Exploration of the potential of high throughput DNA sequencing to quantify fungicide resistance in populations of *Zymoseptoria tritici*

Yaiza Gutiérrez Vázquez

Thesis is submitted to the University of Newcastle

for the degree of Doctor of Philosophy

November, 2021

Fera, Sand Hutton, UK, YO41 1LZ

School of Natural and Environmental Sciences, Newcastle University, NE1 7RU

Abstract

Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* is the most damaging foliar disease of wheat in Northern Europe, causing up to 50% yield loss each year. Common STB control strategies include the application of fungicides. However, intensive fungicide use has driven a selection for resistant isolates and some fungicides have as a result become less effective, or in extreme cases totally ineffective. Moreover, in the last decades there has been a reduction in the availability of fungicides to control this disease due to regulation processes and slow development of new active ingredients, which creates a threat for the EU wheat production. There are three fungicidal modes of action that are commonly used for STB: multi-site inhibitors (MSI), succinate dehydrogenase inhibitors (SDHIs) and demethylation inhibitors (DMIs, usually referred as azoles).

Traditionally, screening of pathogen populations require the isolation and culturing of the pathogen, which is slow and expensive. Here, a different strategy based on sequencing using Oxford Nanopore Technologies (ONT) was explored to understand the haplotype and incidence of resistance within populations of *Z. tritici*. The method was optimized to enable rapid sequencing of fungicide target genes for isolates, populations, and mock communities. The aim was to facilitate a full, rapid assessment of existing resistance status leading to more informed decision-making regarding fungicide use by farmers, such as minimising the risk of further resistant strains.

To understand the genetic mechanisms that underly fungicide resistance, the sensitivity of 49 isolates of *Z. tritici* to four widely used azoles was tested. The level of resistance was determined in culture and linked to the fungicide resistance mechanisms, which included mutations on *CYP51*, insertions within its promoter region and overexpression of the ABC and MFS transporters. Findings showed that the analysis of all three mechanisms, is required for a better understanding of the phenotype. Moreover, the haplotype of *CYP51* and the four *SDH* subunits was tested to determine how the population structure changed throughout the season after spraying single active ingredients.

Acknowledgements

This PhD was funded by the Institute for Agrifood Research Innovation (IAFRI) a joint institute owned by the Newcastle University and Fera Science Ltd. and the contribution of Agrii who allowed access to their field trial programme to facilitate the study.

I would like to thank many people who helped me during this journey:

Firstly, to my supervisors Prof. Neil Boonham, Dr. Roy Sanderson, Dr. Femke van den Berg and David Langton for all their advice and support throughout this PhD. For giving me the opportunity to grow as a researcher and letting me make my own decisions and mistakes. For your patience, for guiding me when I needed it and also letting me know when I was losing track.

Secondly, I would also like to thank the team at Fera Science Ltd. for all their guidance and in particular Catherine Harrison and Lewis Ridley for their lab support and guidance. Also, a very special thank you to the NGS team at Fera, Ian Adams, John Walshaw and especially Sam McGreig who introduced me to bioinformatics, my new passion, and for guiding me thoughout this learning process.

Lastly, but not least, to my family and friends who supported me on each of the steps and never questioned my absence during the last pandemic year. Espeacially Sara, who took the time to read my thesis and for being that extra pair of eyes who could see what I could not. Thanks to my family, who always believe in me. A special thanks to Chris for putting up with me when I could not put up with myself.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABC: ATP-binding cassette

AHDB: Agriculture and Horticulture Development Board

CE: Control environment

ddNTP: dideoxy-nucleotide

ddPCR: Droplet Digital PCR

DMI: Demethylation inhibitor

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

dNTP: deoxy-nucleotide

dsDNA: double strand DNA

EC_{50}: Effective Concentration 50

EU: European Union

FAO: Food and Agriculture Organization of the United Nations

FRAC: Fungicide Resistance Action Committee

GHG: Glasshouse emissions

GITC: Guanidinium Thiocyanate

GLM: Generalized Linear Model

GPM: General purpose media

GPU: Graphics processing unit

GS: Growth stage

HAD: Healthy area duration

HGCA: Home Grown Cereals Authority

HRM: High Resolution Melting

HTS: High Throughput Sequencing

IGV: Integrative Genomics Viewer

IPM: Integrated Pest Management

IUPAC: International Union of Pure and Applied Chemistry

KASP: Kompetitive Allele Specific PCR

Kb: kilobases

LAMP: loop mediated isothermal amplification

MBC: Methyl Benzimidazole Carbamates

MDR: Multidrug resistance

MFS: Major Facilitator Superfamily

ml: mililitre

MoA: Mode of Action

MSI: Multi-site inhibitors

N: Nitrogen

NMDS: Non-metric multidimensional scaling

NTSR: Non-target site resistance

nt: nucleotide

ONT: Oxford Nanopore Technologies

PBS: Phosphate-buffered saline

PCA: Principal Components Analysis

PCD: Programmed cell death

PCR: Polymerase chain reaction

PDA: Potato dextrose agar

PDR: Pleiotropic drug resistance

PGM: Personal Genome Machine

ppm: part per million

QoI: Quinone outside inhibitors

qPCR: quantitative P

RF: Resistance Factor

RT-qPCR: Quantitative reverse transcription PCR

SDH : Succinate Dehydrogenase

SDHI: Succinate Dehydrogenase Inhibitors

SMRT: single-molecule real-time

SNP: Single nucleotide polymorphism

STB: Septoria tritici blotch

TSM: Target site mutation

TSO: Target site overexpression

TSR: Target site resistance

Chapter 1. General Introduction

1.1 Background

It is estimated that in the next 30 years we will need to produce 50% more food to feed the rapidly increasing world population (FAO, 2018). For instance, the production of staple crops such as wheat should increase as it is the most widely grown cereal crop in Europe, producing over 22 million ha in 2020 (http://epp.eurostat.ec.europa.eu). Every year severe attacks of leaf diseases in winter wheat give rise to significant and economically important losses which can be up to 50% of yield (Oerke, 2005; Jørgensen *et al.*, 2014). As a consequence of the yield losses, an excess of Nitrogen (N) is generated in the field, caused by the applied Nitrogen not being utilised, which negatively impacts the environment. The excess of Nitrogen can be lost through leaching, volatilization or biological processes, contributing to glasshouse emissions (GHG) (Bajgai *et al.*, 2019). Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* (anamorph, *Septoria tritici* or *Mycosphaerella graminicola*) is seen as the most serious leaf disease in Northern Europe (Fones and Gurr, 2015). The economic value of *Z. tritici* yield losses in the UK were calculated to be between £110 and £220 million each year (Fones and Gurr, 2015), with the final figure depending on the severity of infection.

1.2 Strategies to control *Zymoseptoria tritici*

Several Integrated Pest Management (IPM) strategies can be used to help with *Z. tritici* control. IPM strategies prioritize the use of biological, physical and non-chemical methods and minimise the use of fungicides. These strategies include: crop rotation, soil tillage, sowing time, use of resistant cultivars, seed treatment and strategic fungicide applications.

1.2.1 Crop rotation, soil tillage and sowing time

Crop rotation and soil tillage are used to control leaf diseases in wheat. Crop rotation benefits include improved soil structure, nutrient management and decreases pest populations by breaking their reproductive cycle (Hennessy, 2006) and subsequently reduces the need of pesticides. Crop rotation primarily benefits the control for soil-borne diseases, with the reduced benefit for air borne diseases due to the ability of the inoculum to travel long distances. On STB control, crop rotation and soil tillage aim at limiting the primary infection caused by pycnidiospores which can remain dormant for

long periods of times in the crop debris (Suffert *et al.*, 2011). Some studies have demonstrated that the STB incidence was decreased by 3-5 year crop rotation (Eyal 1987). An experiment developed by Gladders *et al.* (2001) also demonstrated that a later sowing date reduced disease severity. For example, the disease risk was reduced in almost 5% of the fields if the winter wheat was sown in October when compared to September and another additional 5% if sown in November.

IPM strategies should be considered for all pests and diseases, not only focusing on a single disease. The IPM strategies presented above can aid with STB control, some of these strategies may help to control other diseases while others will not. For example, in the wheat and barley disease management guide, the Agriculture and Horticulture Development Board (AHDB) reports that early sowing has a high impact on the disease risk for septoria tritici, yellow rust and brown rust but a low impact on powdery mildew (https://ahdb.org.uk/knowledge-library/wheat-and-barley-disease-management-guide).

1.2.2 Use of resistant cultivars

Alongside the agronomic practices, the use of STB resistant wheat cultivars plays a key role in the control of *Z. tritici* and preventing yield losses. The creation of resistant cultivars requires the finding of novel STB resistant genes or alleles. STB resistance can be classified depending on the cultivar-isolate interaction. Quantitative resistance happens when STB is controlled by the effect of several loci, while in qualitative resistance it is controlled by a single major gene (Brown *et al.*, 2015). Currently, a total of 21 qualitative resistant genes and 167 genomic regions providing quantitative resistance have been reported (Brown *et al.*, 2015).

It has also been demonstrated that cultivar mixtures have the ability to reduce fungal diseases in cereal crops (Borg *et al.*, 2018) and can reduce the severity of STB (Ben M'Barek *et al.*, 2020). Kristoffersen *et al.* (2020) reported that cultivar mixtures reduce the risk that *Z. tritici* populations develop fungicide resistance mechanisms and consequently this strategy can help in reducing fungicide applications. However, host resistance to *Z. tritici* might be compromised with the susceptibility of the host to other pathogens (Brown, 2002).

In addition, STB control through the use of resistant cultivars is affected by the rapid adaptation of the pathogen to the environment caused by the high evolutionary

potential of *Z.tritici* (Leroux and Walker, 2011; Steven Kildea, 2021) which irrevocably leads to the use of fungicides to prevent yield loss.

1.2.3 Use of fungicides

Fungicides play a key role in the control of STB and these can be applied as seed or foliar treatments. Whilst seed treatments are not normally used to target specific pathogens, most of them can decrease fungal pathogen levels by suppressing the sporulation for over 6 weeks after sowing (Sundin *et al.*, 1999). In contrast, foliar fungicides are widely used to target specific pathogens.

In order to effectively control *Z. tritici* the timing of fungicide applications is important, as spraying too early or too late will give poorer disease control. For winter wheat crops, yield is correlated to the total green leaf area of the top five leaves and therefore, spray timings should aim to optimise the healthy area duration (HAD) of these leaf layers. Wheat leaves are numbered from the top to the bottom, for a mature plant the top leaf (flag leaf) is referred to as “leaf 1” and the bottom leaf is referred to as “leaf 5” (Figure 1.1).

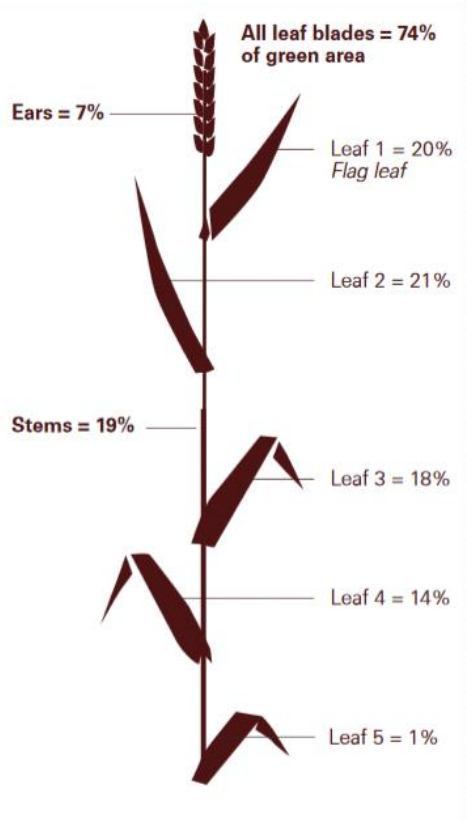


Figure 1.1 Diagram of a 'mature' plant numbering system for leaves, the percentage indicates the distribution of green area at flowering (HGCA, 2008)

AHDB recommend up to four spray timings to protect crops from different fungi. These spray are determined by the growth stage (GS) of the plant, following the Zadoks scale (Zadoks *et al.*, 1974). The T0 spray is generally applied when leaf 4 has emerged (GS30); the T1 spray when leaf 3 has emerged (GS31-32; targeting the recently emerged leaf 3 and sometimes diseases on leaf 4), the T2 spray when the flag leaf (leaf 1) has emerged (GS37-39) and the T3 spray at ear emergence (GS59-61; targetting ear diseases and providing additional control of disease on the top two leaves). The main sprays to control *Z. tritici* are T1 and T2, whereby T1 is aimed at controlling *Z. tritici* on leaves 3 and 4 whilst T2 is aimed at controlling disease on the top two leaves, which contribute approximately 65% of yield. It is recommended to apply a T2 spray when most flag leaves on the main tillers have emerged. T0 and T3 are still recommended for the control of other fungi as well as contributing to *Z. tritici* control in severe disease seasons.

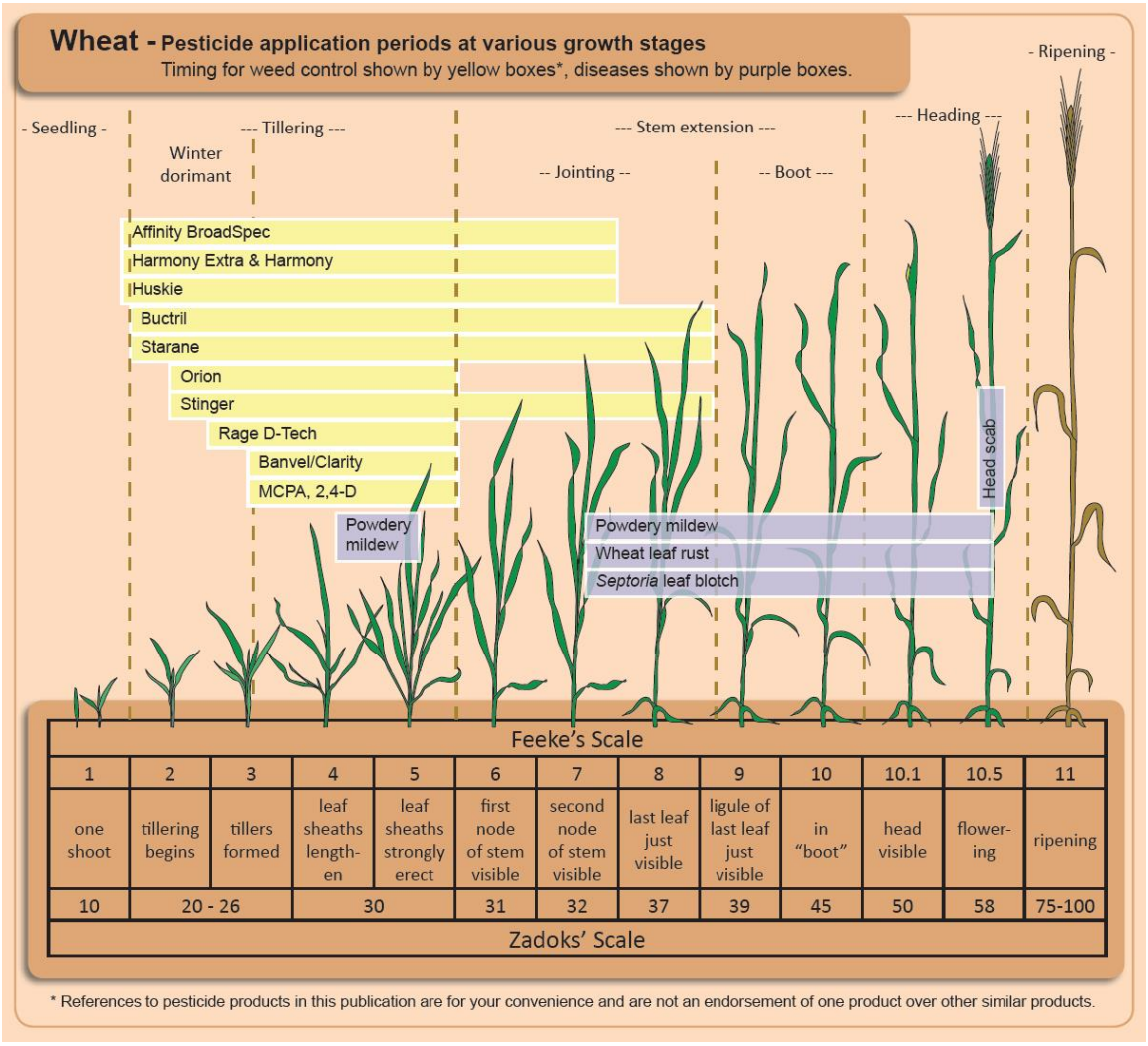


Figure 1.2 Wheat growth stages (Smith, 2013)

Fungicides can be classified according to their mode of action (MoA), which indicates the site or protein within the pathogen where the fungicide targets, or according to their efficacy. Currently there are three main fungicidal MoA available for STB control multi-site inhibitors, demethylation inhibitors (DMIs) also known as azoles and succinate dehydrogenase inhibitors (SDHIs). Moreover, we could classify these MoA according to they interact with a single site or with several sites, known as multi site fungides. If we were to classify according to their efficacy, a fungicide can provide protectant activity when it prevents the growth of the pathogen or can have a curative or eradicant activity which can directly kill the pathogen. Protectant fungicides are applied on healthy leaves, they stay on the surface of the tissue and they do not enter the plant tissue. However, curative or eradicant fungicides act in the infection, before symptoms are visible, they enter the plant tissue and stops or reduces the fungal growth (Paveley et al., 2003). Protectant and curative efficacies depend on the fungicide, some fungicides have just protectant activity while others can be better recommended by its curative efficacy or sometimes both. Fungicides with curative control are only used in the early stages of the disease as they are effective only during the early stages of the latent period. As an example, Sanssené *et al.* (2011) determined that prothioconazole had a protective efficacy which was similar on all the isolates tested but that the curative efficacy was dependant on the isolates.

1.3 Problems of fungicide control

The constant use of fungicides have the disadvantage of developing fungicide resistance and consequently making those fungicides less effective, or in extreme cases totally ineffective. Fungicide resistance occurs when a fungal pathogen can survive and reproduce in the presence of a fungicide, after a period of exposure of the pathogen to the fungicide. Moreover, in the last decades there has been a reduction in the range of fungicides available to control STB (Jonathan Blake, 2011; Jess *et al.*, 2014). This is due to lack of development of new active ingredients and the loss of fungicides due to regulation.

As van den Bosch postulated “the effective life of a fungicide is maximised by reducing the rate at which resistance develops whilst maintaining good disease control”, which is achieved by minimising the selection coefficient (van den Bosch *et al.*, 2014), by reducing the rates of increase of both resistant and sensitive strains, by reducing the

rate of increase of resistant strains in relation to sensitive strains or by reducing the time of exposure to the fungicides. Additionally, when using fungicides with different modes of action resistance development can often be limited through alternation or the use of fungicide mixtures given that most strains resistant to one component of a mixture, are likely to be sensitive to the other component, and vice versa (van den Bosch *et al.*, 2014).

Several studies have shown shifts in the sensitivity of *Z. tritici* to fungicides worldwide (Stammler *et al.*, 2008; Leroux and Walker, 2011; Allioui *et al.*, 2016; Blake *et al.*, 2017; Jørgensen *et al.*, 2020b) as a result of the intensive use of fungicides. The adaptation of *Z. tritici* to azoles has developed faster in countries such as Ireland and the UK, where there is a high disease pressure and an intensive fungicide program (Jørgensen *et al.*, 2020b). The three main resistance mechanisms that have been described as the source of azole resistance are: 1) target site mutations (TSM) in the target gene (*CYP51*) 2) target site overexpression (TSO) and 3) multi drug resistance (MDR) due to upregulation of the transporters (Roohparvar *et al.*, 2007; Omrane *et al.*, 2015).

1.3.1 The evolutionary potential of Z. tritici

The ability of *Z. tritici* to develop fungicide resistance is in part a result of its life-cycle (

Figure 1.3). *Z. tritici,* has a high evolutionary potential due to its mixed reproductive system, where it is able to reproduce sexually as well as asexually (Kema *et al.*, 1996a). Sexual reproduction plays a key role in the ability of the pathogen to develop fungicide resistance, generating new haplotypes through recombination. Additionally, it has a polycyclic lifecycle, allowing the pathogen to sporulate and start a new cycle (Zhan *et al.*, 1998). Fones and Gurr (2015) reportesthat a single spore could infect, sporulate, and spread, up to six times in the winter wheat growing season. Infection normally occurs when airborne spores (ascospores, sexual spores) land on leaves. However, pycnidiospores (sexual spores) are spread due to rain splash which help the spores to spread between the leaves. The fungus starts the initial growth of the hyphae on the leaf surface, followed by host penetration via natural openings, through the stomata. Subsequently, during its biotrophic phase, spores germinate by forming germ tubes and grow very slowly within the leaf without causing any damage or symptoms (Kema *et al.*, 1996b). After this latent period, the pathogen switches from the biotrophic phase to the necrotrophic phase, causing necrotic lesions, fungal growth spikes,

pycnidia and pycnidospores (asexual spores) (Figure 1.4) (Rudd *et al.*, 2015). The transition from biotrophic to necrotrophic stage is associated with induction of host defence responses which might have shared characteristics with host programmed cell death (PCD) and differential regulation of signalling pathways (Keon *et al.*, 2007). The latent period lasts between 14 and 28 days, the length of which depends on a range of different factors such as the host, the isolate, weather conditions and the interaction between isolates. Moreover, it can remain dormant on crop debris and re-infect plants with the arrival of warmer weather (Goodwin, 2012). Despite this, the Fungicide Resistance Action Committee (FRAC) considers *Z. tritici* as a medium risk pathogen in terms of its ability to develop resistance to fungicides. FRAC considers that *Z. tritici* resistance is not a major problem due to it being a slow process and because the control of fungicide resistant populations is still achievable. However, the limited choice of fungicides available for use on cereals already places severe pressures on winter wheat production, and if any of the available fungicides for STB were to succumb to resistance, this would exacerbate the situation.

