Mohd-Assaad, McDonald, and Croll 2019; Richards et al. 2019) and fungicide resistance loci in fungal pathogens (Zhong et al. 2017; Mohd-Assaad, McDonald, and Croll 2019; Steinhauer et al. 2019; Hartmann et al., n.d.; Pereira et al. 2020). Genetic mapping studies can also inform the development process of new fungicides by the early identification of resistance “hotspot” genes (Steinhauer et al. 2019). A

critical factor for successful GWAS applications is the availability of high-quality genomic resources representing the genetic diversity of the pathogen species. Recent pangenome analyses of plant pathogens have laid the foundation for such work (Badet and Croll 2020).

Deciphering complex host-pathogen interactions – a way forward

The application of imaging combined with computational techniques enabled enormous progress in breeding resistant crops and detecting the emergence of new pathogen threats. To reduce complexity, most studies until now focused either on variation on the plant or the pathogen side. However, the nature of host-pathogen interactions can vary across space and time (Frantzeskakis et al. 2020b; de Vries, Stukenbrock, and Rose 2020). In addition to such genotype-by-genotype interactions, environmental conditions (humidity, temperature, competing microbes, *etc.*) influence the outcome of infections. The complexity arises in part from complex pathogen infection cycles starting from initial host contact to transmission to new hosts (Lievens et al. 2018; Hall, Bento, and Ebert 2017). Each infection stage is likely governed, to some extent, by distinct genes. Improving our abilities to predict infection outcomes across many different plant-by-pathogen interactions and environments is therefore a critical step in improving crop resistance and controlling disease outbreaks in the long term.

Integration of different *omics* datasets to build biological networks (*e.g.* gene co-expression or protein-protein interaction networks) has become a powerful approach to unravel genetic factors controlling biological interactions (Peyraud et al. 2017; Mishra, Kumar, and Mukhtar 2019; Castro-Moretti et al. 2020). Beyond this, innovative applications of GWAS can help to identify causal genes by studying host and pathogen genotypes simultaneously in an infection matrix (M. Wang et al. 2018; MacPherson, Otto, and Nuismer 2018) (Fig 3). Applying such an approach to the *Arabidopsis thaliana* - *Xanthomonas arboricola* interaction, Wang et al. (M.

Wang et al. 2018) identified specific genes for quantitative disease resistance in the host that are effective only against a specific set of pathogen strains. Studying crop diseases as an infection matrix can help to identify genomic regions underlying host-pathogen co-evolution and genes responsible for specific phenotypes (Zaidem, Groen, and Purugganan 2019; Beilsmith et al. 2019; Märkle and Tellier 2020). However, applications in the agricultural context are largely missing for now as such approaches are experimentally demanding and costly. Improved pathogen inoculation techniques such as detached leaf assays (Mulema and Denby 2012; El-Mor et al. 2018; Barbacci et al. 2020) or head assays (Goddard et al. 2020) combined with automated image analysis (Karisto et al. 2018) will reduce the experimental burden. Consequently, collecting precise phenotypic data from a large infection matrix of hosts and pathogens has become increasingly accessible (Mulema and Denby 2012). Beyond increased sample sizes of the analyzed host and pathogen genotypes, special care is required to adequately cover host and pathogen genetic diversity. Using a global set of populations of the major wheat pathogen *Z. tritici*, a large set of loci associated with pathogen virulence and reproduction on different hosts were identified (Hartmann et al. 2017; A. Dutta et al. 2020). A large infection matrix of 98 strains of *Botrytis cinerea* and 90 plant genotypes of eight species revealed a highly polygenic architecture of pathogen virulence and host specialization (Caseys et al. 2020). Hence, simultaneously extending genetic diversity screens of the host and pathogen provides a powerful approach to comprehensively map the genetic architecture of virulence and resistance. Particularly relevant pathogens to investigate are the ones capable of infecting multiple plant organs (*e.g.* blackleg in canola) or at different growth stages. Bigger sample sizes in both host and pathogen will also improve heritability estimates (*e.g.* as shown in human studies (H. Kim et al. 2017)). Overall, expanding the breadth of host and pathogen genetic diversity in screenings will help to identify previously unknown resistances/susceptibilities in crop germplasm.

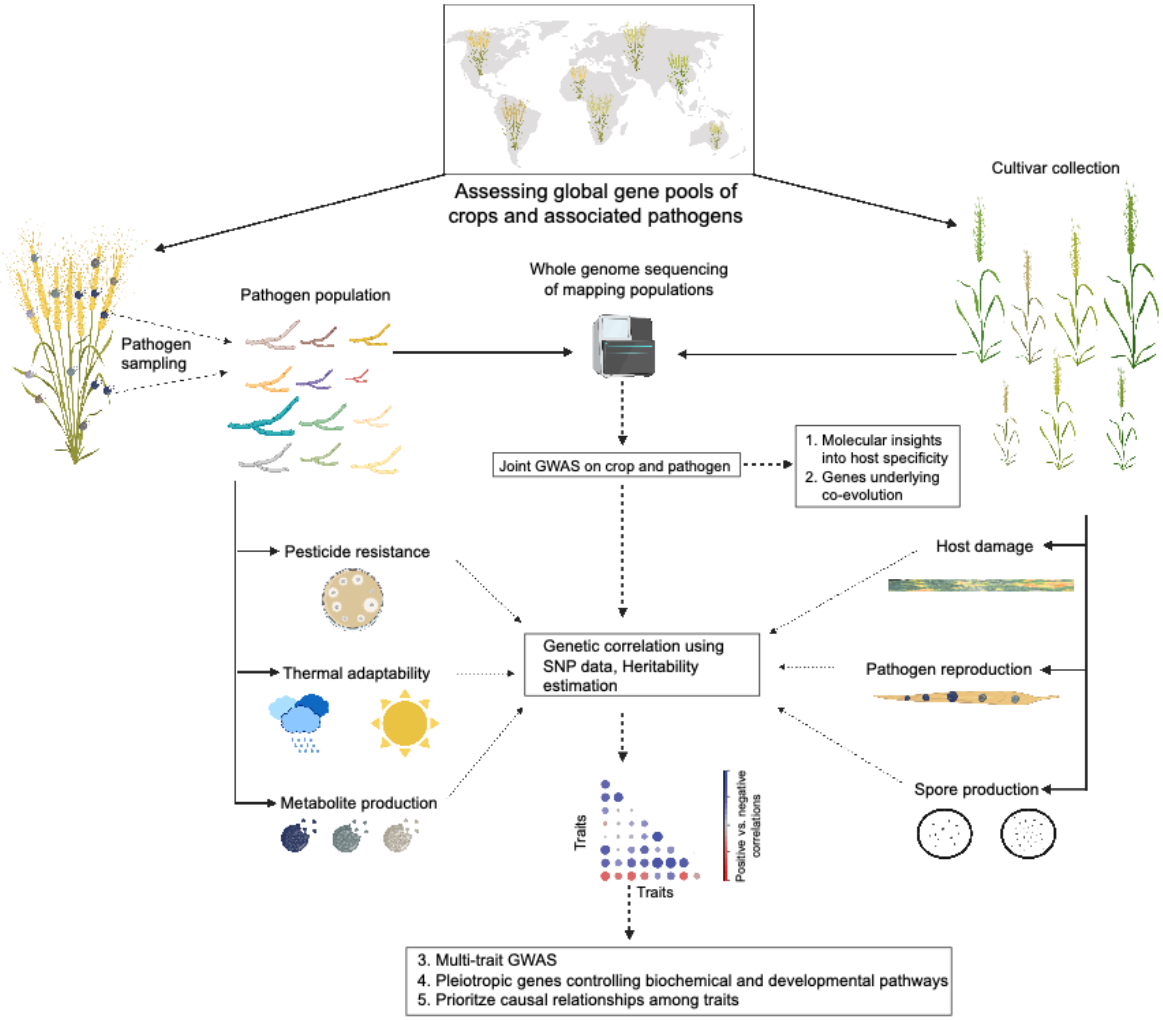


Figure 3: A comprehensive framework for determining the genetic basis of crop-pathogen interactions. Genetically diverse pathogen populations and crop cultivars from different geographies form the basis of the screening. Genome sequencing enables to conduct joint genome-wide association studies (GWAS) to determine the genetic architecture of virulence (in pathogens) and resistance (in crops). Global populations of both pathogen and crop will capture most relevant genetic variation. Beyond virulence, pathogen populations can be screened for loci underlying pesticide resistance, thermal adaptation and metabolite production (*e.g.* melanin). Identifying genetic correlations among pathogen traits facilitates the identification of pleiotropic genes governing trade-offs. Some illustrations were provided by biorender.com according to their usage conditions.

Exploiting pathogen weakness driven by trade-offs

How successful pathogens cope in diverse environments depends on life-history trade-offs, which arise from resource allocation dilemmas and antagonistic gene actions. A trade-off

indicates that an increase in one trait is associated with a decrease in another trait. Such trade-offs are typically dependent on the host genetic background and abiotic conditions. Some pathogen strains have evolved specialization on certain hosts or climatic conditions to maximize their performance. Classic examples include the wild pathosystem of *Plantago lanceolata-Podosphaera plantaginis* interactions (A-L Laine 2007; Anna-Liisa Laine 2008). Trade-offs are likely a key factor maintaining polymorphisms in pathogen populations (Agrawal 2020; Anik Dutta et al. 2020). Studies have already identified trade-offs between foliar damage and asexual transmission (Pasco et al. 2016), sporangia size and number (Mariette et al. 2018), latency in spore production, size and sporulation rates (Delmas et al. 2016). Analyses of demethylation inhibitors (DMI) fungicide-sensitive and resistant isolates of the sugar beet pathogen *C. beticola* have shown that resistant isolates have significantly lower virulence and spore production than sensitive isolates (Karaoglanidis, Thanassoulopoulos, and Ioannidis 2001). No differences were found for incubation periods, mycelial growth, germination of conidia and germ tube length. Similarly, in the bacterial pathogen *Ralstonia solanacearum* mutants lacking the gene for synthesizing an exopolysaccharide virulence factor show increased growth rates compared to the wild-type strain (Peeters et al. 2013; Peyraud et al. 2016). Besides, agricultural pathogens likely face many additional, yet unknown trade-offs because pathogens must also survive outside of annual crop hosts (Fones et al. 2020; Velásquez, Castroverde, and He 2018). The genetic basis of trade-offs remains largely unknown hindering the exploitation of inherent pathogen weaknesses. This is partly because trade-offs are generally expressed by phenotypic trait correlations, which are often confounded by environmental variation and genetic substructure. Identifying genetic correlations among traits can circumvent the above challenges as it reflects the direct effect of genetic factors controlling trade-offs and are robust to confounding factors (van Rheenen et al. 2019). A recent study investigated genetic trade-offs based on genetic correlations in the wheat fungal pathogen

Z. tritici. Performance of a global strain collection on twelve wheat varieties and in various abiotic conditions revealed a broad pleiotropic control of pathogen performance on and off the host (A. Dutta et al. 2020). Weaknesses of pathogens in specific environments will inform more efficient designs of disease control and prevent future resistance breakdowns. Hence, comprehensive maps of genetic trade-offs will possibly enable innovative disease control strategies (Fig 3). Beyond revealing trade-offs, correlated traits can be combined to perform multi-trait GWAS to pinpoint pleiotropic genes and determine causality among traits in specific environments (Porter and O'Reilly 2017; Chhetri et al. 2019) (Fig 3).

Summary and outlook

Technological progress in assessing susceptibility of large collections of crop plants to pathogen damage is crucial for modern resistance breeding efforts. A variety of image capture techniques allow to monitor plant damage at the cellular, leaf, whole plant or field level. Most applications focus on the visible spectrum but hyperspectral imaging platforms have recently gained the ability to detect pathogen infestation even before the appearance of symptoms. A major area of going research is to improve image analyses algorithms to detect and classify pathogen damage. Variation within individual pathogen species can be highly informative about the rise of new virulence or pesticide resistance. Robotics applied to automate the culturing of thousands of pathogen strains enables screening for metabolic variation, drug susceptibility and production of secondary metabolites improving our understanding how pathogens cope with the agricultural environment. Both high-throughput plant and pathogen phenotyping efforts can be combined with genome sequencing and GWAS applications. Unraveling the genetic basis of host resistance helps to speed up breeding efforts through marker-assisted selection. Analysis of pathogen populations can be informative about possible trade-offs in the emergence of virulence or pesticide resistance.