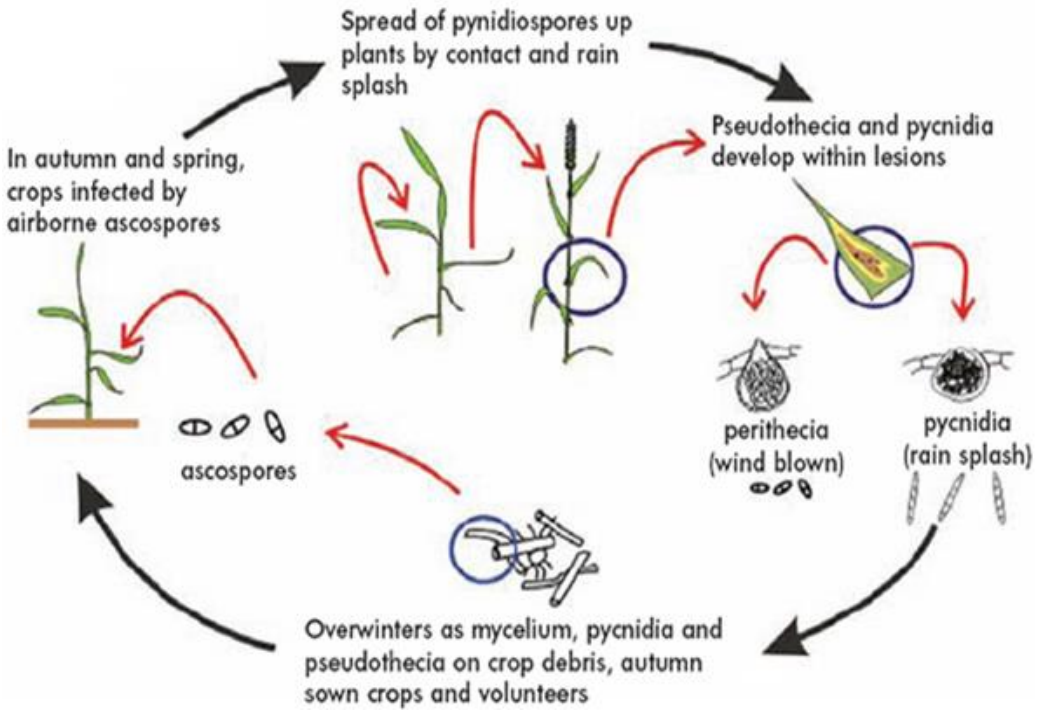


Figure 1.3 Disease cycle of Z. tritici on wheat (HGCA, 2012)

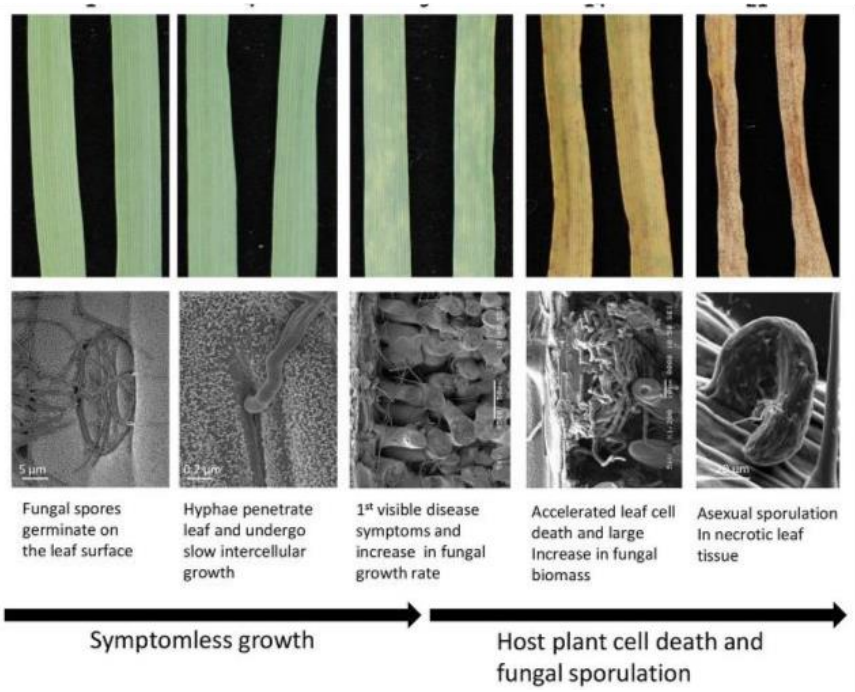


Figure 1.4. Time course of infection of Z. tritici on wheat (Rudd et al, Plant Physiology 2015)

1.4 Introduction to fungicides

In the past decades there has been several groups of fungicides with different MoA used to control STB. Multi-site fungicides were introduced in the 1950's (Russell, 2005), they are broad spectrum and have a protectant action, by impeding spore germination (Leroux P, 2005). For this reason, multi-site fungicides are recommended for use at T0, to prevent new infections and sometimes as a mixing partner with SDHIs and azoles at T1. According to FRAC multi-site fungicides are generally considered a low risk group without any signs of resistance developing. There are several MoA within the single site fungicides, for example Methyl benzimidazole carbamates (MBCs), which interfere in the mitosis by binding to the β-tubulin. Quinone outside inhibitors (QoI), targets the complex III of the cytochrome b, which is part of the respiration process. Demethylation inhibitors (DMIs) commonly known as azoles were introduced in the early 1970's. They interfere with the biosynthesis of ergosterol, the predominant sterol in fungal membranes by interacting with the haem iron in 14α-demethylase (CYP51). Lastly, succinate dehydrogenase inhibitors (SDHIs) were introduced in 2005. They interact with the Succinate Dehydrogenase by inhibiting the complex II of aeorobic respiration (Scalliet *et al.*, 2012). Succinate Dehydrogenase is

composed of four subunits, *SdhA, SdhB, SdhC* and *SdhD* (Stammler *et al.*, 2008), of which *SdhB, SdhC* and *SdhD* interact with SDHI fungicides.

1.5 Fungicide resistance mechanisms

Fungicide resistance can be caused by two different types of mechanisms, target-site resistance (TSR) and non-target-site resistance (NSTR). In TSR, a mutation or the combination of several mutations in the genes encoding the target protein of the fungicide confers fungicide resistance or decreased sensitivity of the pathogen to the fungicide. However, NTSR can be caused by different mechanisms including the overexpression of the target sites caused by rearrangement or mutations in the promoter region of the target protein or overexpression of the transporters, which decreases the fungicide concentration.

1.5.1 Target-site resistance

The first major development of resistance in *Z. tritici* to affect fungicide efficacy involved methyl benzimidazole carbamates (MBCs). In 1984 (Griffin MJ, 1985) resistance to MBCs was reported as described to be mediated by a single nucleotide polymorphism. In the same way, due to a single nucleotide polymorphism, Quinone outside inhibitors are no longer recommended for STB control. QoIs are however still effective against other diseases (Gosling P, 2016).

Currently there are three fungicidal modes of action (MoA) available for *Z. tritici* control: multi-site inhibitors, demethylation inhibitors (DMIs) commonly known as azoles and succinate dehydrogenase inhibitors (SDHIs).

Azoles, interact with the haem iron in 14α-demethylase (CYP51). More than 100 CYP51 haplotypes have been identified (Huf *et al.*, 2018). The many *CYP51* mutations which have been discovered during the past 10-15 years in different combinations have been associated with the most important changes in sensitivity. Different haplotypes of *Z. tritici*, which have been identified, can have different sensitivities to the various DMIs (Leroux *et al.*, 2007; Stammler *et al.*, 2008; Cools and Fraaije, 2013). Since resistance to QoI fungicides developed, the azoles have been seen as the backbone of STB control, but in recent years major changes in the sensitivity of the populations have been observed across Europe (Fraaije *et al.*, 2007).

Despite UK control strategies that primarily rely on the use of SDHIs fungicides, DMIs are still used in combination to SDHIs, as an anti-resistance strategy and to control other diseases. The EU still relies upon the continued effectiveness of the DMIs, which accounts for approximately 50% of the fungicide input in its wheat production, but there has been a clear decline in the protectant activity and curative efficacy of DMIs fungicides over time. Nevertheless, the four most widely used azoles (epoxiconazole, metconazole, prothioconazole and tebuconazole) still provide significant control (Jørgensen *et al.*, 2018).

Similarly, the presence of single nucleotide polymorphisms (SNP) mutations in the different subunits of this enzyme results in a gradual loss of sensitivity to SDHI fungicides.

1.5.2 Non-target site resistance

Non-target site resistance is an alternative mechanism to target-site resistance. Reduced sensitivity to azoles has also been linked to inserts in the promoter, to membrane-bound transporters and, more importantly, to intragenic recombination of *CYP51* (Leroux and Walker, 2011; Cools and Fraaije, 2013). Induction of the transcription of drug efflux transporters or detoxifiying enzymes, and regulation of stress response pathways, are some of the mechanisms involved in non-target site resistance (Hu and Chen, 2021). Figure 1.5 is a modified version of the original figure created by Hu and Chen (2021) representing the main non-target mechanisms involved in fungicide resistance which has here been adapted to fungicides used to control STB. All fungicides can be exported by the drug efflux transporters and all fungicides apart from QoIs are also affected by detoxification. For example, the enzyme glutathione S-transferase (GST) which can biotransform the fungicide chlorothalonil (Kim *et al.*, 2004). However, there are many more mechanisms involved in fungicide resistance. For example, standing genetic variations refer to the presence of alternative genes or alleles, which can be beneficial when environmental conditions change. This mechanism allows the pathogen to adapt to different conditions and therefore increases its potential to survive environmental changes (O'Donnell *et al.*, 2014). Moreover, fungicides can be detoxified by the rapid metabolism through different pathways or chemical reactions (Lu *et al.*, 2020), or molecules which can regulate the response in stress conditions.

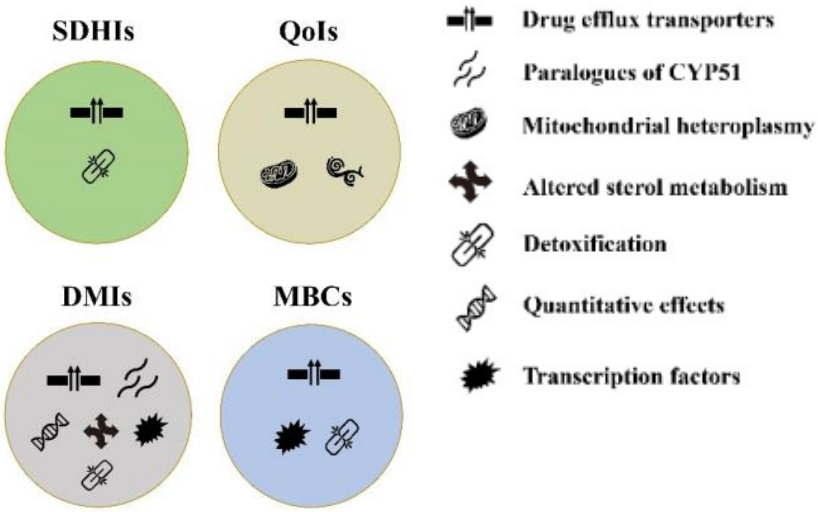


Figure 1.5 Non-target site mechanisms of resistance to fungicides used to control *Zymoseptoria tritici* (modified from Hu et al, 2021)

Each circle represent a fungicide and the icons inside the mechanisms involved in fungicide resistance. SDHIs: Succinate Dehydrogenase inhibitors; QoIs: Quinone outside inhibitors; DMIs: Demethylation inhibitors; MBCs: Methyl benzimidazole carbamates.

1.4.2.1. Drug efflux transporters

Drug efflux transporters are present in all living organisms, including both plant and animal pathogens. One of the most studied drug efflux transporters are the ATP-binding cassette transporters (ABC transporters) which are primary active transporters, capable of transporting small molecules and macromolecules upon ATP hydrolysis. Representatives of the pleiotropic drug resistance (PDR) and multidrug resistance (MDR) subfamilies of ABC transporters are well known for their role in resistance to antifungal agents. The second group of major importance, the Major Facilitator Superfamily (MFS transporters) (Del Sorbo *et al.*, 2000), are secondary carriers capable of transporting solutes in response to chemiosmotic ion gradients (Paulsen *et al.*, 1996). MFS transporters can have low substrate specificity and therefore, also play a role in sensitivity to drugs such as fungicides. Overexpression of such transporters can lead to MDR against microorganisms (de Waard and van Nistelrooy, 1979). In *Z. tritici*, the overexpression of MgMFS1 is linked to azole and SDHI resistance (Roohparvar *et al.*, 2007; Omrane *et al.*, 2015).

1.6 Analysing mechanisms involved in fungicide resistance in *Z. tritici*

Over the last decades different methods have been developed to analyse the TSR and NTSR mechanisms in *Z. tritici* involved in fungicide resistance. Methods designed to characterise TSR aim to analyse and/or sequence the target genes of the fungicides. However, the complexity of the analysis of NTRS methods is higher, as they normally require quantitative methods.

1.6.1 Conventional methods designed to analyse target-site resistance

The analysis of target-site resistance requires the analysis or sequencing of the fungicide target genes. DNA sequencing originated in the mid 1970s, during which two different techniques for directly sequencing DNA were developed; the well-known and still used Sanger chain-termination method and the almost forgotten chemical sequencing, known as Maxam-Gilbert sequencing (Maxam and Gilbert, 1977; Sanger *et al.*, 1977).

The former method was designed by Frederick Sanger, who developed a gel-based method which combined a DNA polymerase, nucleotides (dNTPs) and termination nucleotides (ddNTPs). The random addition of both types of nucleotides generated DNA fragments of different sizes, each terminated by a ddNTP. Four reactions were needed to generate fragments terminated at each of the four nucleotides, adenine (A), cytosine (C), guanine (G) and thymine (T). As each reaction only contained one of the four different ddNTPs, many different legths strand were obtained. The reaction was followed by high resolution, acrylamide gel electrophoresis to resolve the fragments that were 1nt different in size from each other and by loading adjacent lanes with the DNA fragments terminated by each nucleotide, the DNA sequence could be read. This method was improved throughout the years with the most significant modifications being replacement of radioactive labels with fluorescently-labelled nucleotides and gels replaced with acrylamide filled capillaries. These changes largely resulted in the automated process we use today, probably the most used method to analyse *Z. tritici (Leroux et al., 2007; Estep et al., 2015; McDonald et al., 2019)*.

Since the 1970s, several methods have been developed which allow the analysis of genetic material. Figure 1.6 represents a timeline of the methods developed in the last decades which involve DNA sequencing. Along with Sanger sequencing pyrosequencing has also been used to analyse *Z. tritici* (Wieczorek et al., 2015).

Pyrosequencing is based on the sequential addition of nucleotides and the detection of the luminescence of the pyrophosphate when a nucleotide is incorporated to the sequence. However, not all the methods used to analyse the resistant populations of *Z. tritici* involve sequencing. Methods such as loop mediated isothermal amplification (LAMP) (Fraaije, 2014a) which was developed by Notomi *et al.* (2000) are currently widely used. LAMP can amplify DNA with high specificity and in a very rapid manner. This technique uses multiple primers designed to create continuous loops during DNA amplification. Amplification can be visualized by the naked eye due to the presence of fluorescent dyes or by using portable devices that allow the use of this technique in the field by combining it with quick DNA extractions.

Other methods using fluorescence to analyse the genotype of the target proteins are qPCR and KASP (Taher *et al.*, 2014; Vagndorf *et al.*, 2018; Hellin *et al.*, 2020). Quantitative PCR (qPCR) is often used to quantify the amplified DNA which is monitored by the increase in fluorescence during the amplification. Kompetitive allele-specific PCR (KASP) requires three elements, a KASP assay mix, which contains allele specific tailed forward primers and a common reverse primer; a KASP mastermix, containing the elements of a convetional PCR and two fluorescent reporter cassettes for labelling which are complementary to the tailed primers and the DNA template. As in a conventional PCR, several cycles of the denaturation of the DNA template, anealling and extension will be performed, in which, the allele specific primer will amplify the DNA template, in the next cycles complementary sequences of the tailed primer will be synthesized and in further cycles fluorescent reporter which are complementary to the allele specific tailed primer will bind and generate fluorescence. Last but not least, High Resolution melting (HRM), which is based on detecting the differences in the qPCR melting curves can also be used as an alternative to sequencing methods (Curvers *et al.*, 2015). However, all these methods are low throughput and therefore they require the analysis of high number of samples, especially to resolve pathogen populations. As an example, it took over a decade and a billion dollars to complete the first human genome to be sequenced using Sanger sequencing (Adams *et al.*, 2001; Birren *et al.*, 2001).

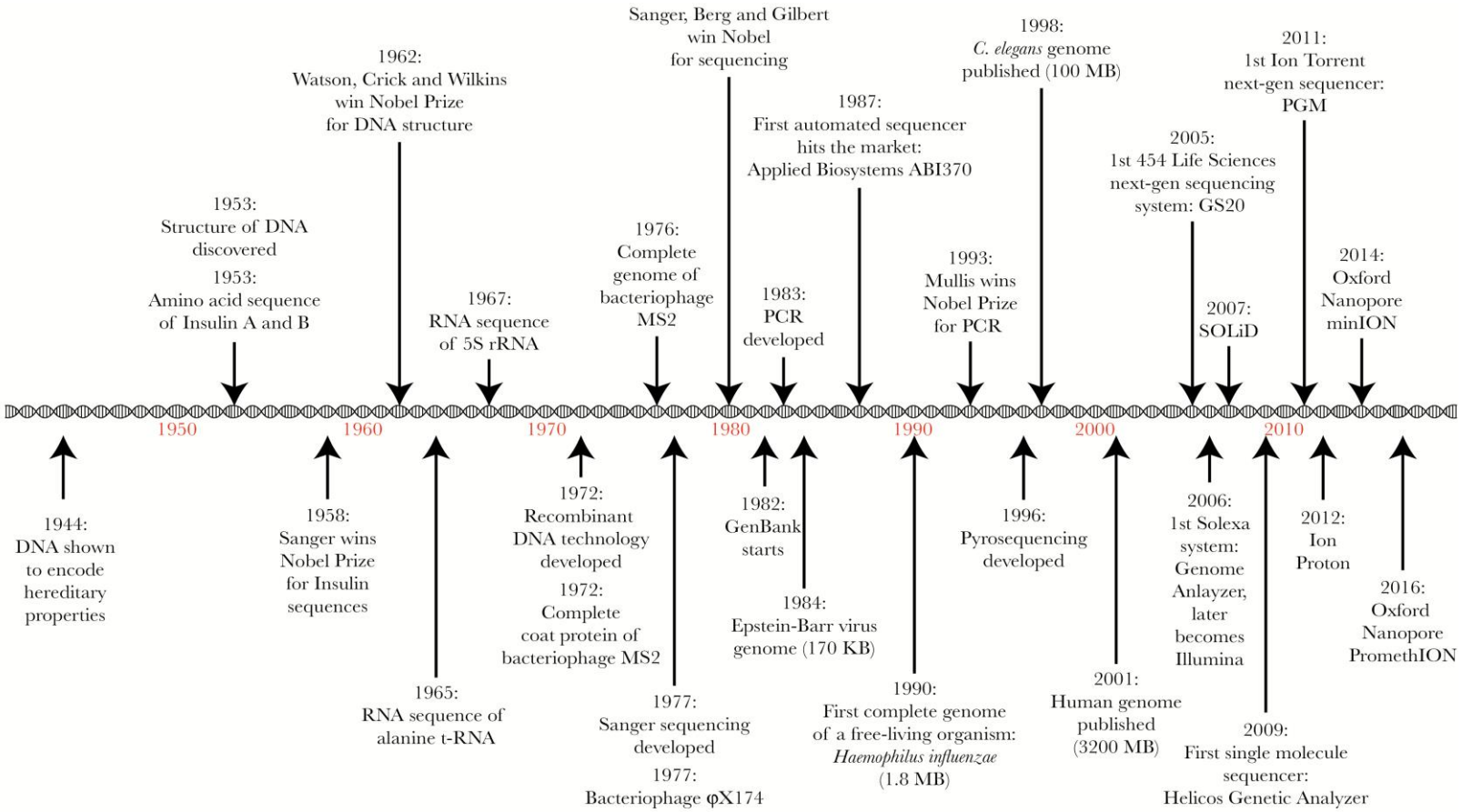


Figure 1.6 Timeline of scientific discoveries and technological advances that have enabled the rapid development of sequencing approaches (Leguia *et al.*, 2020).

DNA, deoxyribonucleic acid; PCR, polymerase chain reaction; rRNA, ribosomal ribonucleic acid.

1.6.2 Use of High Throughput Sequencing to analyse target-site resistance

In the last two decades, new techniques have been developed which are capable of sequencing multiple samples in parallel and create large data sets, also known as high throughput sequencing (HTS). Moreover, these techniques have become more affordable (Leguia *et al.*, 2020). In 2006, the first Solexa system was created, which has since become known as Illumina, the mostused HTS platform (Loch *et al.*, 2008). The concepts behind Illumina sequencing is similar to Sanger sequencing described previously. Both are based on the synthesis of the DNA and use of fluorescent nucleotides which are incorporated one by one. Each incorporated nucleotide is recognised by its fluorecescent tag. However, Illumina sequencing allows the analysis of multiple samples in parallel in three steps. Firstly, library preparation, where DNA is fragmented and adapters are added to the sample. These adapters contains regions which are complementary to the flow cell oligos, indexes and sequencing primers. Secondly, clustering, is the process where each fragment is amplified. The library is loaded into a flow cell where the DNA binds to the complementary oligo in the flow cell. Each fragment is amplified into clonal clusters through bridge amplification. Once the cluster generation is completed the read is ready to be sequenced. Thirdly, fluorescent nucleotides compete to bind to the sequence, when the nucleotide is added to the sequence it will be detected by the emitted fluorescence.