Future directions of research should focus on a set of complementary research areas.

- Create efficient pipelines merging imaging and molecular assays for pathogen detection. Such integrated systems could help farmers deploy appropriate counter-measures in the field prior to widespread damage and reduce overall pesticide application.
- The susceptibility of crop cultivars to major pathogens should be re-assessed continuously to detect changes in the virulence profile of the prevalent pathogens. Rapid evolution in pathogens can lead to catastrophic resistance breakdowns and must be detected early enough. High-throughput imaging systems capturing disease symptoms should be combined with machine learning to effectively recognize changes in virulence profiles. The lack of curated and open access disease image databases is currently slowing progress.
- Regional monitoring efforts of resistance breakdowns or the loss of pesticide efficacy can be achieved by high-throughput genomic screening of infected leaf material. Efficient genotyping assays focusing on major genes are likely to scale well to broad applications. Bioinformatic procedures for such genomic data analyses are largely in place. A successful implementation of such genomic monitoring will also help to detect the arrival of new pathogens early enough to deploy resistant cultivars or implement changes in pesticide application regimes.
- The systematic identification of trade-offs faced by pathogens adapting to pesticides and resistant crop cultivars could lead to more durable control measures.

Acknowledgements

We are grateful to Luzia Stalder for help with the design of a figure. NKS, AD and GP are supported by funding from the Swiss National Science Foundation (grant 173265), the Federal Office for Agriculture (NAP-PGREL Project Nr. 627000640) and Innosuisse with grants to DC.

References

Abdulkhair, Waleed M., and Mousa A. Alghuthaymi. 2016. “Plant Pathogens.” In *Plant Growth*, edited by Everlon Cid Rigobelo. InTech. https://doi.org/10.5772/65325.

Abdulridha, Jaafar, Yiannis Ampatzidis, Sri Charan Kakarla, and Pamela Roberts. 2020. “Detection of Target Spot and Bacterial Spot Diseases in Tomato Using UAV-Based and Benchtop-Based Hyperspectral Imaging Techniques.” *Precision Agriculture* 21 (5): 955–78. https://doi.org/10.1007/s11119-019-09703-4.

Abdulridha, Jaafar, Reza Ehsani, and Ana De Castro. 2016. “Detection and Differentiation between Laurel Wilt Disease, Phytophthora Disease, and Salinity Damage Using a Hyperspectral Sensing Technique.” *Collection FAO: Agriculture* 6 (4): 56. https://doi.org/10.3390/agriculture6040056.

Agrawal, Anurag A. 2020. “A Scale-dependent Framework for Trade-offs, Syndromes, and Specialization in Organismal Biology.” *Ecology*. https://doi.org/10.1002/ecy.2924.

Ahmar, Sunny, Rafaqat Ali Gill, Ki-Hong Jung, Aroosha Faheem, Muhammad Uzair Qasim, Mustansar Mubeen, and Weijun Zhou. 2020. “Conventional and Molecular Techniques from Simple Breeding to Speed Breeding in Crop Plants: Recent Advances and Future Outlook.” *International Journal of Molecular Sciences* 21 (7). https://doi.org/10.3390/ijms21072590.

Alizon, S., A. Hurford, N. Mideo, and M. Van Baalen. 2009. “Virulence Evolution and the Trade-off Hypothesis: History, Current State of Affairs and the Future.” *Journal of Evolutionary Biology*. https://doi.org/10.1111/j.1420-9101.2008.01658.x.

Andargie, Mebeaselassie, and Jianxiong Li. 2016. “Arabidopsis Thaliana: A Model Host Plant to Study Plant–Pathogen Interaction Using Rice False Smut Isolates of Ustilaginoidea Virens.” *Frontiers in Plant Science* 7. https://doi.org/10.3389/fpls.2016.00192.

Apan, Armando, Alex Held, Stuart Phinn, and John Markley. 2004. “Detecting Sugarcane ‘orange Rust’disease Using EO-1 Hyperion Hyperspectral Imagery.” *International Journal of Remote Sensing* 25 (2): 489–98. https://www.tandfonline.com/doi/abs/10.1080/01431160310001618031.

Aranzana, María José, Sung Kim, Keyan Zhao, Erica Bakker, Matthew Horton, Katrin Jakob, Clare Lister, et al. 2005. “Genome-Wide Association Mapping in Arabidopsis Identifies Previously Known Flowering Time and Pathogen Resistance Genes.” *PLoS Genetics* 1 (5): e60. https://doi.org/10.1371/journal.pgen.0010060.

Arens, Nadja, Andreas Backhaus, Stefanie Döll, Sandra Fischer, Udo Seiffert, and Hans-Peter Mock. 2016. “Non-Invasive Presymptomatic Detection of Cercospora Beticola Infection and Identification of Early Metabolic Responses in Sugar Beet.” *Frontiers in Plant Science* 07 (September): 1–14. https://doi.org/10.3389/fpls.2016.01377.

Arsenovic, Marko, Mirjana Karanovic, Srdjan Sladojevic, Andras Anderla, and Darko Stefanovic. 2019. “Solving Current Limitations of Deep Learning Based Approaches for Plant Disease Detection.” *Symmetry* 11 (7): 939. https://doi.org/10.3390/sym11070939.

Ashourloo, Davoud, Hossein Aghighi, Ali Akbar Matkan, Mohammad Reza Mobasheri, and Amir Moeini Rad. 2016. “An Investigation Into Machine Learning Regression Techniques for the Leaf Rust Disease Detection Using Hyperspectral Measurement.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9 (9): 4344–51. https://doi.org/10.1109/JSTARS.2016.2575360.

Asker, Dalal. 2018. “High Throughput Screening and Profiling of High-Value Carotenoids from a Wide Diversity of Bacteria in Surface Seawater.” *Food Chemistry* 261 (September): 103–11. https://doi.org/10.1016/j.foodchem.2018.03.109.

Badet, Thomas, and Daniel Croll. 2020. “The Rise and Fall of Genes: Origins and Functions of Plant Pathogen Pangenomes.” *Current Opinion in Plant Biology* 56 (August): 65–73. https://doi.org/10.1016/j.pbi.2020.04.009.

Bai, Geng, Yufeng Ge, David Scoby, Bryan Leavitt, Vincent Stoerger, Norbert Kirchgessner, Suat Irmak, George Graef, James Schnable, and Tala Awada. 2019. “NU-Spidercam: A Large-Scale, Cable-Driven, Integrated Sensing and Robotic System for Advanced Phenotyping, Remote Sensing, and Agronomic Research.” *Computers and Electronics in Agriculture* 160 (May): 71–

81. https://doi.org/10.1016/j.compag.2019.03.009.
Baker, Neil R. 2008. "Chlorophyll Fluorescence: A Probe of Photosynthesis In Vivo." *Annual Review of Plant Biology* 59 (1): 89–113. https://doi.org/10.1146/annurev.arplant.59.032607.092759.
Barbacci, Adelin, Olivier Navaud, Malick Mbengue, Marielle Barascud, Laurence Godiard, Mehdi Khafif, Aline Lacaze, and Sylvain Raffaele. 2020. "Rapid Identification of an Arabidopsis NLR Gene as a Candidate Conferring Susceptibility to Sclerotinia Sclerotiorum Using Time-resolved Automated Phenotyping." *The Plant Journal: For Cell and Molecular Biology* 103 (2): 903–17. https://doi.org/10.1111/tpj.14747.
Barbedo, Jayme G. A. 2018. "Factors Influencing the Use of Deep Learning for Plant Disease Recognition." *Biosystems Engineering* 172 (August): 84–91. https://doi.org/10.1016/j.biosystemseng.2018.05.013.
Barbedo, Jayme Garcia Arnal. 2016. "A Review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images." *Biosystems Engineering* 144 (April): 52–60. https://doi.org/10.1016/j.biosystemseng.2016.01.017.
Bauriegel, E., A. Giebel, M. Geyer, U. Schmidt, and W. B. Herppich. 2011. "Early Detection of Fusarium Infection in Wheat Using Hyper-Spectral Imaging." *Computers and Electronics in Agriculture* 75 (2): 304–12. https://doi.org/10.1016/j.compag.2010.12.006.
Bauriegel, Elke, Antje Giebel, and Werner B. Herppich. 2011. "Hyperspectral and Chlorophyll Fluorescence Imaging to Analyse the Impact of Fusarium Culmorum on the Photosynthetic Integrity of Infected Wheat Ears." *Sensors* 11 (4): 3765–79. https://doi.org/10.3390/s110403765.
Beilsmith, Kathleen, Manus P. M. Thoen, Benjamin Brachi, Andrew D. Gloss, Mohammad H. Khan, and Joy Bergelson. 2019. "Genome-Wide Association Studies on the Phyllosphere Microbiome: Embracing Complexity in Host--Microbe Interactions." *The Plant Journal: For Cell and Molecular Biology* 97 (1): 164–81. https://onlinelibrary.wiley.com/doi/abs/10.1111/tpj.14170.
Berkman, Paul J., Kaitao Lai, Michal T. Lorenc, and David Edwards. 2012. "Next-Generation Sequencing Applications for Wheat Crop Improvement." *American Journal of Botany* 99 (2): 365–71. https://doi.org/10.3732/ajb.1100309.
Berman, Judith, and Damian J. Krysan. 2020. "Author Correction: Drug Resistance and Tolerance in Fungi." *Nature Reviews. Microbiology* 18 (9): 539. https://doi.org/10.1038/s41579-020-0415-y.
Bettgenhaeuser, Jan, Matthew Gardiner, Rebecca Spanner, Phon Green, Inmaculada Hernández-Pinzón, Amelia Hubbard, Michael Ayliffe, and Matthew J. Moscou. 2018. "The Genetic Architecture of Colonization Resistance in Brachypodium Distachyon to Non-Adapted Stripe Rust (Puccinia Striiformis) Isolates." Edited by Rebecca S. Bart. *PLoS Genetics* 14 (9): e1007637. https://doi.org/10.1371/journal.pgen.1007637.
Bevan, Michael W., Cristobal Uauy, Brande B. H. Wulff, Ji Zhou, Ksenia Krasileva, and Matthew D. Clark. 2017. "Genomic Innovation for Crop Improvement." *Nature* 543 (7645): 346–54. https://doi.org/10.1038/nature22011.
Bhattacharya, Shaoni. 2017. "Deadly New Wheat Disease Threatens Europe's Crops." *Nature* 542 (7640): 145–46. https://doi.org/10.1038/nature.2017.21424.
Blake, Jonathan J., Paul Gosling, Bart A. Fraaije, Fiona J. Burnett, Stuart M. Knight, Steven Kildea, and Neil D. Paveley. 2018. "Changes in Field Dose-Response Curves for Demethylation Inhibitor (DMI) and Quinone Outside Inhibitor (QoI) Fungicides against Zymoseptoria Tritici, Related to Laboratory Sensitivity Phenotyping and Genotyping Assays." *Pest Management Science* 74 (2): 302–13. https://doi.org/10.1002/ps.4725.
Bock, C. H., P. E. Parker, A. Z. Cook, and T. R. Gottwald. 2008. "Visual Rating and the Use of Image Analysis for Assessing Different Symptoms of Citrus Canker on Grapefruit Leaves." *Plant Disease* 92 (4): 530–41. https://doi.org/10.1094/PDIS-92-4-0530.
Bock, C. H., G. H. Poole, P. E. Parker, and T. R. Gottwald. 2010. "Plant Disease Severity Estimated Visually, by Digital Photography and Image Analysis, and by Hyperspectral Imaging." *Critical Reviews in Plant Sciences* 29 (2): 59–107. https://doi.org/10.1080/07352681003617285.
Bravo, Cédric, Dimitrios Moshou, Jonathan West, Alastair McCartney, and Herman Ramon. 2003. "Early Disease Detection in Wheat Fields Using Spectral Reflectance." *Biosystems Engineering* 84 (2): 137–45. https://doi.org/10.1016/S1537-5110(02)00269-6.
Brown, James K. M., Laëtitia Chartrain, Pauline Lasserre-Zuber, and Cyrille Saintenac. 2015. "Genetics of Resistance to Zymoseptoria Tritici and Applications to Wheat Breeding." *Fungal*