In 2011, the Ion Torrent personal genome machine (PGM) was released (Thermo Fischer Scientific). Ion Torrent PGM uses a semiconductor chip with over a million wells which capture chemical information, pH, from the DNA sequencing and translate it to digital information or basecalls. Ion Torrent measures the generated voltage caused by the release of an hydrogen ion when a nucleotide is added to the sequence (Bragg *et al.*, 2013). A year later, Ion Torrent Proton was released, which achieves faster sequencing and higher throughput, as a result of a chip with over half a billion wells (Cao *et al.*, 2017).

Although these techniques provide high amounts of data in less time than conventional techniques, they also require of the isolation of the pathogen given that they cannot be used to sequence the full length of the target genes.

In recent years, two different sequencing platforms which allow the sequencing of long fragments were developed, Pacific Biosciences (PacBio) and Oxford Nanopore

Technologies (ONT). These two sequencing methods have the potential to sequence the full length of the target genes (over 3 kb in the case of *CYP51* including its promoter region and over 2 kb in the case of the biggest *SDH* subunit) and therefore have the potential to sequence populations without the need for pathogen isolation. As a result, these two methods can be used to greatly decrease the diagnostic time between the sample collection and results, which can help farmer to choose the best spray programme to control STB. The use of ONT sequencing has the potential advantage that it can be done more rapidly and inexpensively than PacBio sequencing. However, ONT is reported to have high error rates which may prove problematic for resolving SNPs associated with TSR**.** For example, Samils *et al.* (2021) used PacBio to analyse *CYP 51, Sdh B, C* and *D* subunits and *cytochromer b*, main target proteins of azoles, SDHIs and QoIs respectively. However, this experiment used single isolates and therefore the time between the sample collection and the decision making was not reduced.

1.7 Oxford Nanopore Technologies

Oxford Nanopore Technologies (ONT) developed the first hightroughput portable sequencer, the MinION in 2014. ONT uses nanopores, biological proteins which can be found in the environment. Nanopores are embedded in an electrically resistant polymer membrane and as the DNA strand goes through the pore causes a characteristic disruption in the electric current. Each nucleotide base is distinguished by the electric change caused as the DNA passes through the pore. One of the main challenges when using nanopore sequencing is the high error rate. There are two types of errors, random error and recurrent errors. Random errors are not generally an issue, as they do not have a consistent bias, however recurrent errors can be confused with SNPs. The medaka software documentation demonstrated that there is a strand bias https://nanoporetech.github.io/medaka/future.html, which is probably caused by the different patterns in the sequence (Figure 1.7). Figure 1.7 represents the same region of a sequence, on the left the sequence for the forward strand (AACTCTG) and on the right the reverse complement (CAGAGTT) and shows that the pattern in the electric current is different between the forward and the reverse strand.

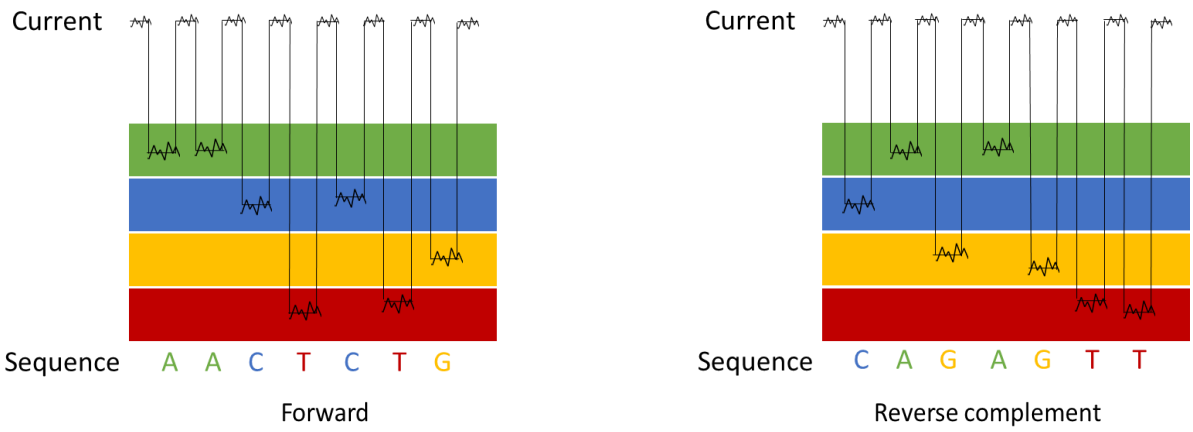


Figure 1.7 Graphical representation of how differences in the electric current can be translated to DNA sequences.

Both pictures represent the same region of the genome, in the left panel the current is measured whilst sequencing forward and on the right panel the current is measured across the same region whilst sequencing with the reverse primer (i.e. across the reverse complement strand of the forward sequence). In addition, the platform has been improved in the recent years, with the introduction of new pores, new basecallers and new chemistries

1.8 Methods designed to analyse non-target resistance

Few studies have determined the mechanisms involved in NTSR. Among the tests performed to analyse NTSR in *Z. tritici* is screening using tolnaftate. Tolnaftate is a fungicide which has been used to indicate the overexpression of transporters (Yamashita and Fraaije, 2018). However, this test does not specify which transporter has been overexpressed. The main techniques used to visualize changes in the gene expression is qPCR which also has be used to quantify the expression of the transporter MgMFS1, the main transporter studied in *Z. tritici* (Roohparvar *et al.*, 2007). A similar technique to qPCR could be used for gene expression experiments is Droplet Digital PCR (ddPCR) (Clarner *et al.*, 2021). This technique is based on the use of water-oil droplets. Individual fragments of DNA separate and amplify in each of the droplets, as a result, the positive droplets can be counted making this technique more accurate, especially for enumerating low copy numbers of DNA.

1.9 Aims and objectives

Zymoseptoria tritici is the causal agent of Septoria leaf blotch, the most important disease affecting wheat crops. The value of these losses can be up to 50% of the yield. Azoles and SDHIs are the main groups of fungicides, with different modes of action,

used to control this disease and the efficacy of these fungicides has decreased in recent years due to resistance development. The many mutations in the target genes, upstream regulation and the overexpression of transporters are known as the main resistance mechanisms.

The aim of the work is to understand the genetic structure of field populations of Septoria and our hypothesis is that by understanding the population structure we will be able to modify spray programmes in real-time to achieve better control. Analysing the main three mechanisms involved in resistance could give us an idea of which product would be the best one to spray to provide good disease control and a low selection pressure. Septoria genotypes and upstream regulation will be analysed using MinIon assays. MinIon creates reads long enough to encompass the whole gene, facilitating the determination of genotypes at a population level.

The research was structured into three objectives, which make up the three experimental chapters as follows:

Objective 1: To develop and validate the use of long-read, nanopore sequencing and analysis methods to enable us to assess haplotypes present within populations of Septoria.

Objective 2: To understand the relationship between the *CYP51* haplotype and the phenotype. This research included the analysis of the azole target gene, including upstream promoter regions and the screening for overexpression of the transporters.

Objective 3: To determine how the *Z. tritici* population changes after the use of single fungicides. Samples were collected prior to each treatment to establish the population structure and the changes in the population after each spray.

Chapter 2. Nanopore sequencing for the identification of *Zymoseptoria tritici* populations

2.1 Introduction

Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* is the most damaging foliar disease of wheat in Northern Europe (Fones and Gurr, 2015) causing yield losses of between 10 and 50% each year (Burke and Dunne, 2008; Jørgensen *et al.*, 2014; Fones and Gurr, 2015). *Z. tritici* spores germinate on the leaf and grow within the leaf without damaging it (Kema *et al.*, 1996b). After the latent period, the leaf tissue is killed and pycnidia are formed within necrotic lesions which decrease the photosynthetic area of the leaf (Rudd *et al.*, 2015). Spores can disseminate long distances and they can remain dormant from the previous crop until optimal growing conditions occur (Kema *et al.*, 1996a).

Common STB control strategies include the application of fungicides. However, intensive fungicide use has in many cases resulted in selection for resistant strains leading to less effective, or in extreme cases totally ineffective fungicides (Dooley *et al.*, 2016; Jørgensen *et al.*, 2017). Mutations in the target protein of the fungicides lead to target site resistance, insertions in the promoter region (TSR), and the upregulation of the efflux pump lead to non-target site resistance (NTSR) (Griffin MJ, 1985; Cools and Fraaije, 2013). For example, in the case of methyl benzimidazole carbamate (MBC) fungicides, which interfere with cellular division by binding to the β-tubulin and inhibiting its polymerization, and quinone outside inhibitors (QoI), which target the complex III of the cytochrome b, which is part of the respiration process, single nucleotide polymorphisms in the target protein have led to *Z. tritici* fungicide resistant populations (Griffin MJ, 1985; Fraaije, 2003). Currently, neither of these fungicide groups are recommended for STB control.

The three fungicidal modes of action (MoA) that remain available for *Z. tritici* control are: multi-site inhibitors, succinate dehydrogenase inhibitors (SDHIs) and demethylation inhibitors (DMIs) commonly known as azoles. Azoles interfere with the biosynthesis of ergosterol by interacting with the haem iron in 14α-demethylase (CYP51), which is a member of the Cytochrome P450 superfamily. Despite azoles still

being considered suitable for control, many CYP51 mutations have been associated with changes in fungicide sensitivity. In *Z. tritici*, different haplotypes, can have different sensitivities to the various azoles (Leroux et al., 2007, Stammler et al., 2008, Cools and Fraaije, 2013), with a reduced sensitivity having been linked to inserts in the promoter, membrane-bound transporters and, more importantly, intragenic recombination of *CYP51* (Leroux and Walker, 2011; Cools and Fraaije, 2013; Kildea *et al.*, 2019).

The limited choice of fungicides available for use on cereals, caused by the lack of development of new active ingredients and new regulations which have limited the number of available fungicides, places severe pressures on winter wheat production. Further development of resistance by *Z. tritici* to the remaining fungicides available for STB control would exacerbate the situation (Jonathan Blake, 2011; Jess *et al.*, 2014). In addition, on the relatively rare occasions that new active ingredients become available (e.g. BASF's Revysol and Corteva's Inatreq, both available since 2020) it is essential to have appropriate resistance management practices in place to maximise the product life-time. Better understanding of the genetic composition of the pathogen populations circulating in cropping systems could help manage pathogens and resistance build up in several ways. If the data can be produced fast enough, it could help to tailor fungicide spray programmes in a responsive manner. Testing the circulating populations at the start of the season and/or between spray applications could establish how effective the previous treatment was and enable the selection of products which are effective for controlling the isolates remaining in the population. The data could contribute to the modelling of population dynamics, giving a better understanding of how different MOAs affect the population structure which could help develop more effective spray programmes. The methods could potentially also be used as a surveillance tool, deployed to detect the presence of new resistant isolates early, when their prevalence is low. Linked to appropriate management practices this could prevent the build-up of these isolates within populations prolonging the time over which the fungicide can provide sufficient disease contro (van den Berg *et al.*, 2013).

Traditionally, screening of resistant populations with mutations in *CYP51* is carried out by using Sanger sequencing (Leroux *et al.*, 2007; Estep *et al.*, 2015; McDonald *et al.*, 2019), Kompetitive allele specific PCR (KASP) (Vagndorf *et al.*, 2018),

pyrosequencing (Wieczorek *et al.*, 2015), quantitative PCR (qPCR) (Taher *et al.*, 2014; Hellin *et al.*, 2020), high resolution melting (HRM) (Curvers *et al.*, 2015), loop-mediated isothermal amplification (LAMP) (Fraaije, 2014a) or Illumina sequencing (McDonald *et al.*, 2015; Pieczul and Wąsowska, 2017). These methods are, time-consuming and they cannot be applied to a mixed field population and therefore rely on isolation and culturing of a large number of isolates in the laboratory, prior to testing. Although very accurate and able to analyse full length sequences of *CYP51*, Sanger sequencing can only sequence a single DNA fragment (below 900 nt) from a single isolate. Whilst current high throughput sequencing methods (HTS) can generate millions of DNA fragments simultaneously, most of them generate sequences of approximately 300 nt that are far shorter than the full length of the target genes (e.g. *CYP51* is around 1632 nt) and therefore can not be used to analyse the haplotypes present in mixed samples or populations. Given that several mutations can occur within *CYP51*, each potentially associated with a different resistance phenotype, a method designed to obtain an accurate representation of the resistance composition of the pathogen population would need to be able to sequence the full *CYP51* gene. The method should also be able to quantify the relative presence of each haplotype within a population.

Over the last few years, two new sequencing methods which allow the analysis of mixtures of long fragments (Laver *et al.*, 2015; Rothfels *et al.*, 2017) and have the potential to sequence the full length of *CYP51* including its promoter region have been developed: the single-molecule real-time (SMRT) sequencing developed by Pacific Bioscience (PacBio) and nanopore sequencing developed by Oxford Nanopore Technologies (ONT). The ONT nanopore sequencing has several potential advantages over the SMRT sequencer. Firstly, the initial capital purchase price is 100x lower for a MinION. Secondly, the platform is scalable from 7 Tb of data on a 48 cell ProMethION through MinION flow-cells that can produce 30 Gb of data, down to an adapter for smaller capacity flow-cells (Flongle) which can produce up to 2 Gb of data for 10x less cost, which is potentially useful for routine analysis labs. Thirdly, the MinION is the only portable HTS device which has the potential to be used in-field (Radhakrishnan *et al.*, 2019) or in remote or resource poor laboratories. Diagnosis of plant diseases in the field using portable devices is highly recommended as the time between sampling, analysis and recommended action can be significantly reduced

(Fraaije 2014). Whilst LAMP tests can be used in the field to identify single nucleotide polymorphisms (SNP's) in the target genes, they require isolation of the pathogen and large numbers of individual tests to achieve a picture at the population level. Platforms like the MinION could be used directly to determine all haplotypes present within populations circulating in the field, enabling a full assessment of risk and facilitating quicker, potentially real-time decision-making regarding fungicide selection.

The aim of this study was to determine whether MinION sequencing can be used to analyse the presence and variability of *CYP51* mutations in *Z. tritici* populations to obtain an accurate reflection of the haplotype composition of the population. Synthetic populations were produced by mixing DNA from isolates in different mock communities and the proportions of isolates as predicted by the sequencing results were subsequently compared to the expected (original) proportions. DNA was also directly extracted from leaves to analyse the frequency of the haplotypes within the "in-crop" population.

2.2 Materials and methods

2.2.1 Isolation and culturing

Z. tritici infected KWS Santiago wheat leaves were collected from a spray trial in Bishop Burton (Yorkshire, UK) in 2019 and from Salisbury and Ludlow (UK) during a previous project undertaken between 2015 and 2017 (Innovate UK project 102088, https://gtr.ukri.org/projects?ref=102088). Three samples obtained from Shropshire and Dorset were used in a preliminary assay in order to test the reliability of nanopore sequencing. Spray trials developed in Bishop Burton were designed to explore the genetic changes in *Z. tritici* populations throughout the season. Eight different fungicides, SDHIs or azoles, were sprayed at two growth stages, GS31-32 and GS37-39 (first node at least 1 cm above tillering node and flag leaf blade all visible, respectively), and a control plot. Additionally each plot was sprayed with Corbel (BASF, Ireland), a systemic fungicide with protectant and curative properties for the control powdery of mildew and rust. Samples were collected at the beginning of the season, just before the first application (GS31-32) and after the last spray (GS71-72, grain watery ripe).

Isolations were performed following a method similar to that of Kildea (2009). Leaves were sequentially rehydrated in sterilized distilled water, sterilized in 10% NaClO, rinsed in water and dried out on sterilized Grade 1 filter paper (Whatman, GE Healthcare, Fischer Scientific, UK). Leaves were placed on water agar exposing pycnidia upwards and incubated for 48 hours at 18-20°C whilst exposed to UV light. Single cirri were picked using a sterilized needle and transferred to Potato Dextrose Agar (PDA) with 1% streptomycin, after which the cultures were incubated for 3-5 days under the previously described conditions. Single colonies were transferred to a new PDA plate and incubated for 3-5 days at 18-20°C in the presence of UV light. Pure cultures were obtained by scraping the isolates from the PDA plate and storing in 10% sterilised semi-skimmed milk at -20°C until required.

2.2.2 Mock populations

Isolates IS132, IS162 and IS172 were chosen from the previous Flonge run by their haplotype. The experiment wanted to test if we could differentiate 1) isolates with different haplotypes but with a single mutation which differ between those hapotypes and 2) isolates with very different haplotypes, which differ in many mutations. DNA of the three isolates with different haplotypes, were combined in two isolate mixtures IS132:IS162, IS132:IS172 and IS162:IS172 of different ratios (95:5, 90:10, 85:15, 75:25, 60:40, 50:50, 40:60, 25:75, 15:85, 10:90 and 5:95) to generate 33 mock populations. Additionally, 6 more mock populations using mixtures of the three isolates IS172:IS162:IS132 were generated, 60:20:20, 80:10:10, 20:60:20, 10:80:10, 20:20:60 and 10:10:80. These mock populations were sequenced using a MinION R9.4.1 flow cell (Oxford Nanopore Technologies) for 48 hours and in a R10.3 flow cell (Oxford Nanopore Technologies) for 72 hours as described below.

2.2.3 Field populations

A total of 16 infected wheat samples from the field trial in Bishop Burton were sequenced in the same experiment as the mock communities. These samples were collected from untreated plants and from plants treated with azoles. Leaves were dried overnight and storage at ambient room temperature (18-22°C) till needed. DNA was extracted as explained below from full leaves, sequenced on a MinION R9.4.1 flow cell for 48 hours and repeated using a R10.3 flow cell for 72 hours.

2.2.4 DNA extraction

DNA was extracted from pure cultures using a QIAmp DNA Blood Mini Kit (Qiagen, Manchester, UK) according to the manufacturer's protocol but replacing PBS with ATL buffer (Qiagen) in the lysis step which was incubated at 56 °C for 3-12 hours instead of 10 min. The extraction protocol from the pure culture to DNA it takes between 5 and 14 hours, depending how long the samples are incubated in ATL buffer.

To extract DNA from full leaves, leaves were surface sterilized in 10% NaClO for 20 seconds, rinsed thoroughly in sterile water and air-dried for 15 minutes. Single leaves were cut into 1 cm pieces using a sterile scalpel and placed in a 5 ml tube (Axygen, Fisher Scientific, UK) with a 10 mm stainless steel ball bearing (Qiagen) and 3 ml of CTAB buffer (2% CTAB, 1.5 M NaCl and 1 M NaH_2PO_4, 1% antifoam in distilled water, Sigma Aldrich, Dorset, UK) before being homogenised using the Geno Grinder 2010 (SPEX Sample Prep, Rickmansworth, UK) at 1,000 rpm for 4 minutes. Supernatant was transferred into a 2 ml tube and centrifuged for 2 minutes at 5,000 rpm. Following centrifugation, 800 µl of the supernatant was transferred to a 2 ml tube containing 250 µl guanidinium isothiocyanate buffer (GITC) (Guanidine Thiocyanate 5.25 M, Tris HCl 50 mM, EDTA 20mM and 1.3% Triton X-100 in distilled water), 750 µl of precipitation solution (Wizard Magnetic DNA Purification System for Food, Promega, UK) was added, vortexed for 20 seconds and the samples centrifuged for 10 minutes at 13,000 rpm. The supernatants (500 µl) were transferred to a 96-well plate for DNA binding along with 400 µl of 60% isopropanol, 50 µl of magnesil beads from the Wizard Magnetic DNA Purification System for Food kit and the centrifuged sample The samples were processed using a Kingfisher Flex automated, magnetic particle processor (Thermo Fisher Scientific, Loughborough, UK) transferring the magnetic beads through and aliquot of 900 µl of GITC buffer, two aliquots of 900 µl 70% ethanol before final elution of DNA into 200 µl of molecular grade water (Severn Biotech Ltd., UK). The extraction from leaves takes about 7 hours, depending on the number of samples that need extractiung. The resulting DNA was quantified using the dsDNA HS Assay Kit and the Qubit 2.0 Fluorometer (Life Technologies) and stored at -20°C until required.

2.2.5 Sanger sequencing

In order to confirm the size of the insertion in the promoter region of the *CYP51* gene of the isolate IS360 from the Bishop Burton trial, a PCR was performed using the primers Mg51-ProF 5' GTGGCGAGGGCTTGACTAC 3' and Mg51-SeqR 5' CTGCGCGAGGACTTCCTGGA 3', designed by Cools *et al.* (2012). The reaction was carried out in a Biorad C1000 thermocycler (Biorad, Watford, UK) using 1 µl of DNA, 12.5 µl of LongAmp Taq 2X Master Mix (New England Biolabs Ltd, Hitchin, UK), 1 µl of each primer at 10 µM and 9.5 µl of molecular grade water. The PCR reaction was performed as follows: 30 seconds of initial denaturation at 94°C followed by 30 cycles of denaturation at 94°C (30 seconds), annealing at 59°C (30 seconds), extension at 65°C (50 seconds), followed by a final extension of 65°C (10 minutes).