Genetics and Biology: FG & B 79 (June): 33–41. https://doi.org/10.1016/J.FGB.2015.04.017.
Cairns, T. C., Y. S. Sidhu, Y. K. Chaudhari, N. J. Talbot, D. J. Studholme, and K. Haynes. 2015. "Construction and High-Throughput Phenotypic Screening ofZymoseptoria Tritici over-Expression Strains." *Fungal Genetics and Biology: FG & B* 79 (June): 110–17. https://doi.org/10.1016/j.fgb.2015.04.013.
Calderón, Rocío, Juan A. Navas-Cortés, and Pablo J. Zarco-Tejada. 2015. "Early Detection and Quantification of Verticillium Wilt in Olive Using Hyperspectral and Thermal Imagery over Large Areas." *Remote Sensing* 7 (5): 5584–5610. https://doi.org/10.3390/rs70505584.
Camargo, A., and J. S. Smith. 2009. "Image Pattern Classification for the Identification of Disease Causing Agents in Plants." *Computers and Electronics in Agriculture* 66 (2): 121–25. https://doi.org/10.1016/J.COMPAG.2009.01.003.
Campbell, Malachy T., Qian Du, Kan Liu, Chris J. Brien, Bettina Berger, Chi Zhang, and Harkamal Walia. 2017. "A Comprehensive Image-Based Phenomic Analysis Reveals the Complex Genetic Architecture of Shoot Growth Dynamics in Rice (Oryza Sativa)." *The Plant Genome* 10 (2): 1–14. https://acsess.onlinelibrary.wiley.com/doi/abs/10.3835/plantgenome2016.07.0064.
Caseys, C., G. Shi, N. Soltis, R. Gwinner, and J. Corwin. 2020. "Quantitative Interactions Drive Botrytis Cinerea Disease Outcome across the Plant Kingdom." *bioRxiv*. https://www.biorxiv.org/content/10.1101/507491v3.abstract.
Castro-Moretti, Fernanda R., Irene N. Gentzel, David Mackey, and Ana P. Alonso. 2020. "Metabolomics as an Emerging Tool for the Study of Plant–Pathogen Interactions." *Metabolites* 10 (2): 52. https://doi.org/10.3390/metabo10020052.
Chakraborty, S., and A. C. Newton. 2011. "Climate Change, Plant Diseases and Food Security: An Overview." *Plant Pathology*. https://doi.org/10.1111/j.1365-3059.2010.02411.x.
Chakravarti, Aravinda, and Tychele N. Turner. 2016. "Revealing Rate-Limiting Steps in Complex Disease Biology: The Crucial Importance of Studying Rare, Extreme-Phenotype Families." *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 38 (6): 578–86. https://doi.org/10.1002/bies.201500203.
Chapman, Scott C., Torsten Merz, Amy Chan, Paul Jackway, Stefan Hrabar, M. Fernanda Dreccer, Edward Holland, Bangyou Zheng, T. Jun Ling, and Jose Jimenez-Berni. 2014. "Pheno-Copter: A Low-Altitude, Autonomous Remote-Sensing Robotic Helicopter for High-Throughput Field-Based Phenotyping." *Agronomy* 4 (2): 279–301. https://doi.org/10.3390/agronomy4020279.
Chawade, Aakash, Joost Van Ham, Hanna Blomquist, Oscar Bagge, Erik Alexandersson, and Rodomiro Ortiz. 2019. "High-Throughput Field-Phenotyping Tools for Plant Breeding and Precision Agriculture." *Agronomy* 9 (5): 1–18. https://doi.org/10.3390/agronomy9050258.
Chen, Dijun, Kerstin Neumann, Swetlana Friedel, Benjamin Kilian, Ming Chen, Thomas Altmann, and Christian Klukas. 2014. "Dissecting the Phenotypic Components of Crop Plant Growth and Drought Responses Based on High-Throughput Image Analysis." *The Plant Cell Online* 26 (12): 4636–55. https://doi.org/10.1105/tpc.114.129601.
Chen, Fengping, Guo-Hua Duan, Dong-Liang Li, and Jiasui Zhan. 2017. "Host Resistance and Temperature-Dependent Evolution of Aggressiveness in the Plant Pathogen." *Frontiers in Microbiology* 8 (June): 1217. https://doi.org/10.3389/fmicb.2017.01217.
Chhetri, Hari B., David Macaya-Sanz, David Kainer, Ajaya K. Biswal, Luke M. Evans, Jin-Gui Chen, Cassandra Collins, et al. 2019. "Multitrait Genome-Wide Association Analysis of Populus Trichocarpa Identifies Key Polymorphisms Controlling Morphological and Physiological Traits." *The New Phytologist* 223 (1): 293–309. https://doi.org/10.1111/nph.15777.
Clark, Rachel L., Blair F. Johnston, Simon P. Mackay, Catherine J. Breslin, Murray N. Robertson, and Alan L. Harvey. 2010. "The Drug Discovery Portal: A Resource to Enhance Drug Discovery from Academia." *Drug Discovery Today*. https://doi.org/10.1016/j.drudis.2010.06.003.
Cobb, Joshua N., Genevieve DeClerck, Anthony Greenberg, Randy Clark, and Susan McCouch. 2013. "Next-Generation Phenotyping: Requirements and Strategies for Enhancing Our Understanding of Genotype-Phenotype Relationships and Its Relevance to Crop Improvement." *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik* 126 (4): 867–87. https://doi.org/10.1007/s00122-013-2066-0.
Corredor-Moreno, Pilar, and Diane G. O. Saunders. 2020. "Expecting the Unexpected: Factors Influencing the Emergence of Fungal and Oomycete Plant Pathogens." *The New Phytologist* 225

(1): 118–25. https://doi.org/10.1111/nph.16007.
Costa, J. Miguel, Olga M. Grant, and M. Manuela Chaves. 2013. "Thermography to Explore Plant–environment Interactions." *Journal of Experimental Botany* 64 (13): 3937–49. https://doi.org/10.1093/jxb/ert029.
Cowger, C., M. E. Hoffer, and C. C. Mundt. 2000. "Specific Adaptation by Mycosphaerella Graminicola to a Resistant Wheat Cultivar." *Plant Pathology*. https://doi.org/10.1046/j.1365-3059.2000.00472.x.
Cox, Kerik D., Kacie Quello, Ryan J. Deford, and Janna L. Beckerman. 2009. "A Rapid Method to Quantify Fungicide Sensitivity in the Brown Rot Pathogen Monilinia Fructicola." *Plant Disease* 93 (4): 328–31. https://doi.org/10.1094/PDIS-93-4-0328.
Croll, Daniel, and Bruce A. McDonald. 2017. "The Genetic Basis of Local Adaptation for Pathogenic Fungi in Agricultural Ecosystems." *Molecular Ecology* 26 (7): 2027–40. https://doi.org/10.1111/mec.13870.
Crute, Ian R. 1994. "Gene-for-Gene Recognition in Plant-Pathogen Interactions." *Philosophical Transactions - Royal Society of London, B* 346 (1317): 345–49. https://doi.org/10.1098/rstb.1994.0151.
Delmas, Chloé E. L., Frédéric Fabre, Jérôme Jolivet, Isabelle D. Mazet, Sylvie Richart Cervera, Laurent Delière, and François Delmotte. 2016. "Adaptation of a Plant Pathogen to Partial Host Resistance: Selection for Greater Aggressiveness in Grapevine Downy Mildew." *Evolutionary Applications* 9 (5): 709–25. https://doi.org/10.1111/eva.12368.
Does, H. Charlotte van der, and Martijn Rep. 2007. "Virulence Genes and the Evolution of Host Specificity in Plant-Pathogenic Fungi." *Molecular Plant-Microbe Interactions: MPMI* 20 (10): 1175–82. https://doi.org/10.1094/MPMI-20-10-1175.
Dutta, A., F. E. Hartmann, C. S. Francisco, and B. A. McDonald. 2020. "Mapping the Adaptive Landscape of a Major Agricultural Pathogen Reveals Evolutionary Constraints across Heterogeneous Environments." *bioRxiv*. https://www.biorxiv.org/content/10.1101/2020.07.30.229708v1.abstract.
Dutta, Anik, Daniel Croll, Bruce A. McDonald, and Luke G. Barrett. 2020. "Maintenance of Variation in Virulence and Reproduction in Populations of an Agricultural Plant Pathogen." https://doi.org/10.1101/2020.04.15.043208.
El-Mor, I. M., R. A. Fowler, G. J. Platz, M. W. Sutherland, and A. Martin. 2018. "An Improved Detached-Leaf Assay for Phenotyping Net Blotch of Barley Caused by Pyrenophora Teres." *Plant Disease* 102 (4): 760–63. https://doi.org/10.1094/PDIS-07-17-0980-RE.
Fahlgren, Noah, Malia A. Gehan, and Ivan Baxter. 2015. "Lights, Camera, Action: High-Throughput Plant Phenotyping Is Ready for a Close-Up." *Current Opinion in Plant Biology* 24: 93–99. https://doi.org/10.1016/j.pbi.2015.02.006.
Fao. 2017. "International Plant Protection Convention Plant Health and Food Security." https://www.ippc.int.
Fernandes, Tiago G., Maria Margarida Diogo, Douglas S. Clark, Jonathan S. Dordick, and Joaquim M. S. Cabral. 2009. "High-Throughput Cellular Microarray Platforms: Applications in Drug Discovery, Toxicology and Stem Cell Research." *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2009.02.009.
Fisher, Matthew C., Nichola J. Hawkins, Dominique Sanglard, and Sarah J. Gurr. 2018. "Worldwide Emergence of Resistance to Antifungal Drugs Challenges Human Health and Food Security." *Science* 360 (6390): 739–42. https://doi.org/10.1126/science.aap7999.
Fones, Helen N., Daniel P. Bebber, Thomas M. Chaloner, William T. Kay, Gero Steinberg, and Sarah J. Gurr. 2020. "Threats to Global Food Security from Emerging Fungal and Oomycete Crop Pathogens." *Nature Food* 1 (6): 332–42. https://doi.org/10.1038/s43016-020-0075-0.
Frantzeskakis, Lamprinos, Antonio Di Pietro, Martijn Rep, Jan Schirawski, Chih-hang Wu, and Ralph Panstruga. 2020a. "Rapid Evolution in Plant–microbe Interactions – a Molecular Genomics Perspective." *New Phytologist*. https://doi.org/10.1111/nph.15966.
Frantzeskakis, Lamprinos, Antonio Di Pietro, Martijn Rep, Jan Schirawski, Chih-Hang Wu, and Ralph Panstruga. 2020b. "Rapid Evolution in Plant--Microbe Interactions--a Molecular Genomics Perspective." *The New Phytologist* 225 (3): 1134–42. https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.15966.

Friesen, Timothy L., Eva H. Stukenbrock, Zhaohui Liu, Steven Meinhardt, Hua Ling, Justin D. Faris, Jack B. Rasmussen, Peter S. Solomon, Bruce A. McDonald, and Richard P. Oliver. 2006. "Emergence of a New Disease as a Result of Interspecific Virulence Gene Transfer." *Nature Genetics* 38 (8): 953–56. https://doi.org/10.1038/ng1839.

Goddard, Rachel, Andrew Steed, Catherine Chinoy, Jéssica Rosset Ferreira, Pedro Luiz Scheeren, João Leodato Nunes Maciel, Eduardo Caierão, et al. 2020. "Dissecting the Genetic Basis of Wheat Blast Resistance in the Brazilian Wheat Cultivar BR 18-Terena." *BMC Plant Biology* 20 (1): 398. https://doi.org/10.1186/s12870-020-02592-0.