DNA extracted from isolates collected in the field trials at Salisbury and Ludlow, were amplified using the primers ST51F1 5' ATGGGTCTCCTCCAGGAAGTCCTCC 3' designed by (Fraaije, 2014b) and ST51R3 5' TTGTGAAAGCAGCGTCTCCCTC 3' designed in the current study, which amplify the full length of *CYP51.* A PCR reaction using 1 µl of each primer at 10 µM, 10 µl of 5X Phusion HF buffer (Thermo Scientific, Oxford, UK), 1 µl of dNTPs at 10 µM, 2.5 µl of each primer at 10 µM, 1.5 µl of DMSO, 0.5 µl of Phusion Hot start Flex DNA polymerase, 31 µl of molecular grade water and 1 µl of the DNA template. PCR conditions were as follows: initial denaturation at 98°C (30 seconds) followed by 35 cycles of denaturation at 98°C (10 seconds), annealing at 68°C (20 seconds), and extension of 72°C (60 seconds), followed by a final extension of 72°C (10 minutes).

Amplification was confirmed using 1% agarose gel electrophoresis. PCR products were purified using QIAquick PCR Purification Kit (Qiagen) following the manufacturer instructions. PCR products were sequenced by Eurofins Genomics (Ebersberg, Germany) with the sequencing primers SF- 5' GCGCAGTTCGACGCGCAATT 3', F3- 5' GCGGACCTCACCACTACCTCGA 3' and SR2- 5' CGCGCTATTCATTAGCATAACATCCACC 3'. Sequences were first checked for quality, contigs made for each isolate and primers removed using Geneious Prime 2019 (version 2019.2.1) (https://www.geneious.com).

2.2.6 Nanopore sequencing

A preliminary analysis was developed in order to test the robustness of nanopore sequencing for the analysis the full length of *CYP51* including the upstream regulatory containing insertions for *Z. tritici* isolates. An isolate of each of the 3 most common genotypes in the Innovate UK project (isolates from Salisbury and Dorset) were selected to perform this analysis. A PCR was performed using primers specific to amplify the full length of the *CYP51* gene including the promoter region. These primers were tested against other fungal pathogens which affect wheat crops and none of them amplified when using this set of primers. Primers were tailed to allow later addition of Nanopore adapters and barcodes, Cy51-MI-F1- 5' TTTCTGTTGGTGCTGATATTGCCTTCGCTGAACACCTCGCTT 3' and ST51-MI-R3- 5' ACTTGCCTGTCGCTCTATCTTCTGAAAGCAGCGTCTCCCTC 3', 1 µl of DNA, 12.5 µl of LongAmp Taq 2X Master Mix, 1 µl of each primer at 10 µM and 9.5 µl of molecular grade water. The PCR conditions were 30 seconds of initial denaturation at 94°C followed by 30 cycles of denaturation at 94°C (30 seconds), annealing at 59°C (30 seconds), extension at 65°C (3 minutes 20 seconds), followed by a final extension of 65°C (10 minutes). Amplification was confirmed by 0.5% agarose gel electrophoresis. PCR products were purified using AMPure XP beads (Beckman Coulter, Wycombe, UK) employing a 0.5X beads to sample ratio. Purified PCR products were tagged using the PCR Barcoding expansion 1-96 (EXP-PBC096, Oxford Nanopore) according to the manufacturer's protocols. Barcoded PCR products were quantified using a Quant-iT Picogreen dsDNA Assay Kit (Thermo Fisher Scientific) in the Fluoroskan Ascent (Thermo Scientific). To estimate size of the amplicons in the library TapeStation D5000 high sensitivity assay (Agilent Technologies, Cheadle, UK) was used. Samples were pooled in equal amounts of DNA to a final quantity of 1-1.5 µg. The libraries were prepared using the SQK-LSK108 sequencing kit (ONT), following the manufacturer's protocols and sequenced on a R9.4 MinION flow cell for 48 hours.

Similarly, libraries using isolates, mock communities and wheat leaves from the Bishop Burton spray trial were prepared as described above but replacing the sequencing kit with the SQK-LSK109 following the manufacturer's instructions and sequenced on MinION R9.4, R10.3 or Flongle flow cells.

2.2.7 Sequence Analysis

MinION raw reads in fast5 format from the preliminary experiment (samples obtained from Salisbury and Dorset) were initially basecalled using Guppy v3.1.5, then re-analysed using a newer version (Guppy v3.4.5) to compare the quality score with further sequencing runs. Flongle and MinION raw reads from the Bishop Burton experiment were base called using the Guppy v3.4.5 GPU basecaller with the appropriate configuration file and demultiplexed, to determine which sequences came from which samples, using the Guppy barcoder (Table 2.1). Nanoplot (version 1.28.1) was used to check the key features and statistics of the sequencing run (De Coster *et al.*, 2018). The most common parameter used to measure the sequence quality of the nucleotide is the Phred score, also known as Q score. The Phred score indicates the probability (p) of an incorrect base call: $Q = -10\log_{10}(p)$. Thus, the higher the value of the Phred score, the higher the quality of the sequence. A Phred score of 10 represents a 10% base call error rate (0.1), i.e. 90% accuracy, while a Phred score of 20 relates to a 1% base call error rate.

Table 2.1 Schematic representation of the bioinformatic pipelines used

Device	**Flow-cell**	**Guppy version**	**Configuration file**	**Quality improvement tool**
Flongle	9.4	3.4.5	9.4.1	Filtlong + Canu
MinION	9.4	3.4.5	9.4.1	Filtlong
	10.3	3.4.5	10.3	Filtlong

Both the MinION and Flongle reads were filtered using Filtlong, for the highest quality 10 million bases, reducing the total number of reads analysed (https://github.com/rrwick/Filtlong). In addition a further correction was applied to the Flongle reads using Canu (version 1.8), which corrects the reads by building a read and overlap database, selecting the most consistent overlaps and subsequently generating corrected reads for which read lengths are estimated (https://github.com/marbl/canu) (Koren *et al.*, 2017). All filtered (and corrected) reads as well as Sanger sequences were mapped to the *Z. tritici* IPO323 reference genome (RefSeq assembly accession: GCF_000219625.1) using Minimap2 (version 2.17) (Li,

2018). A modified version of the IPO323 reference genome was manually created by adding the sequences of the 120 and 800 bp upstream insertions of the *CYP51* gene identified in some strains of this study. This modified version was used as the mapping file for the reads obtained from full leaves. In addition to mapping all reads to the genome, reads were grouped according to whether the read represents the forward or reverse complement strand of the reference sequence. These groups will be referred to in the remainder of the paper as 'forward' and 'reverse', respectively. Alignments were visualised using the IGV software (https://software.broadinstitute.org/software/igv/) (Robinson *et al.*, 2011; Thorvaldsdóttir *et al.*, 2013; Robinson *et al.*, 2017). Consensus sequences from the isolates were obtained with IGV and confirmed using MegaX (Kumar *et al.*, 2018).

To analyse the isolate with a very long PCR product (IS146), we mapped the reads to the original, unmodified IPO323 genome sequence using the LAST software (Frith *et al.*, 2010) using the seeding scheme tuned for relatively high rates of insertion and deletion compared to substitution. Mapped reads with a probability of a miss-mapping $> 10^{-6}$ were discarded. The alignments were then analysed with an in-house script which grouped near-identical reference-sequence segments (to which reads had been mapped). In this method, reference segments were grouped such that each segment has a 5' end ≤ 10 bp from the 5' end of at least one other segment in the same group. Within each group, segments were then sub-grouped such that each segment's 3' end is ≤ 10 bp from the 3' end of at least one other segment in the same group. Each subgroup thus represents a consensus segment of the reference segment, to which at least one (often very many) actual mapped segment belongs. For each consensus segment, the median 5' and 3' coordinates of all of the mapped IS146 reads were then determined. This results in effectively a consensus alignment of each segment of the IS146 amplified region and the corresponding segment of the reference genome.

To analyse mixed samples and populations, positions in the gene where the two isolates differed were annotated and the relative proportion of each nucleotide per position was collected using IGV. Moreover, to determine the haplotypes present in the mixture, how those nucleotides correlate at SNP positions, a script using Pysam (https://pysam.readthedocs.io/en/latest/#) (Li *et al.*, 2009) was used to extract the nucleotides of interest and generate a sequence with the SNPs combinations.

A generalised linear model (GLM) was employed to determine which predictor variables affected the difference in the observed compared to predicted proportions of eachpolymorphism. A GLM, with quasibinomial errors was used using the “glm” function of the vegan package, since the response were proportion (percentage-derived) variables. Predictors were the input proportions (i.e. predicted or expected percentages and observed or read percentages) of each isolate, the isolate type (IS172 and IS162), the position of the polymorphism, and method (R9.4 and R10.3).

2.3 Results

2.3.1 Validation of nanopore sequencing for the analysis of CYP51 and the regulatory region in Z. tritici isolates

The full length sequence of *CYP51*, including the upstream region of three isolates (SEPT1278, SEPT1316 and SEPT1972) was determined using ONT and Sanger sequencing. The following SNPs were identified using both sequencing methods. SEPT1278, 5 SNPs: L50S, D134G, V136A, I381V, Y461H; SEPT1316, 8 SNPs: L50S, V136A, S188N, A379G, I381V, Y459del, G460del, S524T and SEPT1972, 8 SNPs: L50S, S188N, I381V, Y459del, G460del, N513K.

The accuracy per SNP position was calculated for the different pipelines. When raw reads were basecalled using Guppy v3.1.5 the per SNP accuracy is between 84-94% (Figure 1c), for filtered reads with Filtlong this increases to 92-97% (Figure 1d) and for corrected reads further increases to 99-100% (Figure 1e). The per position accuracy was improved when Guppy v3.4.5 was applied with the exception of the SNP present in the position 1,448,554 of the IPO323 reference genome, basecalled raw reads per SNP accuracy was between 82-94%, for filtered reads this increased to 92-99% and for corrected reads increases to 99-100%. Moreover, the median Phred score for the raw reads base called using Guppy v3.1.5 was 8.8, increasing to 13.0 when the reads were filtered. For Guppy v3.4.5, the median Phred score was 10.8 for raw reads and 13.0 with filtered reads.

2.3.2 Nanopore sequencing of isolates

A single Flongle 9.4 flow cell with 41 pores was run for 24 hours to sequence the *CYP51* region including the upstream region of a library containing DNA from 22

isolates from Bishop Burton. A total of 203,356 reads were generated with a median Phred score of 9.2 (88% accuracy). After filtering with Filtlong the number of reads were reduced to 58,052 which increased the median Phred score to 11.1 (92.2% accuracy). Once reads were corrected with Canu, the per position consensus was between 98 and 99%. Before applying Canu, forward, reverse and consensus reads were aligned to the reference genome in order to detect recurrent errors, which includes errors in homopolymer regions. As an example, the position 1,448,812 of the wild type (L50) is not in an homolopymer region (GACAAC), conversely when a SNP occur and the adenine (A) is mutated to a guanine (G) will convert the region to an homopolymer region (GACGAC). As a result, isolates IS132, IS162 and IS172 88-89% of the consensus reads (which includes forward and reverse reads) have a cytosine (C), when reads are divided by the strand, 96-97% of the reverse reads have a C while only 83-84% of the forward reads have a C (Table 2.2).

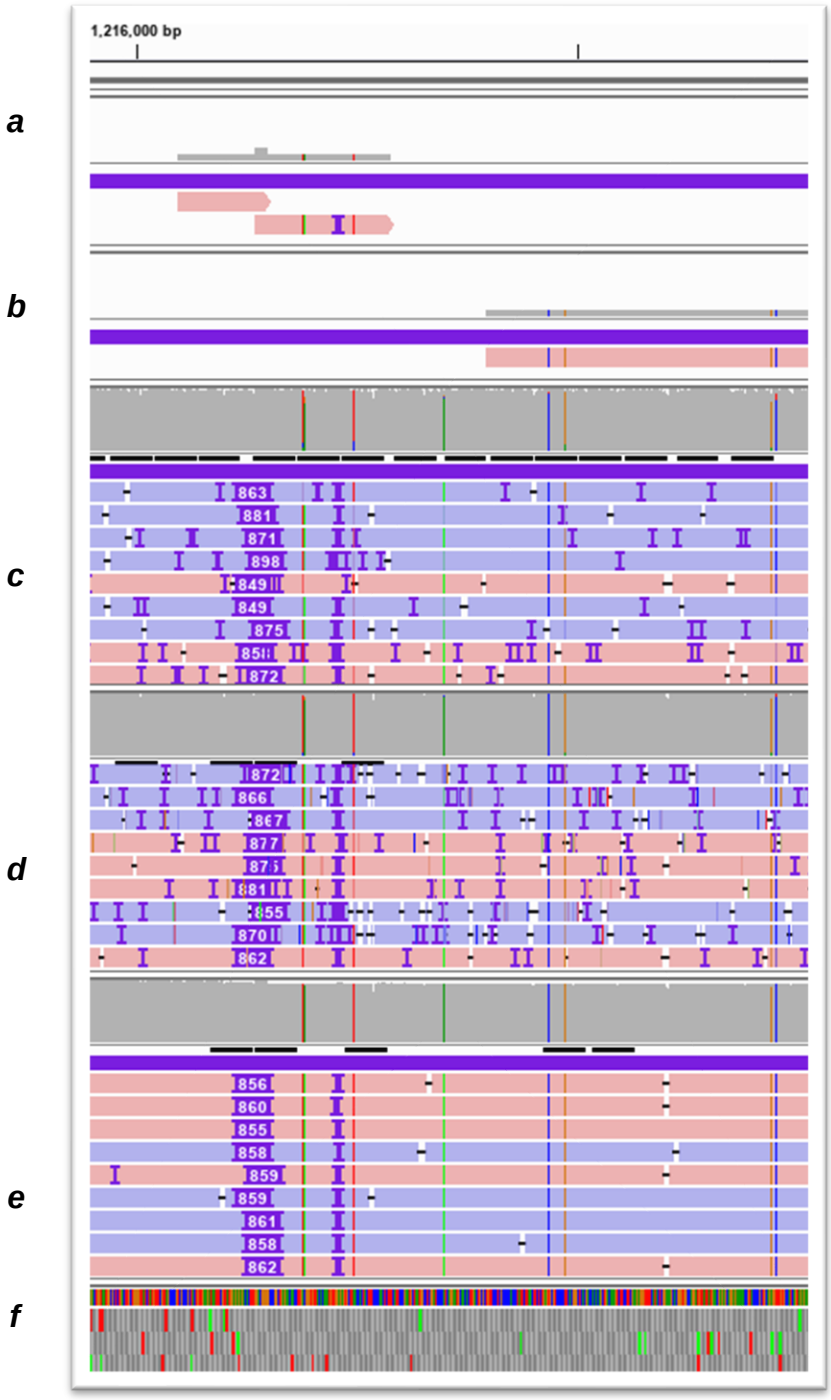
Table 2.2 Percentage of accuracy per SNP position of isolates IS132, IS162 and IS172.

Numbers on the top row indicates the amino acid position, first column indicates the isolate and the second column reads mapped to the IPO323 reference, consensus reads which includes forward and reverse reads and divided by strand (forward and reverse).

Isolate	Reads	50	134	136		188	379	381	460	513	524
IS132	Consensus	89	98	98	96	92	95	96	99	96	92
	Forward	84	99	99	97	93	95	95	99	97	95
	Reverse	96	98	97	94	99	96	97	99	96	89
IS162	Consensus	88	98	98	96	99	99	97	98	98	94
	Forward	83	99	99	98	99	99	97	97	97	97
	Reverse	97	97	98	93	98	99	96	99	99	89
IS172	Consensus	88	98	99	96	99	99	97	97	98	97
	Forward	83	99	99	98	99	99	97	95	97	97
	Reverse	96	98	99	94	99	98	96	99	99	98

Figure 2.1. SEP1276 reads for CYP51 mapped to the IPO323 reference using minimap2 and visualised in IGV.

a) Sanger sequence of CYP51 gene, b) Sanger sequence of the upstream insertion, c) MinION reads, d) MinION reads filtered with Filtlong, e) MinION reads corrected with Canu and f) IPO323 reference genome. Reads were colored by strand, red and blue reads were aligned to the forward and reverse strand, respectively. Within each read, dashes indicates deletions, while purple blocks indicates insertions. The grey coloured coverage track, which is above the coloured reads, represents the nucleotide coverage of each position. Coloured vertical lines indicate possible SNP positions, in green adenosine, blue cytosine, brown guanine and red thymine.



Consensus sequences from the 22 isolates were translated and amino acid substitutions were identified by comparison to the reference sequence of IPO323 (Table 1). No wild type sequences were found, and each isolate showed at least 4 amino acid changes. Amongst the 22 isolates, a total of 11 polymorphic sites were identified (L50, D134, V136, S188, A379, I381, Y459, G460, Y461, N513 and S524), with the polymorphism combinations representing 12 different haplotypes altogether. Isolates with the same polymorphism combinations were considered part of the same haplotype. The most common CYP51 haplotype contained the following amino acid substitutions V136C, I381V, Y461H and S524T, which were identified in isolates IS19, IS46, IS146, IS167, IS370, IS434 and IS487. The other polymorphism combinations were present in either one or maximum of two isolates. The haplotype with the most mutations was represented by isolates IS132 and IS472 and was characterized by 10 mutations (L50S, D134G, V136C, S188N, A379B, I381V, Y459del, G460del, N513K, S524T) when compared to the reference.

From the 22 samples, 18 had an 800 bp insertion in the upstream regulatory region and 2 samples had a 120 bp insertion. In addition to the most common insertions (120 bp and 800 bp), a 171 bp insertion was detected by ONT and confirmed by Sanger sequencing in isolate IS360. Moreover, IS146 showed genome rearrangements in the upstream region of *CYP51.*

Due to the unexpected pattern in the upstream region of the IS146 when amplicon sequences were visualized in IGV, we analysed the mapping of these sequences to the reference genome as described, ignoring all reference genome segments that were mapped by fewer than 100 read segments. The results indicated that this isolate exhibits genome rearrangements which result in the amplified sequence being almost 6 kbp long (the highest median 3' coordinate value of a consensus read segment is 5,925). This segment was located in the upstream region were we would normally find the 200 bp or 800 bp insertion. This segment contained substantial segments of approximately 200 bp to 900 bp length corresponding to various segments of chromosomes 2, 10 and 13 flanked by large segments at the 5' (2,200 bp) and 3' (670 bp) ends corresponding to chromosome 7 as expected. A 400 bp segment prior to the 3' chromosome 7 segment is unaccounted for by these mappings. Some of the read segments are essentially sub-segments of others, while some read segments which

map to different chromosomes overlap; e.g. a 200 bp overlap between a 370 bp segment mapping to chromosome 10 and a 500 bp segment mapping to chromosome 13. This is consistent with the chromosome sequences of the reference genome itself being the result of ancestral duplications and rearrangements.

Table 2.3 CYP51 haplotypes identified using nanopore sequencing on a Flongle flow cell.

The first row indicates the amino-acid position and the second row the reference insertion and amino acid (IPO323). No letter means no change, a dash reflects the deletion of the amino acid and a capital letter is the amino acid observed in that position. * means genome rearrangements.

Position	**Insertion**	**50**	**134**	**136**	**188**	**379**	**381**	**459**	**460**	**461**	**513**	**524**
IPO323	0	L	D	V	S	A	I	Y	G	Y	N	S
IS19	800			C			V			H		T
IS46	800			C			V			H		T
IS146	*			C			V			H		T
IS167	800			C			V			H		T
IS370	800			C			V			H		T
IS434	800			C			V			H		T
IS487	800			C			V			H		T
IS132	800	S	G	A	N	G	V	-	-		K	T
IS472	800	S	G	A	N	G	V	-	-		K	T
IS138	800	S	G	A			V			H		
IS172	800	S	G	A			V			H		
IS162	800	S	G	A			V			H		T
IS476	800	S	G	A			V			H		T
IS111	800	S	G	A		G	V	-	-			T
IS300	800	S	G	A		G	V	-	-		K	T
IS360	171			C	N	G	V	-	-			T
IS496	800	S		C	N	G	V	-	-			T
IS495	800			C		G	V	-	-			T
IS394	800			C		G	V			H		T
IS141	120	S			N		V	-	-		K	
IS497	120	S			N		V	-	-		K	
IS444	800	S		A	N	G	V	-	-			T

2.3.3 Nanopore sequencing of CYP51 from mock communities

Different mock communities were generated by mixing DNA from isolates IS132, IS162 and IS172 in different proportions. Isolates IS162 and IS172, differed with regards to a single amino acid (S524T) and IS132 differed in six mutations when compared to IS162 (S188N, A379G, Y459del, G460del, N513K and S524T) and five mutations when compared to IS172 (S188N, A379G, Y459del, G460del and N513K) (Table 2.3). A total of 2,827,220 reads were obtained using a single R9.4 MinION flow cell, with a median Phred score of 10.4 (90.88% accuracy). After filtering, the number of reads decreased to 142,903 and the median Phred score increased to 13.8 (95.8% accuracy). The same experiment was repeated by sequencing the same library in an R10.3 flow cell. A total of 4,854,782,586 reads were obtained with a median quality of 11.1 (92.2% of accuracy). After filtering 133,151 reads were obtained with a median Phred score of 14.7 (96.6% of accuracy) (Table 2.4).

Table 2.4 The read number and quality scores for samples run using R9.4 and R10.3 flow cells.

	R9.4		R10.3	
	Before filtering	After filtering	Before filtering	After filtering
Number of reads	2,827,220	142,903	4,854,782,586	133,151
Median Phred score	10.4	13.8	11.1	14.7
% of accuracy	90.88	95.8	92.2	96.6

2.3.3.1 Comparison of observed versus expected frequencies for individual nucleotides

Reads obtained from the mock populations of isolates IS172 and IS162 sequenced in MinION R9.4 and R10.3 flow cells were visualized using the IGV software. Due to an experimental error in the preparation of the mock community formed by 90% of IS172 and 10% of IS162 the mixture was discarded. All multi-level interactions of the GLM were fitted, and the model was simplified to the minimal number of significant predictors (isolate type, position of the polymorphism and flow cell). Model assumptions were checked with comparison with quantile-quantile (QQ plots) of residuals. The method (R9.4 or R10.3) and the position of the polymorphism did not have an effect in the model. Conversely, the isolate had an effect in the model (p –

value = 0.00159), with isolate IS172 being observed in a higher proportion than IS162 (Figure 2.2).

The two isolates differed in six nucleotides, located in positions 1,446,998, 1,447,003, 1,447,044, 1,447,063, 1,447,093 and 1,447,118 (IPO323 reference genome nucleotide coordinates). The percentage of reads containing the observed versus expected nucleotide were visualised per position in Table 2.5. For R9.4 expected and observed percentages differed between 0% to 9% (Table 2.5a) while for R10.3 the error rate was reduced to between 0% to 7% (Table 2.5b). Due to the better performance of the R10.3 flow cell, the dataset from this flow cell was used for the further analyses.

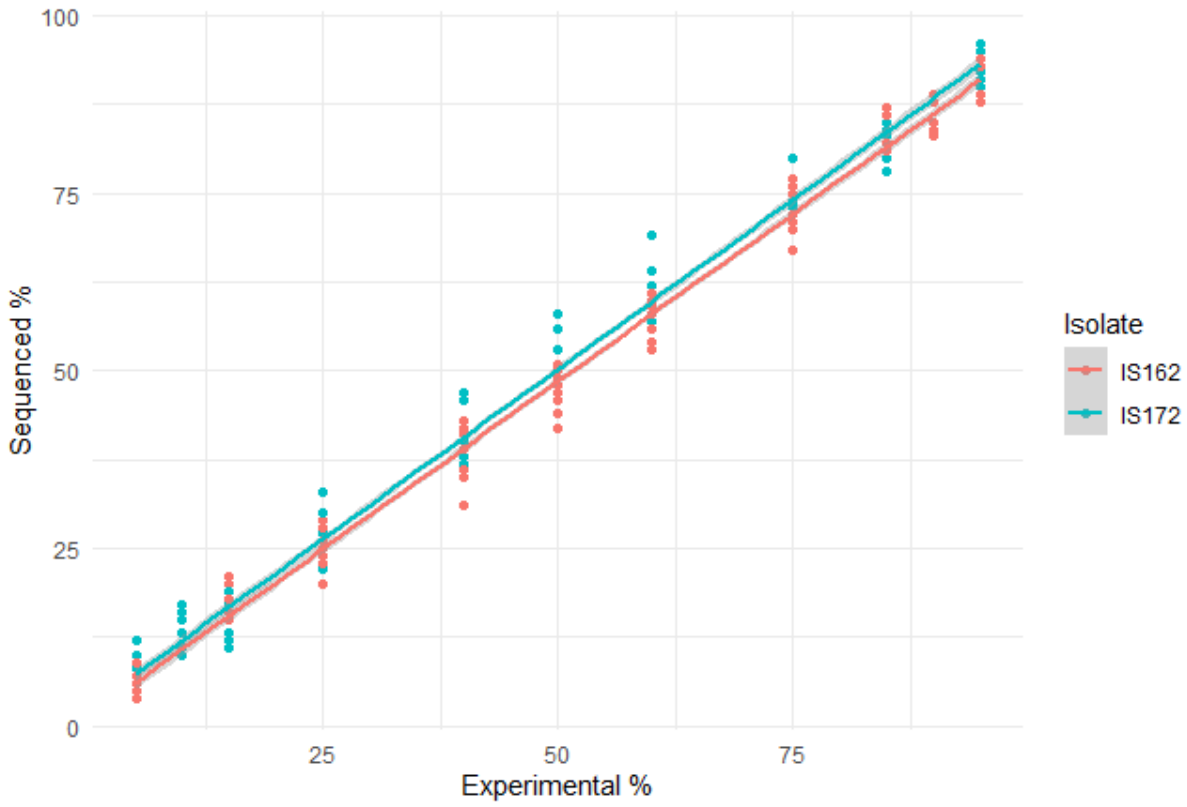


Figure 2.2 Scatter plot and fitted line representing the sequenced percentage of each isolate or observed percentage and the experimental or expected percentage.

Table 2.5 Variability between the expected and observed percentages of each polymorphism in IS172 and IS162 mixtures.

The top row indicates positions in the genome. In the second and last row the nucleotide which differs between the two isolates is marked in bold. Colour codes and values represents the variability between the expected and observed proportion, from green when the values are close, to red when the values differ "-" if IS172 is lower than expected and "+" is IS172 is higher than expected. a) R9.4 and b) R10.3.

a)		1446998	1447003	1447044	1447063	1447093	1447118
% IS172	% IS162	GCTATTC	TCACTAG	CAAGTCT	CTTCTCC	CTTTACT	CGGCTGA
100	0	A	C	G	C	T	C
95	5	1 -	5 -	0	1 +	0	3 -
85	15	1 -	7 -	0	4 +	1 -	3 -
75	25	2 -	5 -	1 -	5 +	1 -	2 -
60	40	1 +	1 -	2 +	9 +	4 +	2 +
50	50	0	3 -	0	8 +	3 +	1 +
40	60	1 -	3 -	0	7 +	2 +	0
25	75	2 -	3 -	2 -	8 +	3 +	0
15	85	3 -	4 -	2 -	4 +	2 +	0
10	90	1 -	0	1 +	7 +	6 +	3 +
5	95	1 +	0	1 +	7 +	5 +	3 +
0	100	G	T	T	T	C	G
		GCTGTTC	TCATTAG	CAATTCT	CTTTTCC	CTTCACT	CGGGTGA

b)		1,446,998	1,447,003	1,447,044	1,447,063	1,447,093	1,447,118
% IS172	% IS162	GCTATTC	TCACTAG	CAAGTCT	CTTCTCC	CTTTACT	CGGCTGA
100	0	A	C	G	C	T	C
95	5	1 -	5 -	1 -	1 +	2 +	0
85	15	4 -	7 -	3 -	0	5 +	2 +
75	25	1 -	4 -	1 -	2 +	2 +	1 +
60	40	3 -	4 -	2 -	4 +	2 +	1 +
50	50	0	2 -	0	6 +	1 +	3 +
40	60	1 -	2 -	0	6 +	1 +	3 +
25	75	1 -	2 -	2 -	5 +	2 +	2 +
15	85	2 -	2 -	2 -	4 +	2 +	1 +
10	90	1 -	1 -	0	6 +	5 +	3 +
5	95	1 +	0	1 +	6 +	7 +	4 +
0	100	G	T	T	T	C	G
		GCTGTTC	TCATTAG	CAATTCT	CTTTTCC	CTTCACT	CGGGTGA

2.3.3.2 Comparison of observed versus expected frequencies for haplotypes

To quantify the haplotypes present in the IS162 and IS172 mixtures, the analysis was focused on the SNPs that differed between the two isolates, at positions 1,446,998, 1,447,003, 1,447,044, 1,447,063, 1,447,093 and 1,447,118 which were extracted and concatenated using Pysam (i.e. only the six nucleotides which differed between the two isolates are being considered; columns 3 and 4 in Table 2.5). Each mixture contained approximately 2500 reads and for each of these reads the nucleotide sequence for the six positions of interest was extracted and subsequently the observed frequency of each sequence calculated. For each mixture, a total of 13-23 sequences were obtained. When haplotypes were determined based on the six nucleotides, an error rate between 27.9% to 45.77% was observed, representing the reads that were not correctly assigned to either of the two isolates (Table 2.6). However, some of the observed mutations were silent when sequences were translated into its representative amino acid and the error rate was reduced to between 4.58 and 11.29%, resulting in a much higher recovery of the expected isolate frequencies (Table 2.6).

Table 2.6 Summary of the observed versus predicted percentages for each isolate based on the six nucleotides which differ between the two isolates (IS172/ACGCTC versus IS162/GTTTCG) and just analysing the SNP which translate into a different amino acid.

Predicted %		Observed %					
		Differing nucleotides			Amino acid associated		
IS172	IS162	ACGCTC (%)	GTTTCG (%)	% Error	% C	% G	% Error
95	5	70.00	2.10	27.90	91.42	4.00	4.58
85	15	61.01	7.35	31.64	80.77	14.74	4.49
75	25	54.57	11.43	34.01	71.40	22.66	5.95
60	40	45.49	19.02	35.49	60.17	33.75	6.08
50	50	37.26	23.34	39.40	48.89	44.59	6.51
40	60	28.80	29.42	41.78	38.84	55.54	5.62
25	75	16.80	34.77	48.42	24.35	68.57	7.08
15	85	8.77	41.79	49.44	14.71	77.11	8.19
10	90	7.50	42.29	50.21	12.62	79.50	7.88
5	95	3.65	45.77	50.58	7.88	83.50	8.62

2.3.4 Nanopore sequencing of more complex mixtures

The haplotypes of the mixtures IS162:IS132, IS172:IS132 and IS132:IS162:172 were extracted as explained in section 2.3.2. The SNP's were extracted, concatenated and frequencies of each haplotype was calculated after discarding haplotypes with a relative abundance was below 5%. The difference between the expected percentage or experimental percentage and the observed percentage or sequenced percentage varied from 0-7% (Table 2.7) in IS172:IS132 and IS162:IS132 mock communities. Moreover, in most of the 95-5% mixtures the haplotype in the lower concentration could not be determined as it was present below the 5% threshold. Similarly, the frequency of the haplotypes above 5% in IS132:IS162:IS172 mixtures were extracted (Table 2.8). The variability between observed and expected percentages in mock communities were 3 isolates where mixed was between 0.7 and 6.6%.

Table 2.7 Summary of the observed versus predicted percentages of each haplotype extracted from the IS172:IS132 and IS162:IS132 mock communities.

ND represent mixtures where haplotypes were not extracted due to an experimental error.

Predicted %		**Observed %**		**Predicted %**		**Observed %**	
IS172	IS132	IS172	IS132	IS132	IS162	IS162	IS132
95	5	100	0	95	5	100	0
90	10	ND	ND	90	10	89	11
85	15	89.5	10.5	85	15	82.9	17.1
75	25	77.9	22.1	75	25	76.3	23.7
60	40	67	33	60	40	58	42
50	50	49.7	50.3	50	50	55.9	44.1
40	60	39.2	60.8	40	60	35.2	64.8
15	85	19.4	80.5	15	85	18.6	81.4
10	90	15.4	84.6	10	90	10	90
5	95	5.2	94.8	5	95	0	100

Table 2.8 Summary of the observed versus predicted percentages of each haplotype extracted from the IS132:IS162:IS172 mock communities.

Predicted %			Observed %		
IS132	IS162	IS172	IS132	IS162	IS172
20	60	20	18.9	64.5	16.2
10	80	10	9.3	82.9	7.8
20	20	60	24.2	16.4	59.4
10	10	80	15.8	8.5	75.7
60	20	20	53.5	20.7	25.8
80	10	10	83.2	<5	16.6

2.3.5 Nanopore sequencing of CYP51 from infected leaves

DNA from 11 wheat leaves was extracted and the CYP51 region, including the promoter region of *CYP51* was sequenced. CYP51 haplotypes present above 5% in these leaves, including mutations in the target gene and the insertion in the promoter region were identified (Table 2.9). Each population had between 1 and 6 haplotypes above the 5% threshold. The relative proportion of each haplotype was calculated by dividing it by the sum of the haplotypes above 5%. For example, for POP1 a total of 30 haplotypes were identified, of which four haplotypes were present at > 5%. Of these four haplotypes 37.5% had the mutations L50S, V136C, S188N, Y459del, G460del, N513K and the 120 bp insertion in the upstream region, 25.1% had L50S, S188T, A379G, Y459del, G460del, S524T and the 800 bp insertion, 22.6% L50S, D134G, V136A, S524T and the 800bp insertion and 15% had S524T and 800 bp (Table 2.10). The four haplotypes represented 70% of the total number of sequences recovered.

Table 2.9 CYP51 haplotypes identified in different populations (POP1-POP11). The first row indicates the amino-acid position and the second row the reference insertion and amino acid (IPO323). No letter means no change, a dash reflects the deletion of the amino acid and a capital letter is the amino acid observed in that position.

	Frequency	Insertion	50	134	136	188	379	459	460	461	381	513	524
	(%)	(bp)	L	D	V	S	A	Y	G	Y	I	N	S
POP1	29.3	800	S	G	A					H	V		T
POP1	20.0	800			C					H	V		T
POP1	24.7	800	S		C	N	G	-	-		V		T
POP1	25.9	120	S			N		-	-		V	K	
POP2	27.5	800	S		C	N	G	-	-		V		T
POP2	29.4	800	S		A	N	G	-	-		V		T
POP2	27.2	120 + 800	S	G	A					H	V		T
POP2	16.0	120	S			N		-	-		V	K	
POP3	13.3	800	S	G	A		G	-	-		V	K	T
POP3	12.2	800	S	G	A					H	V		
POP3	26.7	800	S	G	A					H	V		T
POP3	22.5	800			C	N				H	V		T
POP3	25.4	120	S					-	-		V	K	
POP4	25.4	800	S	G	A					H	V		T
POP4	15.3	800			C					H	V		T
POP4	26.5	800	S		C	N	G	-	-		V		T
POP4	32.8	120	S			N		-	-		V	K	
POP5	10.4	120				N		-	-		V	K	
POP5	89.6	120	S			N		-	-		V	K	
POP6	74.5	800	S		C	N				H	V		T
POP6	25.5	120	S			N		-	-		V	K	
POP7	10.2	800		G	A					H	V		T
POP7	79.1	800	S	G	A					H	V		T
POP7	10.7	800	S		C	N	G	-	-		V		T
POP8	52.1	800	S	G	A					H	V		T
POP8	47.9	800			C					H	V		T
POP9	21.7	800	S	G	A					H	V		T
POP9	17.2	800			C					H	V		T
POP9	61.1	120	S			N		-	-		V	K	
POP10	13.6	120	S			N		-	-		V	K	
POP10	86.4	120	S		A					H	V		T
POP11	74.8	800			C					H	V		T
POP11	25.2	800	S		A	N	G	-	-		V	K	T

Table 2.10 Example of the sequences and calculation of the frequency of population POP1

Sequence	*Freq*	*Above 5%*	*Haplotype %*
-G--CTCATG--CC	17		
-GCCGCCATA--CC	11		
-GCCGCGCCG--CC	16		
C---GTACTGAG	19		
CC--CTACTGAG	28		
CC--GCACTGAG	20		
CC--GTACTGAG	515	515	37.45
CC--GTCATGAG	15		
CC--GTGCTGAG	14		
CCCCGTACTGAG	19		
CG--CTCATG--CC	12		
CGCCGCCATA--CC	14		
CGCCGCGCCG--CC	71		
CT--GTACTGAG	15		
GC--CCGCCG--CC	24		
GG--CCCATG--CC	17		
GG--CT--TG--CC	63		
GG--CTACTG--CC	25		
GG--CTCATG--CC	346	346	25.16
GG--CTGCTG--CC	21		
GGCCCCGCCG--CC	11		
GGCCCTCATG--CC	11		
GGCCGC--TA--CC	54		
GGCCGCACCG--CC	20		
GGCCGCACTA--CC	17		
GGCCGCCATA--CC	204	204	14.84
GGCCGCGC-G--CC	51		
GGCCGCGCCG--CC	310	310	22.55
TC--GTACTGAG	14		
TGCCGCGCCG--CC	11		
Total	*1985*	*1375*	*100*

2.4 Discussion

Nanopore sequencing was successfully used to analyse *CYP51* and the associated upstream region for isolates, mock populations and field populations of *Z. tritici* infected leaves. A total of twelve different haplotypes were identified formed by the combination of different mutations, supporting previous findings of the very diverse nature of *Z. tritici* resistant isolates (Cools and Fraaije, 2013; Dooley, 2015). Although other studies have shown that the same polymorphism can have insertions of a different size in the promoter region (Leroux and Walker, 2011; Cools and Fraaije, 2013; Kildea *et al.*, 2019), in our study the insertion in the promoter region did not contribute to an increase in the number of haplotypes, which is probably due to the relatively low number of isolates sequenced.

There is good agreement between observed and expected results from mock mixtures of isolates when nonsynonymous substitutions were being compared (1-7% error rate). However, the error rate increases significantly when haplotypes with multiple differing nucleotides were compared (23-45%). Mutations found within or adjacent too homopolymer regions (both single nucleotides or repeating dinucleotides) resulted in higher error rates which was consistent with previous findings of homopolymer regions being a source of errors (Zeng *et al.*, 2013; delaGorgendiere *et al.*, 2018). To minimize the effect of the homopolymer regions Oxford nanopore developed new pores with a longer barrel and two electric current reading points (R10.3), compared with the pores in the R9.4 flow cells which have a single reading point (Tytgat *et al.*, 2020). Our analysis has shown that the error rates are reduced by using this newer flow cell, from 4.2% error rate to 3.4% error rate, especially with homopolymers. In cases where the population is predominantly wild type, the goal would be to to detect the emergence of resistance so policy makers can alter regulations and limit the spread of resistance. There will need further investigation to detect if 5% is enough to take action. However, this situation is highly unlikely since the use of fungicides have been exploited for a long time, up to the point of not finding wild types in the fields. On the contrary, in cases where the resistance is widespread, we are likely to have haplotypes with different RF to different active ingredients. In this case, we would want to find active ingredients which would have the best efficacy in the predominant population taking into account cost-effectiveness and the potential and the potential for further resistance evolution.

In this situation, using nanopore sequencing to analyse the Septoria population is useful to track how the population is changing after using the sprays which tackle the most predominant resistant haplotype to see which of the remining haplotypes will be the next to be predominant.

One of the main challenges when analysing mock communities and populations by nanopore sequencing is dealing with the high error rate. In this experiment, the errors found were of two types, as previously reported, i.e. random and recurrent errors (Malmberg *et al.*, 2019; Wick *et al.*, 2019). Random errors were sporadic and do not follow a pattern, while recurrent errors occurred in the same positions (mainly in homopolymer regions). In a population, random errors would not affect the results from the single nucleotide polymorphism comparisons, i.e. in this experiment (after filtering the reads) each mixture had over 2,000 sequences and random errors would represent a very low proportion of the reads. On the other hand, recurrent errors could be confused with single nucleotide polymorphisms. In order to distinguish between recurrent errors and SNPs, we followed the methods outlined by Medaka (https://nanoporetech.github.io/medaka/future.html), which highlighted that recurrent errors have a strand bias, therefore these SNPs can be distinguished from errors by aligning the consensus sequence to the reference and splitting the reads by the sequencing orientation.

To analyse the haplotypes present in a population we need to know how the SNPs are related, which mutations are from the same isolate. SNPs were extracted from each read and reads with similar SNP profiles were identified and clustered together. SNP extraction was tested in the different mock communities. From each IS172:IS162 mock community, haplotype errors, when only looking a the nucleotides which will translate into a different amino acid, represented between 4.58% and 8.62% of reads, indicating that isolates present at a low frequency within populations would not be discriminated from sequencing errors. Similarly, SNPs were extracted from different populations, DNA extracted from full leaves, and the haplotypes present above 5% were detected and quantified.

We also have shown that even with a fungal genome of high plasticity (Goodwin, 2011), the Nanopore platform can be used to characterise novel inserts. In this case, complex genome rearrangements and/or duplications result in unexpectedly long

amplified sequences, in this case of around 2.5 x the most common insertion found in this experiment (800bp) and involving segments apparently originating from three other chromosomes with reference to the IPO323 genome.

It has been hypothesised previously that the use of in-field diagnostics for the identification of resistant isolates would be beneficial, enabling within season modification of a spray programmes (Radhakrishnan, 2019). For this to be effective the method needs to enable the identification of the most prevelant resistant haplotypes within populations and the time between sample collection and generation of results should be sufficiently rapid to facilitate decision making between spray timings. Rapid, in-field assays based on LAMP amplification have been developed previously {Fraaije, 2014 #60}, these could be used in two ways. Firstly, to identify SNPs present within a population, or secondly, to identify haplotypes present in a population by using multiple assays on DNA from individual isolates of the pathogen following *in-vitro* culture. The former doesn't resolve the haplotypes present in a population, whilst the latter approach is slow, requiring the isolation, culturing and testing of a large number of individual isolates with a large number of SNP tests to deliver population scale data. Radhakrishnan et al. (2019) utilised Nanopore sequencing to rapidly identify strains of *Puccinia striiformis f.sp. tritici* in the field. They showed that individual rust pustules contained DNA from single isolates and that by extracting DNA from a large number of individual pustules followed by sequencing, the strain makeup of the population could be elucidated (https://nanoporetech.com/events/agrigenomics-solutions-online-event-qa). Previous work in our laboratory (unpublished data) working on *Z. tritici* showed that clonal DNA could not be extracted reliably from individual lesions. However, we present data showing that by making single DNA extracts from whole leaves we were able to directly elucidate the presence and incidence of haplotypes making up the population. The method was suitable for estimating prevelance to approximately 5% of the population and could be achieved within 72 hours of sampling.

It was envisaged that a better understanding of the presence and frequency of haplotypes within fungal pathogen populations could potentially be used for decision making related to fungicide spray applications in two ways. Firstly, it could be used to identify the most frequent haplotypes within a population, which could be used to guide within season modifications of the spray programmes. For this to be effective the time

between sample collection and generation of results should be sufficiently rapid to facilitate decision making between spray timings. The method developed here may prove suited to this application. Secondly, the method could be used to develop a surveillance programme to identify new azole resistance mutations, which can then guide the implementation appropriate resistance management practises to limit the development of further resistance. For this to be successful, the method ought to be able to detect haplotype frequencies at very low frequencies as it has previously been shown that a fungicide's effective life (number of consecutive growing seasons that the fungicide is able to provide effective control) is strongly affected by the initial resistance frequency (van den Berg et al., 2013). Given that the current method cannot accurately identify haplotypes present at a frequency below 5%, at this point in time the method offers limited value to this application. In terms of delivery of routine diagnostic testing service, the platform should be accessible, inexpensive and rapid in use, enabling the results to be relevant for decision making in the field. In these terms the ONT platform is an attractive proposition. The ONT platform is scalable from 7 Tb of data on a 48 cell ProMethION through MinION flow-cells that can produce 30 Gb of data each, down to an adapter for smaller capacity flow-cells (Flongle) which can produce up to 2 Gb of data. Flongle flowcells currently cost £75 making them suitable for a small number of samples. This is a major advantage over many other HTS platforms were higher reagent costs mean that more samples are needed to per run to keep the costs practice. This scalability is coupled with low capital costs for the lower throughput devices making them an affordable proposition for many labs. From a practical perspective this allows diagnostic services to access the platform suited to the level of service they aim to provide. Data generation using the ONT platform is rapid, it is possible to generate data from samples within a single day which is sufficiently rapid to enable the between modification of spray programmes (Fraaije, 2014).