Guo, Wei, Tokihiro Fukatsu, and Seishi Ninomiya. 2015. "Automated Characterization of Flowering Dynamics in Rice Using Field-Acquired Time-Series RGB Images." *Plant Methods* 11 (1): 7. https://doi.org/10.1186/s13007-015-0047-9.

Guyot, Stéphane, Patrick Gervais, Michael Young, Pascale Winckler, Jennifer Dumont, and Hazel Marie Davey. 2015. "Surviving the Heat: Heterogeneity of Response in Saccharomyces Cerevisiae Provides Insight into Thermal Damage to the Membrane." *Environmental Microbiology* 17 (8): 2982–92. https://doi.org/10.1111/1462-2920.12866.

Hall, Matthew D., Gilberto Bento, and Dieter Ebert. 2017. "The Evolutionary Consequences of Stepwise Infection Processes." *Trends in Ecology & Evolution* 32 (8): 612–23. https://doi.org/10.1016/j.tree.2017.05.009.

Hansen, Anders S., Nan Hao, and Erin K. O'Shea. 2015. "High-Throughput Microfluidics to Control and Measure Signaling Dynamics in Single Yeast Cells." *Nature Protocols* 10 (8): 1181–97. https://doi.org/10.1038/nprot.2015.079.

Hansen, Matthew C., and Thomas R. Loveland. 2012. "A Review of Large Area Monitoring of Land Cover Change Using Landsat Data." *Remote Sensing of Environment* 122 (July): 66–74. https://doi.org/10.1016/j.rse.2011.08.024.

Hartmann, Fanny E., Andrea Sánchez-Vallet, Bruce A. McDonald, and Daniel Croll. 2017. "A Fungal Wheat Pathogen Evolved Host Specialization by Extensive Chromosomal Rearrangements." *The ISME Journal* 11 (5): 1189–1204. https://doi.org/10.1038/ismej.2016.196.

Hartmann, Fanny E., Tiziana Vonlanthen, Nikhil Kumar Singh, Megan McDonald, Andrew Milgate, and Daniel Croll. n.d. "The Complex Genomic Basis of Rapid Convergent Adaptation to Pesticides across Continents in a Fungal Plant Pathogen." https://doi.org/10.1101/2020.07.24.220004.

Haueisen, Janine, Mareike Möller, Christoph J. Eschenbrenner, Jonathan Grandaubert, Heike Seybold, Holger Adamiak, and Eva H. Stukenbrock. 2019. "Highly Flexible Infection Programs in a Specialized Wheat Pathogen." *Ecology and Evolution*. https://doi.org/10.1002/ece3.4724.

He, Dun-Chun, Jia-Sui Zhan, and Lian-Hui Xie. 2016. "Problems, Challenges and Future of Plant Disease Management: From an Ecological Point of View." *Journal of Integrative Agriculture* 15 (4): 705–15. https://doi.org/10.1016/S2095-3119(15)61300-4.

He, L., S. Qi, J. Duan, T. Guo, W. Feng, and D. He. 2020. "Monitoring of Wheat Powdery Mildew Disease Severity Using Multiangle Hyperspectral Remote Sensing." *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, 1–12. https://doi.org/10.1109/TGRS.2020.3000992.

Hindorff, Lucia A., Praveen Sethupathy, Heather A. Junkins, Erin M. Ramos, Jayashri P. Mehta, Francis S. Collins, and Teri A. Manolio. 2009. "Potential Etiologic and Functional Implications of Genome-Wide Association Loci for Human Diseases and Traits." *Proceedings of the National Academy of Sciences of the United States of America* 106 (23): 9362–67. https://doi.org/10.1073/pnas.0903103106.

Holland, Sara L., Tom Reader, Paul S. Dyer, and Simon V. Avery. 2014. "Phenotypic Heterogeneity Is a Selected Trait in Natural Yeast Populations Subject to Environmental Stress." *Environmental Microbiology*. https://doi.org/10.1111/1462-2920.12243.

Horrocks, Kiran J., Darren Ward, and David M. Suckling. 2020. "Can Natural Enemies of Current Insect Pests Provide Biotic Resistance to Future Pests?" *Agricultural and Forest Entomology* 22 (1): 20–29. https://doi.org/10.1111/afe.12353.

Inoue, Yoshihiro, Trinh T. P. Vy, Kentaro Yoshida, Hokuto Asano, Chikako Mitsuoka, Soichiro Asuke, Vu L. Anh, et al. 2017. "Evolution of the Wheat Blast Fungus through Functional Losses in a Host Specificity Determinant." *Science* 357 (6346): 80–83.

https://doi.org/10.1126/science.aam9654.
Islam, M. Tofazzal, Daniel Croll, Pierre Gladieux, Darren M. Soanes, Antoine Persoons, Pallab Bhattacharjee, Md Shaid Hossain, et al. 2016. “Emergence of Wheat Blast in Bangladesh Was Caused by a South American Lineage of Magnaporthe Oryzae.” *BMC Biology* 14 (1): 84. https://doi.org/10.1186/s12915-016-0309-7.
Jain, Archana, Surendra Sarsaiya, Qin Wu, Yuanfu Lu, and Jingshan Shi. 2019. “A Review of Plant Leaf Fungal Diseases and Its Environment Speciation.” *Bioengineered* 10 (1): 409–24. https://doi.org/10.1080/21655979.2019.1649520.
Jones, Hamlyn G., Rachid Serraj, Brian R. Loveys, Lizhong Xiong, Ashley Wheaton, and Adam H. Price. 2009. “Thermal Infrared Imaging of Crop Canopies for the Remote Diagnosis and Quantification of Plant Responses to Water Stress in the Field.” *Functional Plant Biology: FPB* 36 (11): 978–89. https://doi.org/10.1071/FP09123.
Jones, Jonathan D. G., and Jeffery L. Dangl. 2006. “The Plant Immune System.” *Nature* 444 (7117): 323–29. https://doi.org/10.1038/nature05286.
Kado, Clarence I., ed. 2016. “CHAPTER 4: Symptoms Caused by Plant-Pathogenic Bacteria.” In *Plant Bacteriology*, 63–77. The American Phytopathological Society. https://doi.org/10.1094/9780890545034.004.
Karaoglanidis, G. S., C. C卄 Thanassoulopoulos, and P. M. Ioannidis. 2001. “Fitness of Cercospora Beticola Field Isolates--Resistant and--Sensitive to Demethylation Inhibitor Fungicides.” *European Journal of Plant Pathology / European Foundation for Plant Pathology* 107 (3): 337–47. https://link.springer.com/article/10.1023/A:1011219514343.
Karisto, Petteri, Andreas Hund, Kang Yu, Jonas Anderegg, Achim Walter, Fabio Mascher, Bruce A. McDonald, and Alexey Mikaberidze. 2018. “Ranking Quantitative Resistance to Septoria Tritici Blotch in Elite Wheat Cultivars Using Automated Image Analysis.” *Phytopathology* 108 (5): 568–81. https://doi.org/10.1094/PHYTO-04-17-0163-R.
Kefauver, Shawn C., Rubén Vicente, Omar Vergara-Díaz, Jose A. Fernandez-Gallego, Samir Kerfal, Antonio Lopez, James P. E. Melichar, María D. Serret Molins, and José L. Araus. 2017. “Comparative UAV and Field Phenotyping to Assess Yield and Nitrogen Use Efficiency in Hybrid and Conventional Barley.” *Frontiers in Plant Science* 8 (October): 1733. https://doi.org/10.3389/fpls.2017.01733.
Kim, Hwasoon, Alexander Grueneberg, Ana I. Vazquez, Stephen Hsu, and Gustavo de Los Campos. 2017. “Will Big Data Close the Missing Heritability Gap?” *Genetics* 207 (3): 1135–45. https://doi.org/10.1534/genetics.117.300271.
Kim, Song Lim, Nyunhee Kim, Hongseok Lee, Eungyeong Lee, Kyeong-Seong Cheon, Minsu Kim, Jeongho Baek, et al. 2020. “High-Throughput Phenotyping Platform for Analyzing Drought Tolerance in Rice.” *Planta* 252 (3): 38. https://doi.org/10.1007/s00425-020-03436-9.
Kirchgessner, Norbert, Frank Liebisch, Kang Yu, Johannes Pfeifer, Michael Friedli, Andreas Hund, and Achim Walter. 2017. “The ETH Field Phenotyping Platform FIP: A Cable-Suspended Multi-Sensor System.” *Functional Plant Biology*. https://doi.org/10.1071/fp16165.
Kitchen, James L., Frank van den Bosch, Neil D. Paveley, Joseph Helps, and Femke van den Berg. 2016. “The Evolution of Fungicide Resistance Resulting from Combinations of Foliar-Acting Systemic Seed Treatments and Foliar-Applied Fungicides: A Modeling Analysis.” *PloS One* 11 (8): e0161887. https://doi.org/10.1371/journal.pone.0161887.
Klodt, Maria, Katja Herzog, Reinhard Töpfer, and Daniel Cremers. 2015. “Field Phenotyping of Grapevine Growth Using Dense Stereo Reconstruction.” *BMC Bioinformatics* 16 (1): 143. https://doi.org/10.1186/s12859-015-0560-x.
Koenig, Kristina, Bernhard Höfle, Martin Hämmerle, Thomas Jarmer, Bastian Siegmann, and Holger Lilienthal. 2015. “Comparative Classification Analysis of Post-Harvest Growth Detection from Terrestrial LiDAR Point Clouds in Precision Agriculture.” *ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing* 104: 112–25. https://doi.org/10.1016/j.isprsjprs.2015.03.003.
Köhl, Jürgen, Rogier Kolnaar, and Willem J. Ravensberg. 2019. “Mode of Action of Microbial Biological Control Agents Against Plant Diseases: Relevance Beyond Efficacy.” *Frontiers in Plant Science* 10: 845. https://doi.org/10.3389/fpls.2019.00845.
Kuska, Matheus, Mirwaes Wahabzada, Marlene Leucker, Heinz-Wilhelm Dehne, Kristian Kersting,

Erich-Christian Oerke, Ulrike Steiner, and Anne-Katrin Mahlein. 2015. “Hyperspectral Phenotyping on the Microscopic Scale: Towards Automated Characterization of Plant-Pathogen Interactions.” *Plant Methods* 11 (1): 28. https://doi.org/10.1186/s13007-015-0073-7.

Kwolek-Mirek, Magdalena, and Renata Zadrag-Tecza. 2014. “Comparison of Methods Used for Assessing the Viability and Vitality of Yeast Cells.” *FEMS Yeast Research* 14 (7): 1068–79. https://doi.org/10.1111/1567-1364.12202.

Laine, A-L. 2007. “Pathogen Fitness Components and Genotypes Differ in Their Sensitivity to Nutrient and Temperature Variation in a Wild Plant–pathogen Association.” *Journal of Evolutionary Biology* 20 (6): 2371–78. https://doi.org/10.1111/j.1420-9101.2007.01406.x.

Laine, Anna-Liisa. 2008. “Temperature-Mediated Patterns of Local Adaptation in a Natural Plant--Pathogen Metapopulation.” *Ecology Letters* 11 (4): 327–37. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1461-0248.2007.01146.x.

Lambrechts, Louis. 2010. “Dissecting the Genetic Architecture of Host–Pathogen Specificity.” Edited by Glenn F. Rall. *PLoS Pathogens* 6 (8): e1001019. https://doi.org/10.1371/journal.ppat.1001019.

Leach, Michelle D., and Leah E. Cowen. 2014. “To Sense or Die: Mechanisms of Temperature Sensing in Fungal Pathogens.” *Current Fungal Infection Reports* 8 (2): 185–91. https://doi.org/10.1007/s12281-014-0182-1.

Leucker, Marlene, Mirwaes Wahabzada, Kristian Kersting, Madlaina Peter, Werner Beyer, Ulrike Steiner, Anne-Katrin Mahlein, and Erich-Christian Oerke. 2016. “Hyperspectral Imaging Reveals the Effect of Sugar Beet Quantitative Trait Loci on Cercospora Leaf Spot Resistance.” *Functional Plant Biology: FPB* 44 (1): 1–9. https://doi.org/10.1071/FP16121.

Levin-Reisman, Irit, Ofer Fridman, and Nathalie Q. Balaban. 2014. “ScanLag: High-Throughput Quantification of Colony Growth and Lag Time.” *Journal of Visualized Experiments*. https://doi.org/10.3791/51456.

Lievens, Eva J. P., Julie Perreau, Philip Agnew, Yannis Michalakis, and Thomas Lenormand. 2018. “Decomposing Parasite Fitness Reveals the Basis of Specialization in a Two-Host, Two-Parasite System.” *Evolution Letters* 2 (4): 390–405. https://doi.org/10.1002/evl3.65.

Li, Tingting, Chunshan Jiang, Zhenqi Bian, Mingchang Wang, and Xuefeng Niu. 2020. “A Review of True Orthophoto Rectification Algorithms.” *IOP Conference Series: Materials Science and Engineering* 780 (2): 022035. https://doi.org/10.1088/1757-899X/780/2/022035.

Liu, Yu, Xinqi Zheng, Gang Ai, Yi Zhang, and Yuqiang Zuo. 2018. “Generating a High-Precision True Digital Orthophoto Map Based on UAV Images.” *ISPRS International Journal of Geo-Information* 7 (9): 333. https://doi.org/10.3390/ijgi7090333.

Li, Zhenbo, Ruohao Guo, Meng Li, Yaru Chen, and Guangyao Li. 2020. “A Review of Computer Vision Technologies for Plant Phenotyping.” *Computers and Electronics in Agriculture* 176 (September): 105672. https://doi.org/10.1016/j.compag.2020.105672.

Ludovisi, Riccardo, Flavia Tauro, Riccardo Salvati, Sacha Khoury, Giuseppe Mugnozza Scarascia, and Antoine Harfouche. 2017. “UAV-Based Thermal Imaging for High-Throughput Field Phenotyping of Black Poplar Response to Drought.” *Frontiers in Plant Science* 8 (September): 1681. https://doi.org/10.3389/fpls.2017.01681.

MacPherson, Ailene, Sarah P. Otto, and Scott L. Nuismer. 2018. “Keeping Pace with the Red Queen: Identifying the Genetic Basis of Susceptibility to Infectious Disease.” *Genetics* 208 (2): 779–89. https://doi.org/10.1534/genetics.117.300481.

Madec, Simon, Fred Baret, Benoît de Solan, Samuel Thomas, Dan Dutartre, Stéphane Jezequel, Matthieu Hemmerlé, Gallian Colombeau, and Alexis Comar. 2017. “High-Throughput Phenotyping of Plant Height: Comparing Unmanned Aerial Vehicles and Ground LiDAR Estimates.” *Frontiers in Plant Science* 8 (November): 2002. https://doi.org/10.3389/fpls.2017.02002.

Mahlein, A-K, U. Steiner, H-W Dehne, and E-C Oerke. 2010. “Spectral Signatures of Sugar Beet Leaves for the Detection and Differentiation of Diseases.” *Precision Agriculture* 11 (4): 413–31. https://doi.org/10.1007/s11119-010-9180-7.

Mahlein, Anne-Katrin, Ulrike Steiner, Christian Hillnhütter, Heinz-Wilhelm Dehne, and Erich-Christian Oerke. 2012. “Hyperspectral Imaging for Small-Scale Analysis of Symptoms Caused by Different Sugar Beet Diseases.” *Plant Methods* 8 (1): 3. https://doi.org/10.1186/1746-4811-8-

3.
Majid, Asma, G. A. Parray, Shabir H. Wani, Mojtoba Kordostami, N. R. Sofi, Showkat A. Waza, A. B. Shikari, and Shazia Gulzar. 2017. “Genome Editing and Its Necessity in Agriculture.” *International Journal of Current Microbiology and Applied Sciences*. https://doi.org/10.20546/ijcmas.2017.611.520.
Maleki, Sedigheh, Farhad Khormali, Jahangir Mohammadi, Patrick Bogaert, and Mohsen Bagheri Bodaghabadi. 2020. “Effect of the Accuracy of Topographic Data on Improving Digital Soil Mapping Predictions with Limited Soil Data: An Application to the Iranian Loess Plateau.” *Catena* 195 (December): 104810. https://doi.org/10.1016/j.catena.2020.104810.
Mariette, Nicolas, Alexander Kröner, Romain Mabon, Josselin Montarry, Bruno Marquer, Roselyne Corbière, Annabelle Androdias, and Didier Andrivon. 2018. “A Trade-Off Between Sporangia Size and Number Exists in the Potato Late Blight Pathogen Phytophthora Infestans, and Is Not Altered by Biotic and Abiotic Factors.” *Frontiers in Plant Science* 9 (December): 1841. https://doi.org/10.3389/fpls.2018.01841.
Märkle, Hanna, and Aurélien Tellier. 2020. “Inference of Coevolutionary Dynamics and Parameters from Host and Parasite Polymorphism Data of Repeated Experiments.” *PLoS Computational Biology* 16 (3): e1007668. https://doi.org/10.1371/journal.pcbi.1007668.
Martin, Anke, Paula Moolhuijzen, Yongfu Tao, Judy McIlroy, Simon R. Ellwood, Ryan A. Fowler, Greg J. Platz, Andrzej Kilian, and Lisle Snyman. 2020. “Genomic Regions Associated with Virulence in Pyrenophora Teres F. Teres Identified by Genome-Wide Association Analysis and Biparental Mapping.” *Phytopathology®*. https://doi.org/10.1094/phyto-10-19-0372-r.
Martinelli, Federico, Riccardo Scalenghe, Salvatore Davino, Stefano Panno, Giuseppe Scuderi, Paolo Ruisi, Paolo Villa, et al. 2015. “Advanced Methods of Plant Disease Detection. A Review.” *Agronomy for Sustainable Development* 35 (1): 1–25. https://doi.org/10.1007/s13593-014-0246-1.
McAusland, Lorna, Jonathan A. Atkinson, Tracy Lawson, and Erik H. Murchie. 2019. “High Throughput Procedure Utilising Chlorophyll Fluorescence Imaging to Phenotype Dynamic Photosynthesis and Photoprotection in Leaves under Controlled Gaseous Conditions.” *Plant Methods* 15 (1): 109. https://doi.org/10.1186/s13007-019-0485-x.
McCartney, H. A., and B. D. L. Fitt. 1998. “Dispersal of Foliar Fungal Plant Pathogens: Mechanisms, Gradients and Spatial Patterns.” In *The Epidemiology of Plant Diseases*, edited by D. Gareth Jones, 138–60. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-3302-1_7.
McDonald, Bruce A., and Eva H. Stukenbrock. 2016. “Rapid Emergence of Pathogens in Agro-Ecosystems: Global Threats to Agricultural Sustainability and Food Security.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 371 (1709). https://doi.org/10.1098/rstb.2016.0026.
Mishra, Bharat, Nilesh Kumar, and M. Shahid Mukhtar. 2019. “Systems Biology and Machine Learning in Plant-Pathogen Interactions.” *Molecular Plant-Microbe Interactions: MPMI* 32 (1): 45–55. https://doi.org/10.1094/MPMI-08-18-0221-FI.
M., Kishore, and S. B. Kulkarni. 2015. “Hyperspectral Imaging Technique for Plant Leaf Identification.” *2015 International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, 209–13. https://doi.org/10.1109/ERECT.2015.7499014.
Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. 2016. “Using Deep Learning for Image-Based Plant Disease Detection.” *Frontiers in Plant Science* 7 (September): 1419. https://doi.org/10.3389/fpls.2016.01419.
Mohd-Assaad, Norfarhan, Bruce A. McDonald, and Daniel Croll. 2019. “The Emergence of the Multi-Species NIP1 Effector in Rhynchosporium Was Accompanied by High Rates of Gene Duplications and Losses.” *Environmental Microbiology* 21 (8): 2677–95. https://doi.org/10.1111/1462-2920.14583.
Möller, Mareike, Michael Habig, Michael Freitag, and Eva H. Stukenbrock. 2018. “Extraordinary Genome Instability and Widespread Chromosome Rearrangements During Vegetative Growth.” *Genetics* 210 (2): 517–29. https://doi.org/10.1534/genetics.118.301050.
Monteiro, Maria Cândida, Mercedes de la Cruz, Juan Cantizani, Catalina Moreno, José R. Tormo, Emilia Mellado, J. Ramón De Lucas, et al. 2012. “A New Approach to Drug Discovery: High-Throughput Screening of Microbial Natural Extracts against Aspergillus Fumigatus Using

Resazurin." *Journal of Biomolecular Screening* 17 (4): 542–49. https://doi.org/10.1177/1087057111433459.
Mulema, Joseph M. K., and Katherine J. Denby. 2012. "Spatial and Temporal Transcriptomic Analysis of the Arabidopsis thaliana–Botrytis Cinerea Interaction." *Molecular Biology Reports*. https://doi.org/10.1007/s11033-011-1185-4.
Mulla, David J. 2013. "Twenty Five Years of Remote Sensing in Precision Agriculture: Key Advances and Remaining Knowledge Gaps." *Biosystems Engineering* 114 (4): 358–71. https://doi.org/10.1016/j.biosystemseng.2012.08.009.
Müller-Linow, Mark, Benedikt Janssen, Angelina Steier, and Uwe Rascher. 2015. "3-D Field Phenotyping of Crops Using Laser Scanning and Photogrammetric Approaches." In *Workshop on Laser Scanning Applications*. http://juser.fz-juelich.de/record/188656.
Mutka, Andrew M., and Rebecca S. Bart. 2014. "Image-Based Phenotyping of Plant Disease Symptoms." *Frontiers in Plant Science* 5: 734. https://doi.org/10.3389/fpls.2014.00734.
Mutka, Andrew M., Sarah J. Fentress, Joel W. Sher, Jeffrey C. Berry, Chelsea Pretz, Dmitri A. Nusinow, and Rebecca Bart. 2016. "Quantitative, Image-Based Phenotyping Methods Provide Insight into Spatial and Temporal Dimensions of Plant Disease." *Plant Physiology* 172 (2): 650–60. https://doi.org/10.1104/pp.16.00984.
Mwadzingeni, Learnmore, Hussein Shimelis, Ernest Dube, Mark D. Laing, and Toi J. Tsilo. 2016. "Breeding Wheat for Drought Tolerance: Progress and Technologies." *Journal of Integrative Agriculture*. https://doi.org/10.1016/s2095-3119(15)61102-9.
Nagar, Shashwat Deepali, Bhavye Aggarwal, Shikha Joon, Rakesh Bhatnagar, and Sonika Bhatnagar. 2016. "A Network Biology Approach to Decipher Stress Response in Bacteria Using Escherichia Coli As a Model." *Omics: A Journal of Integrative Biology* 20 (5): 310–24. https://doi.org/10.1089/omi.2016.0028.
Neumann, M., L. Hallau, B. Klatt, K. Kersting, and C. Bauckhage. 2014. "Erosion Band Features for Cell Phone Image Based Plant Disease Classification." In *2014 22nd International Conference on Pattern Recognition*, 3315–20. https://doi.org/10.1109/ICPR.2014.571.
Odilbekov, Firuz, Eva Edin, Hadis Mostafanezhad, Hilde Coolman, Laura J. Grenville-Briggs, and Erland Liljeroth. 2019. "Within-Season Changes in Alternaria Solani Populations in Potato in Response to Fungicide Application Strategies." *European Journal of Plant Pathology / European Foundation for Plant Pathology* 155 (3): 953–65. https://doi.org/10.1007/s10658-019-01826-8.
Oerke, E-C, P. Fröhling, and U. Steiner. 2011. "Thermographic Assessment of Scab Disease on Apple Leaves." *Precision Agriculture* 12 (5): 699–715. https://doi.org/10.1007/s11119-010-9212-3.
Oerke, E-C, U. Steiner, H-W Dehne, and M. Lindenthal. 2006. "Thermal Imaging of Cucumber Leaves Affected by Downy Mildew and Environmental Conditions." *Journal of Experimental Botany* 57 (9): 2121–32. https://doi.org/10.1093/jxb/erj170.
Palloix, A., V. Ayme, and B. Moury. 2009. "Durability of Plant Major Resistance Genes to Pathogens Depends on the Genetic Background, Experimental Evidence and Consequences for Breeding Strategies." *The New Phytologist* 183 (1): 190–99. https://doi.org/10.1111/j.1469-8137.2009.02827.x.
Pasco, Claudine, Josselin Montarry, Bruno Marquer, and Didier Andrivon. 2016. "And the Nasty Ones Lose in the End: Foliar Pathogenicity Trades off with Asexual Transmission in the Irish Famine Pathogen Phytophthora Infestans." *The New Phytologist* 209 (1): 334–42. https://doi.org/10.1111/nph.13581.
Peeters, Nemo, Alice Guidot, Fabienne Vailleau, and Marc Valls. 2013. "R Alstonia Solanacearum, a Widespread Bacterial Plant Pathogen in the Post-Genomic Era." *Molecular Plant Pathology* 14 (7): 651–62. https://bsppjournals.onlinelibrary.wiley.com/doi/abs/10.1111/mpp.12038.
Pereira, Danilo, Daniel Croll, Patrick C. Brunner, and Bruce A. McDonald. 2020. "Natural Selection Drives Population Divergence for Local Adaptation in a Wheat Pathogen." *Fungal Genetics and Biology: FG & B* 141 (August): 103398. https://doi.org/10.1016/j.fgb.2020.103398.
Perilla-Henao, Laura M., and Clare L. Casteel. 2016. "Vector-Borne Bacterial Plant Pathogens: Interactions with Hemipteran Insects and Plants." *Frontiers in Plant Science* 7 (August): 1163. https://doi.org/10.3389/fpls.2016.01163.
Petit-Houdenot, Yohann, and Isabelle Fudal. 2017. "Complex Interactions between Fungal Avirulence