In this study it was shown that nanopore sequencing could be used to elucidate the haplotype composition of a *Z. tritici* field population, in relation to azole resistance using single DNA extracts from whole leaves. The method was able to accurately identify haplotypes present at a frequency of approximately 5% of the total population with results being available within 72 hours after receipt of samples. This should enable the identification of the major members of a population cycling within a crop, between sprays, facilitating more accurate selection of MOAs best suited to control the active

population in real-time. This could contribute to better disease control in the short term and a reduction in the resistance selection pressure in the long term.

2.5 Future Research

This experiment shows that ONT has a great potential for the analysis of mixed samples. However, isolates present in a low concentration within a mixture could be mistaken with errors. There are several ways where sequencing quality could be improved. The first suggestion would be improving the basecalling quality by training the basecaller using Taiyaki (Wick, 2019). An initial trial was developed during the current experiment just using the reference of the *CYP51* region of the IPO323 isolate but did not work as expected. As a future experiment it will be suggested to use several *Z. tritici* genomes (data not available at the time) instead of using amplicons. Another quality improvement method could be to create a bigger library where recurrent errors might happen on the SNPs of interest, see which strand is more accurate and modify the python script where the pysam module is used and select the SNPs from the most accurate strand.

Chapter 3. Understanding the relationship between the *CYP51* haplotype and azole resistance

3.1 Introduction

In 2004, the International Union of Pure and Applied Chemistry (IUPAC) defined a bioassay as the "association between dose and the resulting magnitude of a continuously graded change, either in an individual or in a population" (M. Nordberg, 2004) and hence bioassays, otherwise known as sensitivity tests, require the measurement of a response at different dose levels (Holland-Letz and Kopp-Schneider, 2015). Sensitivity tests are widely used in pharmacology and toxicology, to determine the dose-response relationship of a drug. For example, sensitivity tests have been used to characterise an analog of the lysergic acid diethylamide (Halberstadt *et al.*, 2019), to check the immunomodulatory effects of organophosphate in mammals (Mitra *et al.*, 2017). Equally, they have been used to determine the resistance of an organism to a fungicide, biocide or pesticide, for example, to determine the sensitivity of *Zymoseptoria tritici* to azoles, SDHIs and even for the screening of the overexpression of the transporters (Omrane *et al.*, 2015; Yamashita and Fraaije, 2018).

Sesitivity tests are used in fungicide resistance to calculate the effective concentration (EC_{50}), i.e. the concentration at which 50% disease control is achieved, which is the main measure used to calculate the resistance of a fungal pathogen to a fungicide. The results can then be fitted to different models, ranging from simple linear log-dose and exponential models, through to more complex quadratic and logistic models, in order to find dose-response relationships (Pinheiro, 2006). The most appropriate model depends on the application used and needs to be decided on a case by case basis.

A better understanding of the natural *Z. tritici* population structure could help develop a more targeted fungicide programme. In the last decades several studies have tried to understand the relationship between the *Z. tritici* phenotype and its haplotype by linking fungicide resistance, measured by sensitivity tests (phenotype), to the presence of certain mutations in the target gene (haplotype). However, this is still poorly understood. In 2007, Leroux et al. (2007) grouped the different *CYP51* genotypes in three phenotypic classes according to azole sensitivity/resistance: sensitive (TriS), low

resistant (TriLR) and moderate resistant (TriMR). Sentitive isolates were characterized by the absence of mutations in *CYP51* while TriLR and TriMR were characterized by the presence of several mutations. However, four years later, a more detailed classification was published (Leroux and Walker, 2011), which included the impact of the efflux pumps and as a consequence a new phenotypic group was added, i.e. TriHR, which was formed by genotypes which are highly resistant to azoles (

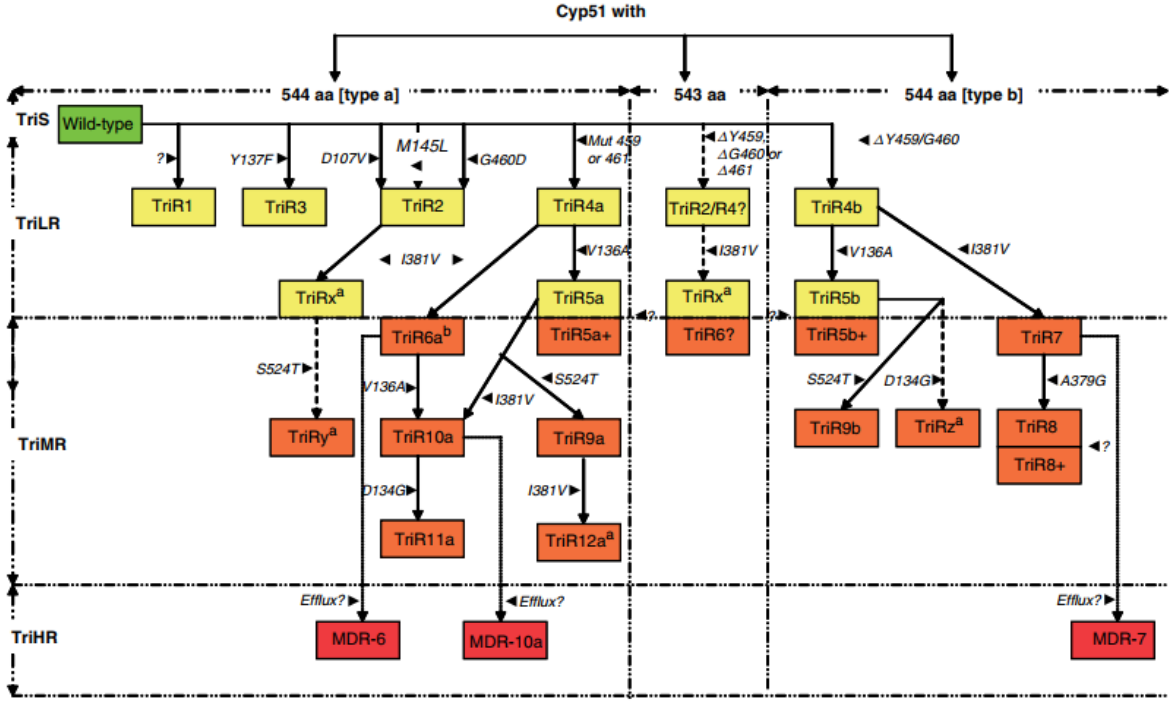


Figure 3.1). Polymorphisms at codons 459, 460 and/or 461 were present in resistant isolates although their role was uncertain, however, the overpression of the *CYP51* gene contributed to a higher resistance. Moreover, an increase of the TriMR isolates was reported, resulting in a decreased efficacy of most DMIs with the exception of epoxiconazole and prothioconazole. However, the efficacy of these fungicides has also been been decreasing in recent years (Blake *et al.*, 2017; Jørgensen *et al.*, 2018).

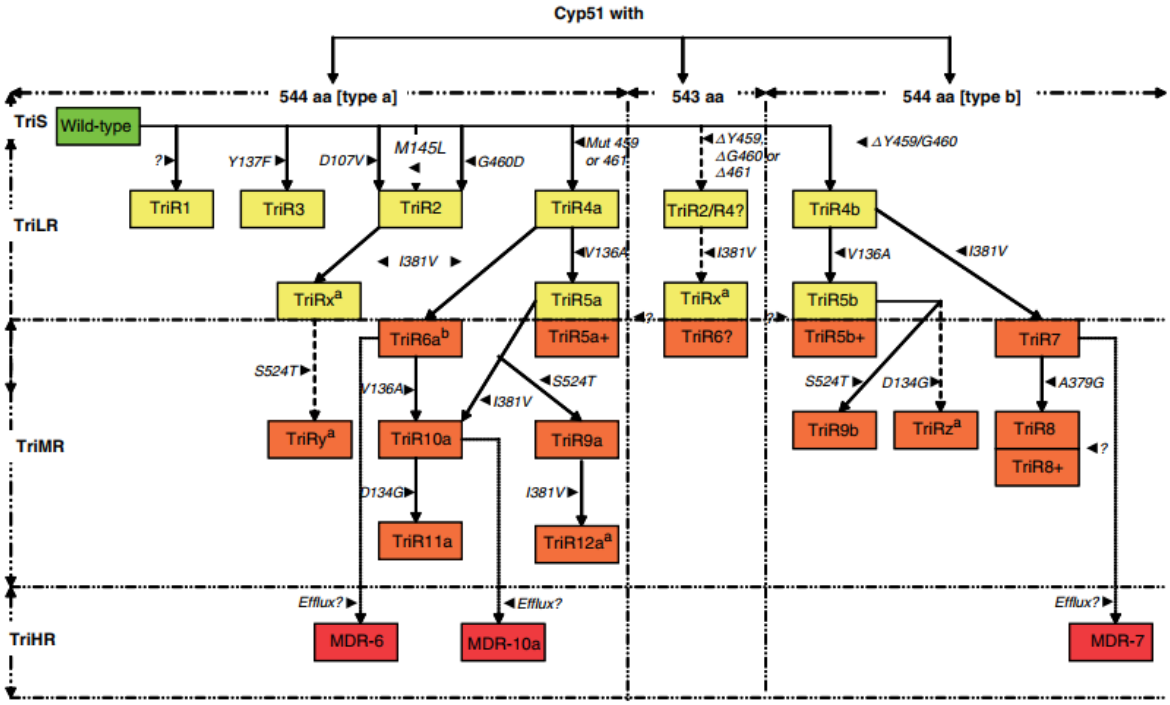


Figure 3.1 Phenotypic classification of *Z. tritici CYP51* genotypes, courtesy of Leroux and Walker (2011)

A more systematic classification of the *Z. tritici* haplotypes, was suggested by Huf *et al.* (2018) which included on a more complex nomenclature consisting of a letter followed by a number. The letter indicates the number of alterations, while the number describes a specific combination of a defined number of amino acid alterations. This study reported that mutations D134G and V136A are frequently found together and are now found across most of Europe. Isolates with mutations D134G and V136A showed a decreased sensitivity to epoxiconazole and prothioconazole-desthio. However, these isolates remained sensitive to other azoles, such as difenoconazole and tebuconazole (Leroux and Walker 2011; Buitrago et al. 2014).

In 2011, Mullins *et al.* (2011) reported the effects of the mutations in the protein structure. The authors, classified the single mutations or combination of mutations in three different groups: alterations that do not affect the azole binding, alterations affecting the size and accessibility of the binding pocket and alterations affecting the localisation of residues interacting with azoles. For example, mutation L50S and combinations of mutations L50S, S188N and N513K have no effect on the azole binding. Isolates with mutation I381V, which decreases the size of the binding pocket, were always observed in combination with L50S, whereas haplotypes with mutations

L50S, I381V and Y461H were also reported to have no effect in azole resistance. However, most of the amino acids, individually or in combination, increase the size of the binding pocket: e.g. haplotypes with mutations L50S, V136A and Y461H result in an small increase, from 1658 to 1971 $Å^3$ and haplotypes with mutations L50S, S188N, ΔY459/G460 and N513K resulting in a much larger volume increase of 5275 $Å^3$. Finally, haplotypes with mutations Y459D, G460D, Y461H and Y461S, lead to a doubling of the cavity size and affect the localisation of residues interacting with azoles, with as a main effect Y137 being withdrawn from the binding pocket.

Developing a better understanding of the *CYP51* population structure in *Z.tritici* isolates is very important to farmers, as the analysis of the mechanisms involved in azole resistance can help them choose a spray programme which will give good control of the disease in the short term and a low resistance development in the long term.

This chapter aimed to test the reproducibility of sensitivity tests and to further investigate the relationship between *Z. tritici* haplotype and its phenotype. As a first step, sensitivity tests were developed under different conditions to calculate the resistance levels of *Z. tritici* isolates in order to verify they are reproducible. As a second step, sensitivity tests were used to calculate the potency of a substance in affecting the growth of the different isolates in an effort to link haplotype to phenotype. This experiment focused on mutations found within *CYP51*, analysis of the insertions in the promoter region and the effect of transporter presence. As the last step, a principal component analysis (PCA) and phylogenetic tree were used to identify pontential patterns in mutations occurence and haplotype grouping. The combination of the analysis of *CYP51* including mutations in the target gene and insertions in the promoter region in combination with testing of the overexpression of transporters using tolnaftate can help to understand the *Z. tritici* phenotype and could lead to a better fungicide spray programme selection. As an example, if two dominant isolates were found within a population, one of them being highly resistant to a single fungicide and the other one sensitive to most of the fungicides then the spray programme should exclude the product to which the first isolate is highly tesistant. This would ensure both good disease control and a reduced further resistance development.

3.2 Materials and methods

3.2.1 Origin of the isolates and isolate selection

DNA and cultures obtained from *Zymoseptoria tritici* infected leaf samples from spray trials performed in South Wales and Lenham during a previous project between 2015 and 2017 (Innovate UK project 102088, https://gtr.ukri.org/projects?ref=102088) were used in the experiments reported on this chapter. As explained on 2.2.1, these spray trials aimed to drive resistance development and track potential *Z. tritici* population changes as a result of the sprays throughout the season. DNA of over 2,000 isolates was sent to Eurofins Genomics (Germany) for Sanger sequencing (DNA isolation, PCR and data analysis explained in section 2.2.5) in order to identify the *CYP51* haplotype, represented by a unique combination of *CYP51* mutations and also to sequence the upstream regulatory region of *CYP51* which ranged from 0 to 800 bp aproximately. A total of 89 different haplotypes were identified based on the sequence of *CYP51* only and independently of the size in the regulatory region.

3.2.2 Isolate selection and preparation for sensitivity tests

A subset of 28 haplotypes was selected to study whether the resistance expressed (phenotype) could be linked to individual mutations (Table 3.1). Haplotypes were selected by the mutations pattern, choosing haplotypes with the same mutations except one single nucleotide polymorphism. As an example Type 3 and Type 14 both have the mutations L50S, V136A, S188N, A379G, I381V, ΔY459/G460 and S524T, except N513K which is only present in Type 14 (Table 3.1). Therefore the aim of comparing these two haplotypes was to see if differences in resistance levels could be explained by the presence/absence of the mutation N513K.

To determine the sensitivity or resistance levels of the isolates to different azoles, bioassays were developed. Cultures were defrosted and 10 µL of each isolate was cultured on Potato Dextrose Agar (PDA) in a light room at 19°C. After 5-7 days spores were suspended in general purpose media (GPM, prepared by disolving 14 g of dextrose, 7.1 g bactopeptone, 1.4 g yeast extract in 1 L of water and autoclaved at 121°C for 15 min) and adjusted to 10^4 spores mL^{-1}.

Table 3.1 *Zymoseptoria tritici* CYP51 haplotypes used in this study.

The first column refers to the name of the isolates or haplotype; the second row represents the amino-acid position and the third row the reference amino acid (IPO323). No letter means no change, a dash refers to the deletion of the amino acid and a capital letter refers to the amino acid observed in that position.

	Amino acid changes													
	50	134	136	145	188	208	377	379	381	459	460	461	513	524
IPO323	**L**	**D**	**V**	**M**	**S**	**S**	**I**	**A**	**I**	**Y**	**G**	**Y**	**N**	**S**
IPO235						N								
IPO94269				F										
Type 1	S	G	A						V			H		
Type 2	S				N				V	-	-		K	
Type 3	S		A		N			G	V	-	-			T
Type 4	S	G	A						V			H		T
Type 5	S	G	A						V	-	-		K	
Type 6	S		A						V			H		T
Type 8	S		A					G	V			S		T
Type 10			A					G	V			S		T
Type 11	S				N			G	V	-	-		K	
Type 14	S		A		N			G	V	-	-		K	T
Type 15	S		C		N			G	V	-	-			T
Type 18	S		A		N				V	-	-			T
Type 19	S		A		N					-	-			T
Type 20	S		A		N					-	-		K	T
Type 22	S		C		N				V			H		T
Type 24	S						V		V			H		
Type 25	S		A		N				V	-	-		K	T
Type 27	S	G	A					G	V	-	-		K	T
Type 32	S		A		N			G	V			S		T
Type 34	S	G	A		N			G	V	-	-		K	T
Type 35	S	G	A		N			G	V	-	-			T
Type 39	S	G	A					G	V	-	-			T
Type 40	S	G	A						V			H	K	T
Type 44	S		A		N			G		-	-		K	T
Type 46	S		A		N			G		-	-			T
Type 59	S				N	T		G	V	-	-		K	
Type 60	S				N			G	V	-	-		K	T
Type 65	S								V			H		

3.2.3 Azole fungicide sensitivity tests

Sensitivity of the isolates to azoles was determined using 96-well microplates (Sarstedt, Nümbrecht, Germany) as described by Kildea (2009). In summary, analytical grade azoles (cyproconazole, metconazole, prothioconazole-desthio and tebuconazole) purchased from Sigma-Aldrich (UK) were each dissolved in 100% methanol (Fisher Scientific) at 1 mg mL^{-1} and stored at 4°C until required. To determine the sensitivity of each isolate to the different fungicides, the growth of the isolates in the presence of a serial dilution of the fungicide was measured. Of the active ingredient, 100 µL was added to the flat-bottomed plate with 150 µL of spore solution at 10^4 spores mL^{-1}. Four replicates per isolate were considered. Fungal growth was determined by measuring the absorbance at 405 nm (Thermo Max microplate reader,

Molecular Devices). Plates were read at time 0 and again after being placed on a shaker incubator at 180 rpm in a light room at 19°C for 7 days.

3.2.4 Assessing the repeatability of the sensitivity tests

To test the repeatability of the bioassays, i.e. to assess the variation in measurements taken by the same person under the same conditions, a total of 4 different haplotypes were randomly selected (3, 4, 14 and 34) and two isolates were tested from each of the haplotypes. Sensitivity of types 3, 14 and 34 to cyproconazole, metconazole and tebuconazole were tested, while for technical reasons, type 4 was only tested against metconazole and tebuconazole. In addition, type 4 was also tested using prothioconazole-desthio. Every individual isolate was cultured on two different PDA plates (i.e. 1A and 1B) and subsequently a sporangial suspension was prepared from each of the cultures (two biological replicates). Each biological replicate was grown on two different microtiter plates (moreover, four replicates within each microtiter plate), resulting in a total of two biological replicates and two sporalgian suspension per isolate (i.e. 1A and 1B were grown in two different plates each) (Figure 3.2). Moreover, two different wild types (IPO323 and IPO268) were used as a control per experiment and to test if similar EC50 values were obtained from the sensitive isolates with no mutations in *CYP51* and without an insertion in the promoter region. The sensitivity of the isolates to the different azoles was tested. Fungicides were dissolved in methanol Cyproconazole, metconazole, prothioconazole-desthio were diluted to final concentrations of 30, 10, 3.33, 1.11, 0.37, 0.12, 0.04 and 0 mg L-1 while tebuconazole was diluted to final concentrations of 75, 25, 8.33, 2.78, 0.93, 0,31, 0.1 and 0 mg mL-1.

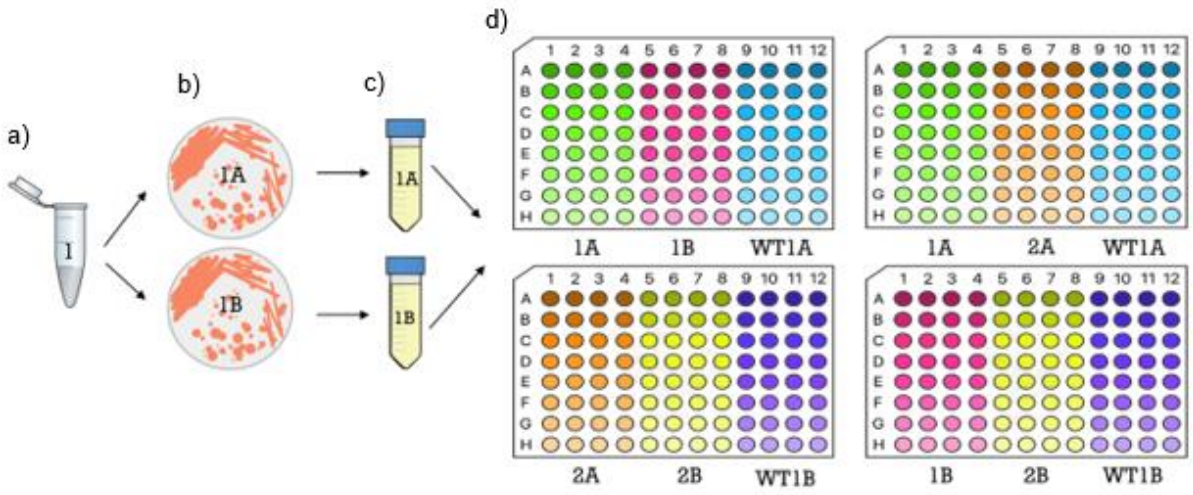


Figure 3.2 Bioassay workflow.

a) Isolate stored in milk, b) isolate cultured on potato dextrose agar c) sporangial suspension in general purpose media at 10^4 spores mL^{-1} and d) microtiter plates with spore suspension and active ingredient, each color represent an isolate and the darkness indicates the concentration of the fungicide (the darker the color the higher concentration of the fungicide). WT = Wild type.