Genes and Their Corresponding Plant Resistance Genes and Consequences for Disease Resistance Management." *Frontiers in Plant Science* 8 (June): 1072. https://doi.org/10.3389/fpls.2017.01072.

Peyraud, Rémi, Ludovic Cottret, Lucas Marmiesse, Jérôme Gouzy, and Stéphane Genin. 2016. "A Resource Allocation Trade-Off between Virulence and Proliferation Drives Metabolic Versatility in the Plant Pathogen Ralstonia Solanacearum." *PLoS Pathogens* 12 (10): e1005939. https://doi.org/10.1371/journal.ppat.1005939.

Peyraud, Rémi, Ullrich Dubiella, Adelin Barbacci, Stéphane Genin, Sylvain Raffaele, and Dominique Roby. 2017. "Advances on Plant-Pathogen Interactions from Molecular toward Systems Biology Perspectives." *The Plant Journal: For Cell and Molecular Biology* 90 (4): 720–37. https://doi.org/10.1111/tpj.13429.

Porter, Heather F., and Paul F. O'Reilly. 2017. "Multivariate Simulation Framework Reveals Performance of Multi-Trait GWAS Methods." *Scientific Reports* 7 (March): 38837. https://doi.org/10.1038/srep38837.

Rahimi, Yousef, Mohammad Reza Bihamta, Alireza Taleei, Hadi Alipour, and Pär K. Ingvarsson. 2019. "Genome-Wide Association Study of Agronomic Traits in Bread Wheat Reveals Novel Putative Alleles for Future Breeding Programs." *BMC Plant Biology* 19 (1): 541. https://doi.org/10.1186/s12870-019-2165-4.

Raza, Shan E. Ahmed, Gillian Prince, John P. Clarkson, and Nasir M. Rajpoot. 2015. "Automatic Detection of Diseased Tomato Plants Using Thermal and Stereo Visible Light Images." Edited by Dragan Perovic. *PloS One* 10 (4): e0123262. https://doi.org/10.1371/journal.pone.0123262.

Rheenen, Wouter van, Wouter J. Peyrot, Andrew J. Schork, S. Hong Lee, and Naomi R. Wray. 2019. "Genetic Correlations of Polygenic Disease Traits: From Theory to Practice." *Nature Reviews. Genetics* 20 (10): 567–81. https://doi.org/10.1038/s41576-019-0137-z.

Richards, Jonathan K., Eva H. Stukenbrock, Jessica Carpenter, Zhaohui Liu, Christina Cowger, Justin D. Faris, and Timothy L. Friesen. 2019. "Local Adaptation Drives the Diversification of Effectors in the Fungal Wheat Pathogen Parastagonospora Nodorum in the United States." *PLoS Genetics* 15 (10): e1008223. https://doi.org/10.1371/journal.pgen.1008223.

Rohr, Jason R., Jenise Brown, William A. Battaglin, Taegan A. McMahon, and Rick A. Relyea. 2017. "A Pesticide Paradox: Fungicides Indirectly Increase Fungal Infections." *Ecological Applications: A Publication of the Ecological Society of America* 27 (8): 2290–2302. https://doi.org/10.1002/eap.1607.

Rumpf, T., A-K Mahlein, U. Steiner, E-C Oerke, H-W Dehne, and L. Plümer. 2010. "Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance." *Computers and Electronics in Agriculture* 74 (1): 91–99. https://doi.org/10.1016/J.COMPAG.2010.06.009.

Rustagi, Anjana, Garvita Singh, Shachi Agrawal, and Prem Kumar Gupta. 2018. "Proteomic Studies Revealing Enigma of Plant–Pathogen Interaction." *Molecular Aspects of Plant-Pathogen Interaction*. https://doi.org/10.1007/978-981-10-7371-7_11.

Saintenac, Cyrille, Wing-Sham Lee, Florence Cambon, Jason J. Rudd, Robert C. King, William Marande, Stephen J. Powers, et al. 2018. "Wheat Receptor-Kinase-like Protein Stb6 Controls Gene-for-Gene Resistance to Fungal Pathogen Zymoseptoria Tritici." *Nature Genetics* 50 (3): 368–74. https://doi.org/10.1038/s41588-018-0051-x.

Salas Fernandez, Maria G., Yin Bao, Lie Tang, and Patrick S. Schnable. 2017. "A High-Throughput, Field-Based Phenotyping Technology for Tall Biomass Crops." *Plant Physiology* 174 (4): 2008–22. https://doi.org/10.1104/pp.17.00707.

Sánchez-Vallet, Andrea, Fanny E. Hartmann, Thierry C. Marcel, and Daniel Croll. 2018. "Nature's Genetic Screens: Using Genome-Wide Association Studies for Effector Discovery." *Molecular Plant Pathology* 19 (1): 3–6. https://doi.org/10.1111/mpp.12592.

Sankaran, Sindhuja, Lav R. Khot, Carlos Zúñiga Espinoza, Sanaz Jarolmasjed, Vidyasagar R. Sathuvalli, George J. Vandemark, Phillip N. Miklas, et al. 2015. "Low-Altitude, High-Resolution Aerial Imaging Systems for Row and Field Crop Phenotyping: A Review." *European Journal of Agronomy*. Elsevier. https://doi.org/10.1016/j.eja.2015.07.004.

Schmidt, Sarah Maria, Joanna Lukasiewicz, Rhys Farrer, Peter van Dam, Chiara Bertoldo, and Martijn Rep. 2016. "Comparative Genomics of *Fusarium Oxysporum* F. Sp. *melonis* Reveals the

Secreted Protein Recognized by the *Fom-2* Resistance Gene in Melon." *The New Phytologist* 209 (1): 307–18. https://doi.org/10.1111/nph.13584.

Schröter, Larissa, and Petra Dersch. 2019. "Phenotypic Diversification of Microbial Pathogens-Cooperating and Preparing for the Future." *Journal of Molecular Biology* 431 (23): 4645–55. https://doi.org/10.1016/j.jmb.2019.06.024.

Shahin, Muhammad A., and Stephen J. Symons. 2012. "Detection of Fusarium Damage in Canadian Wheat Using Visible/near-Infrared Hyperspectral Imaging." *Journal of Food Measurement & Characterization*. https://doi.org/10.1007/s11694-012-9126-z.

Shakoor, Nadia, Scott Lee, and Todd C. Mockler. 2017. "High Throughput Phenotyping to Accelerate Crop Breeding and Monitoring of Diseases in the Field." *Current Opinion in Plant Biology* 38 (August): 184–92. https://doi.org/10.1016/j.pbi.2017.05.006.

Shaw, Michael W., Christy J. Emmanuel, Deni Emilda, Razak B. Terhem, Aminath Shafia, Dimitra Tsamaidi, Mark Emblow, and Jan A. L. van Kan. 2016. "Analysis of Cryptic, Systemic Botrytis Infections in Symptomless Hosts." *Frontiers in Plant Science* 7 (May): 625. https://doi.org/10.3389/fpls.2016.00625.

Singh, Nikhil Kumar, Emilie Chanclud, and Daniel Croll. n.d. "Population-Level Deep Sequencing Reveals the Interplay of Clonal and Sexual Reproduction in the Fungal Wheat Pathogen Zymoseptoria Tritici." https://doi.org/10.1101/2020.07.07.191510.

Spanu, Pietro D., and Ralph Panstruga. 2017. "Editorial: Biotrophic Plant-Microbe Interactions." *Frontiers in Plant Science* 8 (February): 192. https://doi.org/10.3389/fpls.2017.00192.

Spinelli, F., M. Noferini, and G. Costa. 2004. "Near Infrared Spectroscopy (NIRs): Perspective of Fire Blight Detection in Asymptomatic Plant Material." In *X International Workshop on Fire Blight 704*, 87–90. https://www.actahort.org/books/704/704_9.htm.

Steinhauer, Diana, Marie Salat, Regula Frey, Andreas Mosbach, Torsten Luksch, Dirk Balmer, Rasmus Hansen, et al. 2019. "A Dispensable Paralog of Succinate Dehydrogenase Subunit C Mediates Standing Resistance towards a Subclass of SDHI Fungicides in Zymoseptoria Tritici." *PLoS Pathogens* 15 (12): e1007780. https://doi.org/10.1371/journal.ppat.1007780.

Stewart, Ethan L., and Bruce A. Mcdonald. 2014. "Measuring Quantitative Virulence in the Wheat Pathogen Zymoseptoria Tritici Using High-Throughput Automated Image Analysis" 104 (9). https://doi.org/10.1094/PHYTO-11-13-0328-R.

Strange, Richard N., and Peter R. Scott. 2005. "Plant Disease: A Threat to Global Food Security." *Annual Review of Phytopathology* 43: 83–116. https://doi.org/10.1146/annurev.phyto.43.113004.133839.

Suffert, Frédéric, Virginie Ravigné, and Ivan Sachec. 2015. "Seasonal Changes Drive Short-Term Selection for Fitness Traits in the Wheat Pathogen Zymoseptoria Tritici." *Applied and Environmental Microbiology* 81 (18): 6367–79. https://doi.org/10.1128/AEM.00529-15.

Sugiura, Ryo, Shogo Tsuda, Seiji Tamiya, Atsushi Itoh, Kentaro Nishiwaki, Noriyuki Murakami, Yukinori Shibuya, Masayuki Hirafuji, and Stephen Nuske. 2016. "Field Phenotyping System for the Assessment of Potato Late Blight Resistance Using RGB Imagery from an Unmanned Aerial Vehicle." *Biosystems Engineering* 148 (August): 1–10. https://doi.org/10.1016/j.biosystemseng.2016.04.010.

Sundin, George W., Luisa F. Castiblanco, Xiaochen Yuan, Quan Zeng, and Ching-Hong Yang. 2016. "Bacterial Disease Management: Challenges, Experience, Innovation and Future Prospects: Challenges in Bacterial Molecular Plant Pathology." *Molecular Plant Pathology* 17 (9): 1506–18. https://doi.org/10.1111/mpp.12436.

Szymański, Paweł, Magdalena Markowicz, and Elżbieta Mikiciuk-Olasik. 2012. "Adaptation of High-Throughput Screening in Drug Discovery-Toxicological Screening Tests." *International Journal of Molecular Sciences* 13 (1): 427–52. https://doi.org/10.3390/ijms13010427.

Thomas, Stefan, Jan Behmann, Angelina Steier, Thorsten Kraska, Onno Muller, Uwe Rascher, and Anne Katrin Mahlein. 2018. "Quantitative Assessment of Disease Severity and Rating of Barley Cultivars Based on Hyperspectral Imaging in a Non-Invasive, Automated Phenotyping Platform." *Plant Methods* 14 (1): 45. https://doi.org/10.1186/s13007-018-0313-8.

Tisné, S., Y. Serrand, L. Bach, and E. Gilbault. 2013. "Phenoscope: An Automated Large-scale Phenotyping Platform Offering High Spatial Homogeneity." *The Plant*. https://onlinelibrary.wiley.com/doi/abs/10.1111/tpj.12131.

Torriani, Stefano F. F., James P. E. Melichar, Colin Mills, Naomi Pain, Helge Sierotzki, and Mikaël Courbot. 2015. "Zymoseptoria Tritici: A Major Threat to Wheat Production, Integrated Approaches to Control." *Fungal Genetics and Biology: FG & B* 79 (June): 8–12. https://doi.org/10.1016/j.fgb.2015.04.010.

Van de Wouw, Angela P., Vicki L. Elliott, Steven Chang, Francisco J. López-Ruiz, Steven J. Marcroft, and Alexander Idnurm. 2017. "Identification of Isolates of the Plant Pathogen Leptosphaeria Maculans with Resistance to the Triazole Fungicide Fluquinconazole Using a Novel In Planta Assay." *PloS One* 12 (11): e0188106. https://doi.org/10.1371/journal.pone.0188106.

Varshney, Rajeev K., David A. Hoisington, and Akhilesh K. Tyagi. 2006. "Advances in Cereal Genomics and Applications in Crop Breeding." *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2006.08.006.

Velásquez, André C., Christian Danve M. Castroverde, and Sheng Yang He. 2018. "Plant–Pathogen Warfare under Changing Climate Conditions." *Current Biology: CB* 28 (10): R619–34. https://doi.org/10.1016/j.cub.2018.03.054.

Vleeshouwers, Vivianne G. A. A., and Richard P. Oliver. 2014. "Effectors as Tools in Disease Resistance Breeding Against Biotrophic, Hemibiotrophic, and Necrotrophic Plant Pathogens." *Molecular Plant-Microbe Interactions®* 27 (3): 196–206. https://doi.org/10.1094/MPMI-10-13-0313-IA.

Vries, Sophie de, Eva H. Stukenbrock, and Laura E. Rose. 2020. "Rapid Evolution in Plant-Microbe Interactions - an Evolutionary Genomics Perspective." *The New Phytologist* 226 (5): 1256–62. https://doi.org/10.1111/nph.16458.

Wahabzada, Mirwaes, Anne-Katrin Mahlein, Christian Bauckhage, Ulrike Steiner, Erich-Christian Oerke, and Kristian Kersting. 2015. "Metro Maps of Plant Disease Dynamics—Automated Mining of Differences Using Hyperspectral Images." *PloS One* 10 (1): e0116902. https://doi.org/10.1371/journal.pone.0116902.

Walter, James, James Edwards, Jinhai Cai, Glenn McDonald, Stanley J. Miklavcic, and Haydn Kuchel. 2019. "High-Throughput Field Imaging and Basic Image Analysis in a Wheat Breeding Programme." *Frontiers in Plant Science* 10 (April): 449. https://doi.org/10.3389/fpls.2019.00449.

Wang, Junping, Pieter Badenhorst, Andrew Phelan, Luke Pembleton, Fan Shi, Noel Cogan, German Spangenberg, and Kevin Smith. 2019. "Using Sensors and Unmanned Aircraft Systems for High-Throughput Phenotyping of Biomass in Perennial Ryegrass Breeding Trials." *Frontiers in Plant Science* 10: 1381. https://doi.org/10.3389/fpls.2019.01381.

Wang, Miaoyan, Fabrice Roux, Claudia Bartoli, Carine Huard-Chauveau, Christopher Meyer, Hana Lee, Dominique Roby, Mary Sara McPeek, and Joy Bergelson. 2018. "Two-Way Mixed-Effects Methods for Joint Association Analysis Using Both Host and Pathogen Genomes." *Proceedings of the National Academy of Sciences of the United States of America* 115 (24): E5440–49. https://doi.org/10.1073/pnas.1710980115.

Wijekoon, C. P., P. H. Goodwin, and T. Hsiang. 2008. "Quantifying Fungal Infection of Plant Leaves by Digital Image Analysis Using Scion Image Software." *Journal of Microbiological Methods* 74 (2-3): 94–101. https://doi.org/10.1016/j.mimet.2008.03.008.

Wilkinson, Samuel W., Melissa H. Magerøy, Ana López Sánchez, Lisa M. Smith, Leonardo Furci, T. E. Anne Cotton, Paal Krokene, and Jurriaan Ton. 2019. "Surviving in a Hostile World: Plant Strategies to Resist Pests and Diseases." *Annual Review of Phytopathology* 57 (August): 505–29. https://doi.org/10.1146/annurev-phyto-082718-095959.

Xu, Rui, Changying Li, and Andrew H. Paterson. 2019. "Multispectral Imaging and Unmanned Aerial Systems for Cotton Plant Phenotyping." *PloS One* 14 (2): e0205083. https://doi.org/10.1371/journal.pone.0205083.

Yang, Wanneng, Hui Feng, Xuehai Zhang, Jian Zhang, John H. Doonan, William David Batchelor, Lizhong Xiong, and Jianbing Yan. 2020. "Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives." *Molecular Plant* 13 (2): 187–214. https://doi.org/10.1016/j.molp.2020.01.008.

Yang, Wanneng, Zilong Guo, Chenglong Huang, Lingfeng Duan, Guoxing Chen, Ni Jiang, Wei Fang, et al. 2014. "Combining High-Throughput Phenotyping and Genome-Wide Association Studies

to Reveal Natural Genetic Variation in Rice." *Nature Communications* 5: 1–9. https://doi.org/10.1038/ncomms6087.

Young, Sierra N., Erkan Kayacan, and Joshua M. Peschel. 2019. "Design and Field Evaluation of a Ground Robot for High-Throughput Phenotyping of Energy Sorghum." *Precision Agriculture* 20 (4): 697–722. https://doi.org/10.1007/s11119-018-9601-6.

Zaidem, Maricris L., Simon C. Groen, and Michael D. Purugganan. 2019. "Evolutionary and Ecological Functional Genomics, from Lab to the Wild." *The Plant Journal: For Cell and Molecular Biology* 97 (1): 40–55. https://doi.org/10.1111/tpj.14167.

Zhang, Dongyan, Xingen Zhou, Jian Zhang, Yubin Lan, Chao Xu, and Dong Liang. 2018. "Detection of Rice Sheath Blight Using an Unmanned Aerial System with High-Resolution Color and Multispectral Imaging." Edited by Zonghua Wang. *PloS One* 13 (5): e0187470. https://doi.org/10.1371/journal.pone.0187470.

Zhang, Jiaoping, Arti Singh, Daren S. Mueller, and Asheesh K. Singh. 2015. "Genome-Wide Association and Epistasis Studies Unravel the Genetic Architecture of Sudden Death Syndrome Resistance in Soybean." *The Plant Journal: For Cell and Molecular Biology* 84 (6): 1124–36. https://doi.org/10.1111/tpj.13069.

Zhang, Xu, Ronald J. Hause Jr, and Justin O. Borevitz. 2012. "Natural Genetic Variation for Growth and Development Revealed by High-Throughput Phenotyping in Arabidopsis Thaliana." *G3* 2 (1): 29–34. https://doi.org/10.1534/g3.111.001487.

Zhan, J., R. E. Pettway, and B. A. McDonald. 2003. "The Global Genetic Structure of the Wheat Pathogen Mycosphaerella Graminicola Is Characterized by High Nuclear Diversity, Low Mitochondrial Diversity, Regular Recombination, and Gene Flow." *Fungal Genetics and Biology: FG & B* 38 (3): 286–97. https://doi.org/10.1016/s1087-1845(02)00538-8.

Zhong, Ziming, Thierry C. Marcel, Fanny E. Hartmann, Xin Ma, Clémence Plissonneau, Marcello Zala, Aurélie Ducasse, et al. 2017. "A Small Secreted Protein in Zymoseptoria Tritici Is Responsible for Avirulence on Wheat Cultivars Carrying the Stb6 Resistance Gene." *The New Phytologist* 214 (2): 619–31. https://doi.org/10.1111/nph.14434.

Zhou, Qing-Bo, Qiang-Yi Yu, Jia Liu, Wen-Bin Wu, and Hua-Jun Tang. 2017. "Perspective of Chinese GF-1 High-Resolution Satellite Data in Agricultural Remote Sensing Monitoring." *Journal of Integrative Agriculture* 16 (2): 242–51. https://doi.org/10.1016/S2095-3119(16)61479-X.

General Discussion and perspective

In this doctoral thesis, I used population genomics to study the lifestyle and evolution of virulence in *Zymoseptoria tritici* (syn *Mycosphaerella graminicol*a), the most damaging fungal pathogen of wheat in Europe. I first analyzed whole genome sequences of a large set of *Z. tritici* sampled from multiple genetically distinct wheat cultivars replicated in space at three time points of the growing season **(chapter 1)**. Population genomic analysis revealed a high genetic diversity at a single field level with a rapid decay in linkage disequilibrium. In approximately 80% of the population, found one or more accessory chromosome to be missing. The population diversity resulted from sexual recombination and a minor degree of clonality was most likely driven by splash dispersal of spores among adjacent wheat plants. The majority of clonal genotypes were associated with a specific cultivar CH Combin indicating a successful genotype rising in frequency. This chapter retraced the impact of sexual and asexual reproduction on genetic diversity in a single field.

The interplay of sexual and asexual reproduction has important consequences for short-term adaptation, as recombination enables the reshuffling of allelic variation. In **chapter 2,** I used genome-wide association mapping to identify genetic variation linked to virulence of *Z. tritici* on a commonly used wheat cultivar Claro. The genetic basis of virulence was complex, with the strongest association being intergenic and flanked by genes encoding a predicted effector and a serine-type protease. The intergenic region was highly dynamic, consisting of multiple families of transposable elements. Large insertion and deletion events likely driven by TEs generated length variation between the flanking genes by a factor of seven. The TE dynamics also extended in the previously established global collection of *Z. tritici*. The associated locus also showed strong signatures of repeat-induced point mutation (RIP), contributing to the

locus's rapid diversification. This chapter highlights the power of combining GWAS and population-scale genome analyses to investigate major effect loci in pathogens.

Chapter 3 looks at genetic basis of metabolites production variation in *Z. tritici* population. I focused on a panel of 102 isolates showing considerable genetic diversity. GWA mapping of untargeted metabolite profile revealed significantly associated loci in proximity to genes encoding functions related to endosomal functions and transporters. This study provides a significant resource to unravel polymorphism underlying metabolome diversity within a species.

Recombination as a weapon stay ahead in the fight with host defences

Meiotic recombination is an essential process in sexual fungal pathogen that shapes the genetic variation in populations and direct adaptive evolution. Genome plasticity and standing genetic variations generated by recombination allow the pathogen to adapt to selection pressure introduced by changes in the host resistance and artificial or natural environmental change (McDonald & Linde, 2002). Studies have shown that recombination plays a critical role in defining genomic architecture such as nucleotide variations, GC content or transposable element distribution (Muszewska *et al.*, 2017, 2019). Only recently, population genomics facilitated by whole-genome data has been able to correlate the history of recombination to the emergence of new pathogens and adaptive mechanism in many species (Burdon *et al.*, 2016). It has indeed been shown in a previous study on *Z. tritici* that elimination of the sexual stage (recombination) can reduce the effective population size (McDonald & Mundt, 2016). Recombination also breaks up the linkage between alleles, aiding the pathogen to create novel genotypes **(Chapter 1)** and phenotypic diversity **(Chapter 2-3)**. However, the recombination rate can be varying across the genome (Croll *et al.*, 2015), as we observe in our study and can

lead to loss or duplication of non-essential chromosomes **(Chapter 1)**. Our association mapping study on virulence and specialized metabolites exploited the genetic diversity and rapid decay in linkage to detect the genetic basis of phenotypic variations **(Chapter 2-3)**. Combining single field population genomics and association mapping using whole-genome sequencing in this PhD thesis is a unique approach. The population genomics approach helps to understand the genome-wide contribution to diversity and predict evolutionary responses to changing environmental conditions.