3.2.5 Pump upregulation screening

Sensitivity assays with the fungicide tolnaftate (Signam-Aldrich) were used for the screening of the upregulation of the efflux pumps. Tolnaftate has been used in recent years as an indicator of the pumps being upregulated and therefore contributing to the exportation of the fungicide and resulting in a reduced fungal growth inhibition. Four replicates of each isolate were tested against a final concentration of 10, 5, 2.5 and 0 mg L^{-1}. To calculate the EC_{50} of isolates which grew in the presence of tolnaftate these isolates were tested again in a wider range of fungicide concentrations, i.e. 10, 5, 2.5, 1.25, 0.625, 0.313, 0.156 and 0 mg mL^{-1}. Of the active ingredient, 100 µL was added to the flat-bottomed plate with 100 µL of spore solution at 5 x 10^4 spores mL^{-1}. Plates were incubated in the dark at 19°C for 4 days. Fungal growth was determined by measuring the absorbance at 620 nm.

3.2.6 Calculation of the most common parameters in fungicide resistance (Efective concentration 50 and resistance factors)

An R programme was created to automate the EC_{50} calculations by performing a regression of the percentage of control on different concentrations of each active

ingredient. The programme uses the “drc” package, which allows the analysis of dose-response curves (Ritz and Streibig, 2005; Ritz *et al.*, 2015). The EC_{50} can be derived from the equation describing the change in percentage control due to changes in the active ingredient concentration.

$$f(x(b,d,e)) = \frac{d}{1 - \exp\{b(\log(x) - \log(e))\}} \quad (1)$$

A three-parameter logistic model with $f(x(b,d,e))$ the percentage disease control was based on, e as the EC_{50}, d as the upper limit and b as the relative slope around EC_{50}. Under the assumption that disease control can not be lower than 0% or exceed 100% the lower and upper limits were set to 0 and 100%, respectively.

To standardise sensitivity values resistance factors (RFs) were calculated by dividing the EC_{50} of the resistant isolate (or isolate of interest) by the EC_{50} of the wild type (or most sensitive isolate). The wild type isolate IPO323, is considered as the gold standard, but EC_{50} calculations did not fit the model for all fungicides, therefore the average of the EC_{50} values from wild types were calculated and used as sensitive EC_{50}. Scaling in such a manner allowed the classification of the fungicide resistance in 3 different groups: sensitive (RF < 25), moderate resistant (25 < RF < 100) and highly resistant (RF >100) (Leroux and Walker, 2011).

3.2.7 Discovering patterns in mutation occurrence and haplotype grouping

A principal components analysis (PCA), a widely used unconstrained ordination technique to visualise patterns in the data was performed to evaluate potential patterns in mutation occurrence. Given that the PCA focuses on genotype only, this analysis was perfomed using a total of 1,821 isolates which were successfully sequenced from the Innovate UK project. The PCA was generated in R using the ‘rda’ function of the vegan package (J Oksanen, 2020).

Additionally, a phylogenetic tree was generated to assess how haplotypes are grouped together. A single sequence from each of the 28 haplotypes used in the reproducibility experiment were included in the phylogenetic tree. Sequences were aligned using mafft (Katoh *et al.*, 2005); a maximum-likelihood tree was subsequently generated

using the standard parameters on iqtree2 (Minh *et al.*, 2020) and the tree was visualised and exported using FigTree (v1.4.4) (http://tree.bio.ed.ac.uk/software/figtree/).

3.3 Results

3.3.1 Assessing the reproducibility of the fungicide sensitivity tests

The sensitivity of 8 isolates, two biological replicates of haplotypes 3, 4, 14 and 34, were tested in the reproducibility experiment, while wild types were used as a negative control. Wild type isolates grew in the absence of fungicide while they did not grow in the presence of the fungicides, however EC_{50}s could not be calculated since the same fungicide concentrations as the resistance isolates were used and did not fit the model. In this experiment, 81.2% of the sensitivity tests had a significant fit, enabling estimates of EC_{50} to be made (Table 3.2). For all eight isolates, similar EC_{50} values were obtained in all four replicates, indicating that the results were highly repeatable. Moreover, EC_{50} values were also similar within each type, with the exception of Type 14 where EC_{50}s were similar for metconazole (1.265-1.575 mg L^{-1}) but different levels of sensitivity to cyproconazole were found within the isolates. The EC_{50} values of the isolate SEPT1385 were between 16.728 and 16.743 mg L^{-1} while on isolate TRET0953 the EC_{50} varied between 6.567 and 10.268 mg L^{-1}. For tebuconazole the EC_{50}s on isolate SEPT1385 varied between 2.460 and 3.228 mg L^{-1} while the sensitivity of isolate TRET0953 to tebuconazole varied between 14.680 and 16.797 mg L^{-1}. High sensitivity tebuconazole was found within types 3, 4 and 34, with EC_{50} values variying between 0.510 and 3.706 mg L^{-1}. In addition, a high sensitivity to prothioconazole-desthio was detected on Type 4 (0.125 - 0.330 mg L^{-1}). In terms of sensitivity to cyproconazole, type 34 was more sensitive to this fungicide (4.513 - 8.180 mg L^{-1}) than Type 3 (10.655 – 62.639 mg L^{-1}).

Table 3.2 Sensitivity (EC_{50} mg L^{-1}) of isolates from the repeatability experiment.

Dashes indicate insignificant model fits and ND not determined, not tested, EC_{50} values.

	Active ingredient sensitivity (EC50 mg L^{-1})

Haplotype	Isolate	Replicate	Cyproconazole	Metconazole	Prothioconazole-desthio	Tebuconazole
Type 3	SEPT1316	A1	14.922	0.724	ND	2.088
		A2	10.682	-	ND	2.129
		B1	-	1.750	ND	3.415
		B2	-	-	ND	-
	TRET0912	A1	35.726	1.291	ND	0.872
		A2	10.655	1.226	ND	0.510
		B1	-	1.566	ND	1.874
		B2	62.639	1.171	ND	2.564
Type 4	SEPT1417	A1	ND	-	0.193	0.449
		A2	ND	0.387	0.330	0.537
		B1	ND	0.531	0.245	0.902
		B2	ND	0.424	0.223	0.771
	TRET0486a	A1	ND	0.521	0.251	0.797
		A2	ND	0.475	0.366	1.060
		B1	ND	-	0.125	0.848
		B2	ND	0.436	0.263	1.075
Type 14	SEPT1385	A1	16.728	-	ND	3.228
		A2	16.732	-	ND	3.018
		B1	16.743	1.575	ND	2.816
		B2	-	-	ND	2.430
	TRET0953	A1	7.109	1.487	ND	16.797
		A2	8.611	1.265	ND	16.332
		B1	6.567	1.555	ND	16.210
		B2	10.268	1.391	ND	14.680
Type 34	TRET0375	A1	8.180	1.048	ND	2.951
		A2	6.070	2.413	ND	1.170
		B1	-	1.699	ND	3.706
		B2	-	1.156	ND	0.640
	TRET0897	A1	-	3.684	ND	2.885
		A2	4.513	1.749	ND	1.576
		B1	-	-	ND	2.403
		B2	-	0.606	ND	0.876

3.3.2 Discovering patterns in mutation occurrence and haplotype grouping

A total of 1607 samples, collected between 2014 and 2018, were successfully sequenced for all the mutations. A PCA was generated using the 1607 haplotypes in which a central cluster was found (Figure 4.4). The central grouping of points (around the zero-zero position of PC1 and PC2) simply indicates mutations that were

ubiquitous (e.g. L50S and I381V present in over 97% of the isolates), mutations in low frequencies (D107V, T165P, V136G, S208T, N284H, I377V, A410T, Y459D, Y459S, Y461S) or isolates without an insertion in the promoter region of *CYP51* (less than 1% of the population). In Fig. 3.4, mutations S188N, Y459del and Y460del were grouped relatively close together, whilst N513K and 120 bp insertion in the promotor region also had relatively low PC1 scores. Higher PC1 scores were obtained for mutations D134G and Y461H, which were relatively similar in ordination space, and also V136A and the 800 bp insertion in the promoter region.

Moreover, a maximum-likehood tree was generated using one sequence from each haplotype (Figure 3.4). This tree showed two distict groups, 1 and 2, with a subdivision of group 1 into 1A and 1B. There are two haplotypes which were not included in these groups, these are the wild type represented as IPO323a on the top of the tree and type 24 with the mutations L50S, I377V, I381V and Y461H, on the bottom of the tree. Subgroup 1A is characterised by the absence of the mutation S188N, a mutation which is present in subgroup 1B and group 2. Subgroup 1B is characterised by the presence of mutations L50S, V136C, S188N, I381V, Y461H and S524T, with the exception of haplotype 63 which did not contain the mutation V136C. The mutation L50S was present in all haplotypes apart from type 9 and 10, which where annotated as L50.

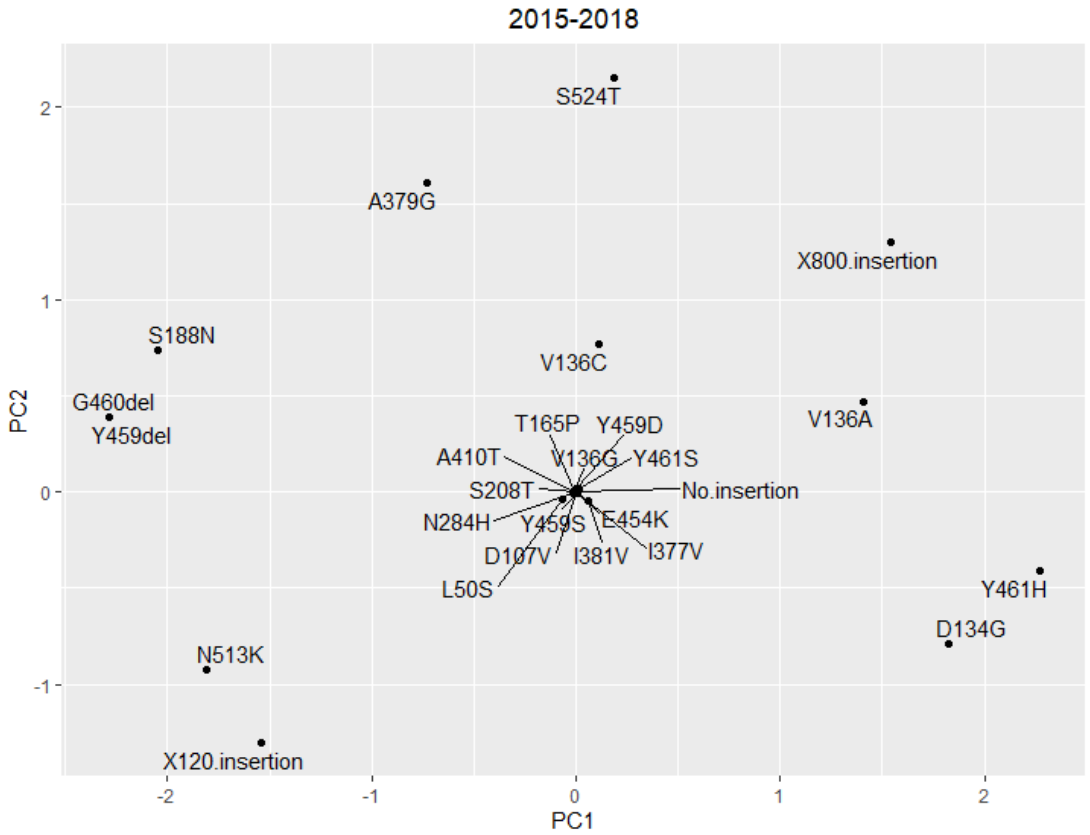


Figure 3.3 Principal Components Analysis representing the distribution of the *CYP51* mutations and the insertion in the promoter region of the gene using isolates collected from 2015 to 2018. The first and second unconstrained axes (PC1, PC2) represent 50.84% and 20.99% of the variation, respectively.

Table 3.3 Z*ymoseptoria tritici* CYP51 mutations found in the 2015-2018 isolates

a) Mutations in the target gene and b) insertions in the promoter region

	Mutation	n	Frequency (%)
a	L50S	1571	97.76
	D107V	2	0.12
	D134G	548	34.10
	V136A	911	56.69
	V136C	287	17.86
	V136G	2	0.12
	T165P	2	0.12
	S188N	972	60.49
	S208T	2	0.12
	N284H	4	0.25
	I377V	5	0.31
	A379G	407	25.33
	I381V	1563	97.26
	A410T	5	0.31
	E454K	3	0.19
	Y459del	813	50.59
	Y459D	4	0.25
	Y459S	3	0.19
	G460del	813	50.59
	Y461H	771	47.98
	Y461S	13	0.81
	N513K	523	32.55
	S524T	815	50.72
b	0 bp	20	1.24
	120 bp	407	25.33
	800 bp	1180	73.43

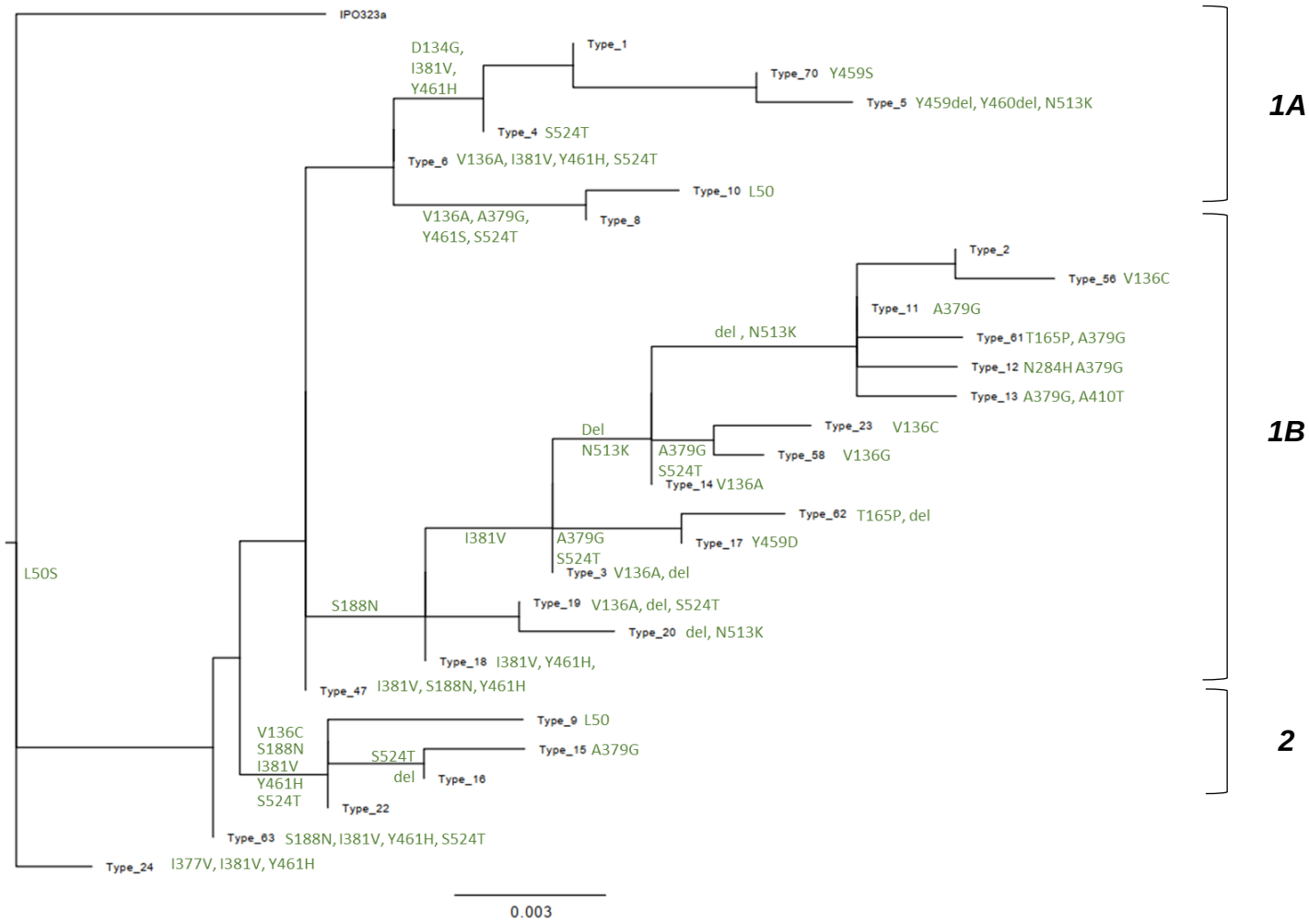


Figure 3.4 Maximum-likelihood tree representing all the haplotypes found between 2015 and 2018. Mutations are indicated in green.

3.3.3 Understanding the relationship between the three known mechanisms of azole resistance and resistance development

The sensitivity of 49 isolates to cyproconazole, metconazole, prothioconazole-desthio and tebuconazole was tested and EC_{50}s were calculated (Table 3.4). The majority of the sensitivity values to cyproconazole ranged from 5 to over 50 mg L^{-1}, with the exemptions of types 10,11, 24 and isolate TRET0175 from type 1 which was between 1.121 and 4.228 mg L^{-1}. Sensitivity values for metconazole are generally high (low EC_{50}) and ranged from 0,152 to 5.583 mg L^{-1} with the exemptions of types 11, 18, 22, 59 and 66 which ranged from 7.394 and 14.330 mg L^{-1}. Sensitivity values for prothioconazole-desthio are below 2 mg L^{-1} with the exemption of isolates TRET0194 an TRET0355a, from type 3 and 18 respectively. Sensitivity to tebuconazole are below 5 mg L^{-1} with the exemptions of types 2, 15, 22, 59, 60, 66 and isolate TRET0953 from type 14 which are above 5 mg L^{-1}

3.3.3.1 Calculating resistance factors (RFs)

The most effective active ingredient used in this experiment was prothioconazole-desthio, where most of the isolates were sensitive except isolates TRET0194 and TRET0355a form types 3 and 18 respectively and type 27 which were moderate resistant. However, the least effective active ingredient was cyproconazole, where most of the types were highly resistant to cyproconazole, with the exemption of type 24 where isolate SEPT1347 had low resistance and TRET0393a which was moderate resistant. Moreover, types 8, 10, 11 and isolates TRET0175 from type 1 and TRET0277 from type 60 are moderate resistance to cyproconazole. The other two active ingredients had similar patterns, most of the isolates were sensitive or moderately resistant to metconazole and/or tebuconazole. A few isolates were highly resistant to metconazole, such as isolate SEPT0222 from type 1, isolate TRET0177 from type 11, isolates TRET0182 and TRET0656 from type 22 and type 59 anf 66. Similarly, isolates SEPT1278, TRET0953, SEPT1392, TRET0182 and TRET0277 from types 2, 14, 15, 22 and 60 respectively, were highly resistant to tebuconazole. Regarding the overexpression of the efflux pumps, tolnaftate is used as an indicator of the transporters being upregulated. These experiments indicate that transporters were highly overexpressed on isolate TRET0953 from type 14 (91.8 fold) and isolate TRET0897 from type 34 (23.096 fold), whereas lower levels of expression were found on isolates TRET0424b (type 4), TRET0182 (type 22) and some other isolates where

RF were close to the wild type, sensitive, isolate, isolates such as SEPT0222, TRET0175, SEPT1347 and 2006-218.

Table 3.4 Sensitivity (EC_{50} mg L^{-1}) of isolates to cyproconazole, metconazole, prothioconazole-desthio and tebuconazole.

Type	Insertion	Isolate	Cyproconazole	Metconazole	Prothioconazole-	Tebuconazole	Tolnaftate
1	800a	SEPT0222	17.780	7.811	-	4.128	0.792
1	800a	TRET0175	4.228	0.450	-	-	0.633
2	120	SEPT1972	8.752	0.768	0.052	6.114	NT[a]
2	120	SEPT1278	22.827	-	0.077	12.463	ND[b]
2	120	TRET0424b	17.745	2.465	0.157	-	4.263
3	800b	SEPT1316	29.859	1.135	0.534	2.713	NT
3	800b	TRET0194	11.996	5.583	2.768	2.281	ND
3	800b	TRET0912	-	1.041	0.248	-	ND
4	800a	SEPT1417	-	-	0.151	1.079	ND
4	800a	TRET0486a	-	1.221	0.954	0.970	ND
4	800a	TRET0924	17.258	-	0.403	1.040	ND
4	800a	TRET460b	-	0.362	0.192	0.477	NT
5	800b	TRET0217	-	0.542	0.059	2.858	NT
6	120	TRET0902	-	1.836	1.484	4.086	ND
6	800a	TRET0965	-	3.287	0.976	3.954	ND
8	0	SEPT1585	6.832	1.364	-	-	ND
10	0	SEPT0841	3.473	0.655	0.154	0.236	NT
11	120	TRET0177		8.726	-	-	ND
11	800b	SEPT1393	2.794	1.062	-	2.758	ND
14	120	TRET0132	-	-	-	1.343	NT
14	800b	SEPT1385	29.582		0.367	1.325	ND
14	800b	TRET0953	10.655	1.500	0.045	14.363	45.900
15	800b	SEPT1392	28.993	-	0.222	17.989	ND
15	800b	TRET0247	-	2.106	0.091	11.927	NT
18	120	TRET0355a	43.587	5.256	5.177	4.771	NT
18	800b	SEPT1164	-	-	0.940	-	ND
19	800b	TRET0180	-	-	-	0.146	ND
22	120	TRET0182	20.702	7.394	0.504	16.384	4.900
22	120	TRET0656	-	12.651	0.335	-	ND
22	800a	TRET0184	13.290	-	-	-	ND
22	800a	TRET0519b	10.276	1.135	0.066	8.684	ND
24	800a	SEPT1347	1.121	0.646	-	-	0.624
24	800a	TRET0393a	2.492	1.477	-	3.033	ND
25	800b	TRET0071	24.480	2.356	1.406	1.797	NT
27	800b	TRET0967	-	3.272	1.878	3.694	ND
32	0	TRET0059	10.093	0.407	-	0.881	ND
34	800b	TRET0897		1.578	0.685	2.340	11.548
35	800b	TRET0292b	39.335	1.886	1.021	2.822	ND
39	800b	TRET0451a	-	1.212	1.136	4.082	ND
40	800a	TRET0526a	13.766	-	0.672	-	ND
44	800b	TRET0235	13.823	1.369	-	1.765	NT
46	800	TRET0193	11.063	-	-	0.388	NT
59	120	TRET0675	17.473	10.274	0.502	7.045	0.750
60	800b	TRET0227	10.032	-	0.064	-	ND
60	800b	TRET0277	6.712	3.450	0.058	13.144	NT
65	800a	2006-218	-	0.152	0.013	1.119	0.528
65	800a	TRET0445a	-	1.118	0.487	1.643	NT
65	800a	TRET519c	>50	0.152	-	1.041	NT
66	800b	SEPT1410	9.802	14.330	0.490	10.574	NT

[a]not tested samples

[b]EC_{50} were not determined in samples which tested negative in the first screening

Dashes indicate insignificant model fits

Table 3.5 Resistance factors.