Studying the spatial and temporal changes in pathogen population

Most of the population genomics studies of fungal plant pathogen have sampling across large geographical scale (across countries and continents), considering space as a reasonable proxy for variation through time. However, short-term response to biotic and abiotic factors likely involves adaptive changes at the scale of agricultural fields and throughout crop growing seasons. Additionally, global population analysis can suffer from biased sample selection or low sample numbers (within each sub-population). Population genomic analyses with samplings on large spatial scales can only provide details about the evolution of pathogen lineages and evolutionary events from the distant past. However, deeply sampling a single field population can additionally help understand "evolution in action" of genetic and phenotypic diversity on a much finer scale. Population genomic analysis of a single field population of *Z. tritici* revealed genetic diversity in a single field exceeding global diversity levels for many other pathogens. The population maintains a vast genetic pool by sexual reproduction throughout the epidemic **(Chapter 1)**. It is evident that even at a smaller spatio-temporal scale the pathogen is generating a repository of standing genetic variations which could fuel rapid evolution. Exact contributions of sexual and asexual cycles towards standing genetic variation individually still need to be explored though.

Fungi use both clonal and sexual reproduction depending on life stage, environmental conditions, substrate, and nutrient availability (Drenth *et al.*, 2019). Although microsatellite-based study is a reliable to detect reproduction mechanism, whole genome enables to explore the degree of genetic difference in a population. In my theis, we found that non-clonal isolates were genetically different in > 15% of the analysed whole genome SNPs **(Chapter 1)**. The mixture of reproduction mode seems also be the deciding factor in producing spore structure optimal for distribution on spatial scale at different time point of the epidemic (Philibert *et al.*, 2011; Drenth *et al.*, 2019). Hence, maintaining a mixed reproduction during the epidemic constitutes an important predictor of future epidemic and possible emerging adaptations to new environmental conditions. Sexual recombination in the pathogen plays a significant role in the purging of deleterious mutations. But this raises the question that weather sexual reproduction facilitates specific genetic architecture for different phenotypic traits and whether these genetic architectures are population specific.

New insight into the genetic architecture for virulence

Genome-wide association studies allows to map heritable genetic variation for a specific phenotype. However, it does not test whether the associated genetic variation is under selection in populations (Pino Del Carpio *et al.*, 2018). The result of mapping study on virulence identified locus rich in transposable elements and is flanked by genes encoding an effector candidate and a serine type protease **(chapter 2)**. Analysis of the same locus on global population showed that dynamics of TEs at the virulence loci found in a single field population also expanded to worldwide population. Effector genes and specialized metabolite gene clusters are often located in the most repetitive regions of the chromosome (Wang *et al.*, 2017; Torres *et al.*, 2020). The proximity to repetitive regions, in particular transposable elements (TEs), increases the likelihood for sequence rearrangements to occur. Indeed, it has been

demonstrated that TEs are de-repressed under stress conditions and in turn upregulate expression of effector genes (Fouché *et al.*, 2020). The de-repressed TEs can also expand the copy numbers in other de-repressed regions during infection cycle of the pathogen. This phenomenon could be one of drivers of genome compartmentalization also known as "two-speed genome" where genes necessary for virulence or specialized metabolites always cluster in close proximity to transposable elements. However, it is still unknown if the TE dynamics in the genome is regulated to counter selection pressure or it is just a random process.

The repeat rich loci associated with virulence revealed presence-absence variation at the TE family level **(Chapter 2)**. The specific TE families present in all isolates (e.g. DTX_MITES_Addanc and RSX_SINE_Reikon) could have been fixed in the region and facilitate further accumulation of TEs in the region. These nested repeat regions may act as a 'breeding ground' where new pathogenicity associated variations can emerge. Another important factor to consider in the TE mediated genome compartmentalization is the role of reproduction mode. Sexual reproduction and hence resultant recombination events have an impact the genomic accumulation of TEs (Kent *et al.*, 2017). Recombination events can facilitate accumulation of highly active TEs or even purge the deleterious ones (Arkhipova & Meselson, 2005). However, studies on *Saccharomyces cerevisiae* the impact of sexual or asexual reproduction has been very contradicting (Ågren *et al.*, 2015; Bast *et al.*, 2019; Chen & Zhang, 2021). It is still unknow how recombination and TE dynamics in fungal pathogens are corelated. Additionally, repeat induced point (RIP) mutations, a mechanism to control the invasion of TEs, is highly active during sexual reproduction stage (Gardiner *et al.*, 2020). Presence of TEs close to virulence related genes can impact sequence composition through RIP or transcriptional regulation (Fouché *et al.*, 2018). Similar possibilities could be a reason for retaining only the beneficial TEs and gene mutations leading to overall gain in fitness.

Identifying the causality of neutral versus adaptive effect of repeat rich region could be achieved with in future with analysis of whole genome analysis.

This thesis highlights the potential of whole genome sequencing of a single field population in understanding the lifestyle and evolution of virulence for an important plant pathogen. Establishing a complete picture of the genomic basis of adaptation in a pathogen population is key to disease management. Most population genomics studies attempt to infer evolutionary dynamics via distant spatial sampling strategies. This strategy, however, does not allow to understand the diversity and evolutionary potential at a local scale. This thesis using deep sequencing of a single field establishes that the pathogen is evolving also at a shorter time scale. In the future, multi-year whole genome analyses will reveal the genotype turnover and change in population structure from one epidemic season to the other, thus helping in the timely application of protective measures. Such sampling strategies could also be useful for testing more specific hypotheses of pathogen adaptation, e.g. adaptation to specific wheat cultivars or fungicides. The method could also provide essential insights into the genomic and eco-evolutionary processes underlying the emergence and adaptation mechanism of obligate clonal pathogens.

Association mapping, similar to pathogen population genomic studies, also focused on collections across a large geographical scale (across countries and continents) and, therefore, looked at evolutionary dynamics from ancient times. However, fungal pathogens adapt to new selection pressures within single seasons. In this thesis, we set out to analyze these short-term dynamics systematically. Specifically, we demonstrate that genomic and metabolomic GWAS of a single field can be used to identify candidate loci for virulence traits. Transcriptional results

from a previous study encourage to further investigate the effector candidate found at the virulence locus.

We also established one of the first studies to standardize metabolic GWAS for a fungal plant pathogen. A more systematic analysis and functional validation of the associated loci would reveal the actual genetic mechanism of the metabolite variations. This can be exploited to design synthetic compounds to specifically target the metabolic pathways essential for pathogen growth. Including virulence phenotypes and metabolome profiles from different conditions will help to identify more potential effectors.

Our study system validates that monitoring of short-term evolutionary changes, as shown here, will become powerful tools to design more durable strategies against crop diseases.

Reference

Ågren JA, Greiner S, Johnson MTJ, Wright SI. 2015. No evidence that sex and transposable elements drive genome size variation in evening primroses. *Evolution* 69: 1053–1062.

Arkhipova I, Meselson M. 2005. Deleterious transposable elements and the extinction of asexuals. *BioEssays* 27: 76–85.

Bast J, Jaron KS, Schuseil D, Roze D, Schwander T. 2019. Asexual reproduction reduces transposable element load in experimental yeast populations. *eLife* 8.

Burdon JJ, Zhan J, Barrett LG, Papaix J, Thrall PH. 2016. Addressing the challenges of pathogen evolution on the world's arable crops. *Phytopathology* 106: 1117–1127.

Chen P, Zhang J. 2021. Asexual experimental evolution of yeast does not curtail transposable elements (S Wright, Ed.). *Molecular Biology and Evolution*.

Croll D, Lendenmann MH, Stewart E, McDonald BA. 2015. The Impact of Recombination Hotspots on Genome Evolution of a Fungal Plant Pathogen. *Genetics* 201: 1213–28.

Drenth A, McTaggart AR, Wingfield BD. 2019. Fungal clones win the battle, but recombination wins the war. *IMA Fungus* 10: 18.

Fouché S, Badet T, Oggenfuss U, Plissonneau C, Francisco CS, Croll D. 2020. Stress-Driven Transposable Element De-repression Dynamics and Virulence Evolution in a Fungal Pathogen (I Arkhipova, Ed.). *Molecular Biology and Evolution* 37: 221–239.

Fouché S, Mence Plissonneau C, Croll D. 2018. The birth and death of effectors in rapidly evolving filamentous pathogen genomes.

Gardiner DM, Rusu A, Barrett L, Hunter GC, Kazan K. 2020. Can natural gene drives be part of future fungal pathogen control strategies in plants? *New Phytologist* 228: 1431–1439.

Kent T V., Uzunović J, Wright SI. 2017. Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372: 20160458.

McDonald BA, Linde C. 2002. PATHOGEN POPULATION GENETICS, EVOLUTIONARY POTENTIAL, AND DURABLE RESISTANCE. *Annual Review of Phytopathology* 40: 349–379.

McDonald BA, Mundt CC. 2016. How Knowledge of Pathogen Population Biology Informs Management of Septoria Tritici Blotch. *Phytopathology* 106: 948–955.

Muszewska A, Steczkiewicz K, Stepniewska-Dziubinska M, Ginalski K. 2017. Cut-and-Paste Transposons in Fungi with Diverse Lifestyles. *Genome Biology and Evolution* 9: 3463–3477.

Muszewska A, Steczkiewicz K, Stepniewska-Dziubinska M, Ginalski K. 2019. Transposable elements contribute to fungal genes and impact fungal lifestyle. *Scientific Reports* 9: 4307.

Philibert A, Desprez-Loustau ML, Fabre B, Frey P, Halkett F, Husson C, Lung-Escarmant B, Marçais B, Robin C, Vacher C, *et al.* 2011. Predicting invasion success of forest pathogenic fungi from species traits. *Journal of Applied Ecology* 48: 1381–1390.

Pino Del Carpio D, Lozano R, Wolfe MD, Jannink J-L. 2018. Genome-Wide Association Studies and Heritability Estimation in the Functional Genomics Era. In: Springer, Cham, 361–425.

Torres DE, Oggenfuss U, Croll D, Seidl MF. 2020. Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fungal Biology Reviews* 34: 136–143.

Wang Q, Jiang C, Wang C, Chen C, Xu J-R, Liu H. 2017. Characterization of the Two-Speed Subgenomes of Fusarium graminearum Reveals the Fast-Speed Subgenome Specialized for Adaption and Infection. *Frontiers in Plant Science* 8.

Acknowledgements

I am sincerely grateful to all the people who contributed to my and supported me throughout my PhD journey. In particular, I would like to thank my supervisor Prof. Daniel Croll who has been an exceptional mentor and an inspiration. Thank you for always being available to discuss and help. You helped me beyond everything to evolve from wet lab microbiology to population genomics. I would like to thank Prof. Pilar Junior and Dr. Remco Stam for agreeing to be a part of my PhD committee.

My sincere thanks to everyone who contributed to my research. I would like to thank Dr. Petteri Karisto for contributing to my first chapter. I would also like to thank Leen Abraham and Dr. Thomas Badet for providing the data and analysis of the second chapter. Thanks to Dr. Gaetan Glauser for all the support in performing and analyzing the metabolomics experiments. It was an exciting experience to work with Anik Dutta and Guido Pucceti on the review article. I would like to thank all the people who took out time to read my manuscripts and provide constructive comments.

I cannot thank enough all the members of Laboratory of Evolutionary Genetics. With your constant support and help my journey would not been so enjoyable. I would really miss all our barbeques and crazy time at cerf among other things. Special thanks to Toto, Leen, Luzia and Sabina for being for being like a family and all the fun time. A big thanks to my office partner Ursula for all the science and non-science discussions we had. Also, I want to thank a former member of the lab, Simone for always motivating people to do some fun activities.

Lastly, but also most importantly, I want to thank all my family. Without your support and confidence in I would not have been able to come so far in my career and evolve as a person.