In green, sensitive isolate (RF < 25), in orange moderate resistant isolates (25 < RF < 100) and in red highly resistant isolates (RF > 100) (Leroux *et al.*, 2007)..

2	Insertion	Isolate	Cyproconazole	Metconazole	Prothioconazole-desthio	Tebuconazole	Tolnaftate
1	800a	SEPT0222	221.316	108.517		34.399	1.584
1	800a	TRET0175	52.633	6.252			1.266
2	120	SEPT1972	108.948	10.674	0.870	50.947	NT
2	120	SEPT1278	284.137		1.283	103.855	ND
2	120	TRET0424b	220.886	34.253	2.623		8.526
3	800b	SEPT1316	371.677	15.765	8.897	22.604	ND
3	800b	TRET0194	149.319	77.563	46.128	19.009	ND
3	800b	TRET0912		14.469	4.136		ND
4	800a	SEPT1417			2.524	8.991	ND
4	800a	TRET0486a		16.970	15.893	8.083	ND
4	800a	TRET0924	214.819		6.722	8.668	ND
4	800a	TRET460b		5.027	3.192	3.972	NT
5	800b	TRET0217		7.530	0.982	23.813	NT
6	120	TRET0902		25.515	24.726	34.052	ND
6	800a	TRET0965		45.665	16.272	32.951	ND
8	0	SEPT1585	85.045	18.953			ND
10	0	SEPT0841	43.227	9.101	2.571	1.968	NT
11	120	TRET0177		121.227			ND
11	800b	SEPT1393	34.782	14.759		22.987	ND
14	120	TRET0132				11.195	NT
14	800b	SEPT1385	368.222		6.122	11.040	ND
14	800b	TRET0953	132.636	20.838	0.753	119.691	91.8
15	800b	SEPT1392	360.894		3.700	149.910	ND
15	800b	TRET0247		29.255	1.519	99.388	NT
18	120	TRET0355a	542.553	73.027	86.289	39.758	NT
18	800b	SEPT1164			15.665		ND
19	800b	TRET0180				1.219	ND
22	120	TRET0182	257.695	102.729	8.408	136.535	9.8
22	120	TRET0656		175.761	5.576*		ND
22	800a	TRET0184	165.432				ND
22	800a	TRET0519b	127.906	15.773	1.106	72.370	ND
24	800a	SEPT1347	13.957	8.982			1.248
24	800a	TRET0393a	31.021	20.527		25.272	ND
25	800b	TRET0071	304.724	32.727	23.437	14.977	NT
27	800b	TRET0967		45.464	31.306	30.781	ND
32	0	TRET0059	125.631	5.652		7.341	ND
34	800b	TRET0897		21.929	11.410	19.504	23.096
35	800b	TRET0292b	489.631	26.198	17.016	23.520	ND
39	800b	TRET0451a		16.843	18.932	34.013	ND
40	800a	TRET0526a	171.351		11.207		ND
44	800b	TRET0235	172.067	19.014		14.712	NT
46	800	TRET0193	137.706			3.234	NT
59	120	TRET0675	217.495	142.741	8.373	58.705	1.5
60	800b	TRET0227	124.872		1.069		ND
60	800b	TRET0277	83.551	47.926	0.964	109.535	NT
65	800a	2006-218		2.118	0.216	9.326	1.056
65	800a	TRET0445a		15.533	8.119	13.695	NT
65	800a	TRET519c	250.000	2.109	0.000	8.671	NT
66	800b	SEPT1410	122.007	199.099	8.170	88.115	NT

[a]not tested samples

[b]EC_{50} were not determined in samples which tested negative in the first screening

3.4 Discussion

Sensitivity tests are needed for the understanding of resistant phenotypes in the population since they are a key assay for estimating how resistant or sensitive the isolate is to each individual fungicide.

The experiments where the reliability of the bioassays were tested showed that they were reproducible within the experiment. However, there were small fluctuations within the different replicates of the same isolate which may be caused by the variation in the number of spores present in the spore suspension, probably caused by clusters of spores which did not disaggregate. In all cases, the wild types or sensitive isolates, were tested at the same concentrations as the rest of the isolates to avoid the addition of a variability factor. Despite this, these isolates were useful as a negative control as they did not grow in the presence of the fungicide, they did not fit the model and as a result EC_{50} could not be calculated. Additionally, similar sensitivity values were obtained within each type, with the exception of Type 14, where the sensitivity of the isolate SEPT1385 to cyproconazole was higher than isolate TRET0953 and the sensitivity to tebuconazole on isolate SEPT1385 was lower that on isolate TRET0953. However this can be caused by the overexpression of the transporters on the isolate TRET0953.

The sensitivity tests of the analysed isolates were more sensitive to metconazole and prothioconazole-desthio, with lower EC_{50}s than cyproconazole and tebuconazole. Fluctuations within resistance factors of the same haplotype can be explained when haplotypes are analysed in combination with the insertion in the promoter region and/or the upregulation of the pumps is also considered. Isolate TRET0355a (Type 18) is highly resistant to cyproconazole and has moderate resistantance to the rest of the fungicides while SEPT1164 (type 18) is sensitive to prothioconazole desthio, or isolates TRET0182 and TRET0656 from type 22 are highly resistant to cyptoconazole, metconazone and tebuconazole while isolates TRET0184 and TRET0519b (type 22) are still highly resistant to cyproconazole but with lower RFs and sensitive to metconazole and moderately resistant to tebuconazole, in both cases the variability in sensitivity can be explained by the 120 bp insertion, which provides higher RFs. Moreover, similar RFs were obtained between isolates SEPT1385 and TRET0953 from type 14, both of them with the 800 bp insertion, but isolate TRET0953 is highly

resistant to tebuconazole while isolate SEPT1385 is sensitive, this can be explained by the upregulation of the pumps which is 91.8 fold higher than a sensitive isolate.

Some authors have suggested that variation in the sensitivity of *Z. tritici* to different isolates can be explained by the presence of single mutations, where the mutation S524T was reported to be resistant to all fungicides (Huf *et al.*, 2018). However, the experiments shown here suggest the analysis of the three mechanisms involved in azole resistance would be a better aproach for understanding the relationship between target and non-target site resistance and the phenotype. The mutation S524T was present in over 78% of the haplotypes present here. However, most of the haplotypes with the S524T were sentitive to prothioconazole-desthio and some of the isolates were also sensitive to metconazole and tebuconazole. Moreover, Huf *et al.* (2018) concluded that haplotypes with mutations D134G and V136A had a decreased sensitivity to prothioconazole-desthio which differs with the results obtained in this chapter where some of the haplotypes with D134G and V136A were sensitive to tebuconazole. In recent years, isolates with mutation A379G in combination with I381V dominated the European *Z. tritici* population (Stammler and Semar, 2011), which agrees with the haplotypes present in this study.

Over the last two decades an increase in the number of mutations and haplotypes have been reported (Cools and Fraaije, 2013; Huf *et al.*, 2018). The 28 different haplotypes analysed here were selected from the total of 72 haplotypes identified from 2015 to 2018, which shows the great potential of *Z. tritici* to adapt to fungicides. Moreover, the phylogenetic tree supports Dooley's thesis where S188N was the mutation which split the tree into two different clades.

In summary, this chapter shows that sensitivity tests were reliable and also the importance of analysing the three main mechanisms involved in azole resistance for a better understanding of the phenotype. This study shows the close relationship of the phenotype and the haplotype. Most of the haplotype and phenotype incongruences here were explained by the insertion in the promoter region or by the overexpression of the transporters. However, the analysis of more isolates within the same type would be needed to confirm this hypothesis and find more significant effects and also the development of qPCR or ddPCR for the confirmation of the overexpression of the transporters.

3.5 Future Research

The sensitivity values of some of the isolates did not fit the model used in this chapter, probably caused by the low range in concentrations, therefore a serial dilution with more concentrations or an initial screening of the isolate to know if the isolate is sensitive or not in order to choose a more appropriate serial dilution could provide a more efficient EC_{50} calculations. Moreover, a better method to prepare spore suspensions could improve accuracy as in some cases some clusters of spores where found and this could affect the overall count of spores. In addition, the analysis of more isolates per type would be suggested for a better understanding of the relationship between the resistance mechanisms and the phenotype. Lastly, another experiment would be needed to confirm the pumps being upregulated. In this chapter tolnaftate was used as an indicator of the pumps being upregulated. However, this will need to be confirmed by developing expression studies to test upregulation of the transporters directly, these experiments could include qPCR and/or ddPCR.

Chapter 4. Exploration of changes in the population structure of *Zymoseptoria tritici* after the use of single fungicides

4.1 Introduction

Azoles and succinate dehygrogenase inhibitors (SDHIs) are the main groups of fungicides used to control Septoria tritici blotch (STB) caused by *Zymoseptoria tritici* in wheat. Over recent decades, there has been a shift in the sensitivity of the pathogen to these fungicides and thus our ability to control this disease. The reduction in the efficacy of these fungicides has been variable worldwide, attributed to when these fungicides were released, the selected wheat variety and weather conditions, where wet conditions require of a higher fungicide application which translate to a higher selection pressure. There are three known mechanisms which contribute to the reduction of the sensitivity to fungicides. Firstly, mutations in the target genes (single nucleotide polymorpshims), which can modify the structure of the binding pocket and reduce the compatibility of the fungicide with the target protein. Insertions in the promoter region have a key function in the regulation of the gene, providing in some cases reduced sentivity to azoles. Lastly, the overexpression of the transporters is a multidrug resistance (MDR) mechanism, which reduce the concentration of the fungicide within cells by exporting it across membranes. This last mechanism is independent of the fungicide used.

Azoles were released in early 1970s, since then, over 30 mutations have been reported in *CYP51*, the target protein of azoles and over 100 different haplotypes (Cools *et al.*, 2012; Price *et al.*, 2015; Huf *et al.*, 2018), whereby haplotypes are characterized by different combinations of these mutations. Moreover, insertions in the promoter region of *CYP51* can also be a source of resistance development, with S524T being one of the latest reported single nucleotide polymorphism (SNP) which is present in highly resistant isolates (Huf *et al.*, 2018; Kildea *et al.*, 2019). SNPs in the coding region of the Succinate Dehydrogenase (SDH) subunits is one on the main mechanisms involved in SDHI resistance. SDHIs were released a few decades later, in 2005 and since then over 50 mutations have been reported in the target genes which form *SDH*. This protein is formed by four subunits, subunits A, B, C and D of which B, C and D are the most studied subunits since they form part of the SDHIs binding pocket. Some

of the reported mutations have only been found following mutagenesis experiments, while others were also found in field isolates. Most of the SNPs affecting the SDHI binding and consequently resistance to SDHIs are reported to be in SDHC (Hellin *et al.*, 2020). The most common reported mutations are C-N33T, C-N34T, C-T79N/I and CN86S/A (FRAC, 2020). However, in recent years the SNP C-H152R was reported to be very frequent resulting in complete resistance to SDHIs (Hellin *et al.*, 2020). Resistance to both groups of fungicides, azoles and SDHIs, can arise, in an analogous manner to bacteria that become resistant to multiple antibiotics. Both types of fungicides are affected by multidrug resistance, one of the most studied MDR mechanism in *Z. trititi* is the overexpression of the MgMFS1 transporter (Omrane *et al.*, 2015).

In the recent years, three new fungicides became available, one azole, Revysol (BASF, Ireland) and two new SDHIs, Inatreq (Corteva, UK) and Revystar (BASF). These will become important in new STB control programmes. As new fungicides are released, novel mutation and/or haplotypes may become more important and monitoring would be highly advisable as the outcome could be used to tailor subsequent spray programmes. Moreover, this information can be used to reduce resistance development or even to control the dominance of specific strains.

The aim of this chapter is to investigate how the *Zymoseptoria tritici* population structure changes throughout the season following the application of a single azole or SDHI fungicide.

4.2 Materials and methods

4.2.1 Origin of isolates

Zymoseptoria tritici infected leaf samples were obtained from a spray trial hosted a Bishop Burton College (2.2.1). This spray trial was managed by Agrii (UK) with the aim of exploring the genetic diversity of *Z. tritici* populations throughout the season. Experimental plots were laid out randomly in 4 blocks containing 8 fungicide treatments and an un-treated plot, used as a control per block (Figure 4.1). Each plot was 1.5 m wide and 12 m long. Treatments consisted of two foliar fungicide applications (T1 at GS31-32 and T2 at GS37-39) with each treatment containing either a single SDHI or a single azole. Additionally each plot was sprayed with Corbel (BASF) in order to control powdery mildew and rust (Table 4.1). The six SDHIs used within this

experiment were: Fluxapyroxad (Imtrex, BASF), Bixafen (Inception, Bayer), Benzovindiflupyr (Ceratavo Plus, Syngenta), Penthiopyrad (Vertisan, DuPont), two new SDHIs (New SDHI 1 and New SDHI 2); and the two azoles: prothioconazole (Proline, Bayer) or a new triazole (New triazole) all sprayed as solo products. The products used at T2 were the same as those used at T1 for each treatment. Product application rates were designed to allow the infection and growth of *Z. tritici*. Leaves were collected in April 2019 before the first spray and on July 2019 after both spray treatments.

Figure 4.1 Field plots layout.

Top number (101-209) indicate the plot number while bottom number (1-9) indicate the type of treatment as explained on Table 4.1

Table 4.1 Treatments used in the spray trial.

Treatment number, name of the commercial product, active ingredient and actual amount of each product at each treatment time (T1 and T2). Same product applied on T1 was applied on T2.

Treatment number	Product/s	Active ingredient	Product applied l/ha
1	Corbel	Control	1.0
2	Imtrex + Corbel	Fluxapyroxad (SDHI)	0.75 + 1.0
3	Inception + Corbel	Bixafen (SDHI)	0.5 + 1.0
4	Ceratavo Plus + Corbel	Benzovindiflupyr (SDHI)	0.375 + 1.0
5	Vertisan + Corbel	Penthiopyrad (SDHI)	1.0 + 1.0
6	New Fungicide 1 + Corbel	New SDHI 1	1.5 + 1.0
7	New Fungicide 2 + Corbel	New SDHI 2	1.0 + 1.0
8	Proline + Corbel	Prothioconazole (Triazole)	0.36 + 1.0
9	New Fungicide 3 + Corbel	New Triazole	0.5 + 1.0

4.2.2 Yield assessments

Yield assessments were carried out using a small plot combine harverster on 26th August 2019 to measure the yield per plot. The length of the plots was measured and together with the moisture content of the grain from each plot the yield corrected to 85% dry matter was determined.

4.2.3 Nanopore sequencing

Z. tritici isolations were performed from May 2019 till September 2019 and subsequent DNA extraction, August 2020, were perfomed as explained in 2.2.1 and 2.2.4. Isolations. A total of 446 isolates were selected for the analysis, 57 isolates from control plots, 86 isolates from the plots treated with triazoles and 303 samples from plots treated with SDHIs (Figure 4.2). Only the *CYP51* region was analysed in the plots treated with azoles and only the 4 subunits of the SDH were analysed in the plots treated with SDHIs, whereas all were analysed in the isolates obtained from the control or untreated plots.

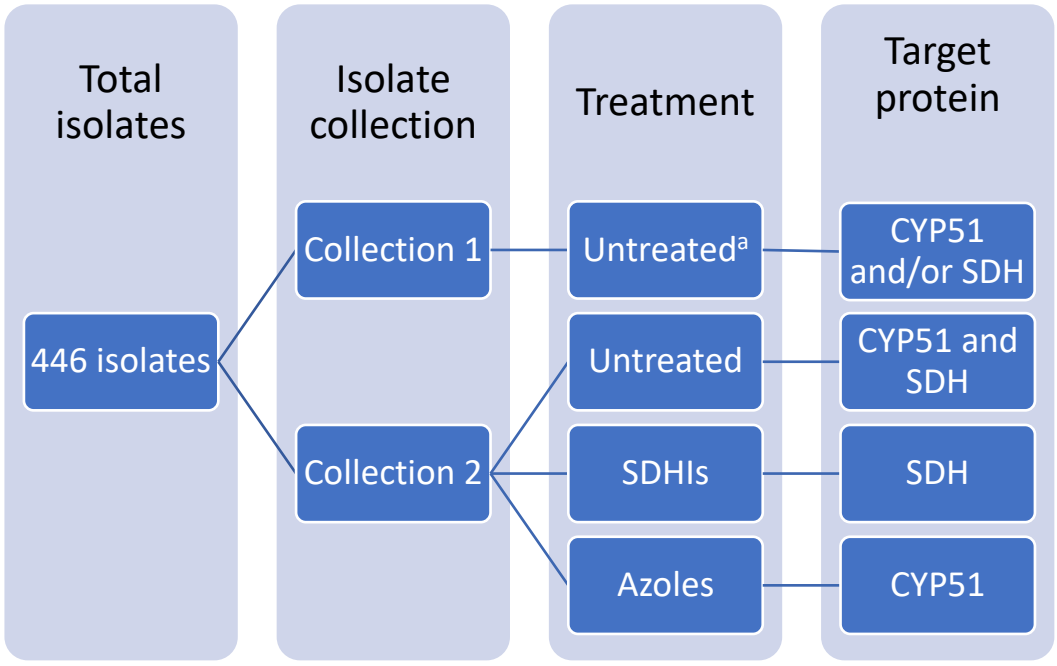


Figure 4.2 Schematic representation of origin of the isolates.

From left to right, total number of analysed isolates; collection 1, isolates collected before any treatment and collection 2, isolates obtained after both spray treatments (T1 and T2); applied treatment and analysed target protein. [a] Samples collected before the first fungicide spray were considered as untreated, but target protein analysis was done according to any subsequent fungicide treatment

PCRs were performed using tailed primers specific to amplify the full length of the *CYP51* gene including the promoter region or the full length of each of the SDH subunits. Reaction mixes were prepared as follows: 1 µl of DNA, 12.5 µl of LongAmp Taq 2X Master Mix, 1 µl of each primer at 10 µM and 9.5 µl of molecular grade water. The PCR conditions varied depending on the gene, whereby all reactions started with an initial denaturation at 94°C for 30 s followed by 30 cycles of denaturation at 94°C for 30 s; annealing at 59°C (CYP51), 56°C (SdhA), 66°C (SdhB and SdhC) or 60°C (SdhD) for 30 s; extension at 65°C for 3 min 20 s (CYP51), 2 min 25 s (SdhA), 1 min 15 s (SdhB), 50 s (SdhC) or 44 s (SdhD), and followed by a final extension at 65°C for 10 min. Amplification was confirmed by 0.5% agarose gel electrophoresis.

Table 4.2 Polymerase chain reaction (PCR) primers used to amplify *CYP51* gene including the promoter region and the genes of the four SDH subunits in *Zymoseptoria tritici*

Primer	Sequence (5' -> 3')	Amplicon size (pb)
Cy51-MI-F1	TTTCTGTTGGTGCTGATATTGCCTTCGCTGAACACCTCGCTT	2,585[a]
ST51-MI-R3	ACTTGCCTGTCGCTCTATCTTCTGAAAGCAGCGTCTCCCTC	
SdhA-F	TTTCTGTTGGTGCTGATATTGCAATCCGATTCTGACTGTCTGCC	2,374
SdhA-R	ACTTGCCTGTCGCTCTATCTTCTTGCTTGCTTTCCCTTGAAGAG	
SdhB-F	TTTCTGTTGGTGCTGATATTGCAACACTCCACGCCTCACGAC	1,494
SdhB-R	ACTTGCCTGTCGCTCTATCTTCTGTTTCTCATTACCGCGGACAC	
SdhC-F	TTTCTGTTGGTGCTGATATTGCCTCTCCGCTCCGGCACAC	797
SdhC-R	ACTTGCCTGTCGCTCTATCTTCGGCACTCCCTTGGGTCCTG	
SdhD-F	TTTCTGTTGGTGCTGATATTGCCCGGCATCATCGTCAAGC	863
SdhD-R	ACTTGCCTGTCGCTCTATCTTCTCCCCCACTCCTCAAACCGT	

[a]amplicon size without an insertion in the promoter region

PCR products were purified using AMPure XP beads (Beckman Coulter, Wycombe, UK) employing a 0.5X beads to sample ratio. Purified PCR products were tagged using the PCR Barcoding expansion 1-96 (EXP-PBC096, Oxford Nanopore) according to the manufacturer's protocols. Barcoded PCR products were quantified using a Quant-iT Picogreen dsDNA Assay Kit (Thermo Fisher Scientific) in the Fluoroskan Ascent (Thermo Scientific). Samples were pooled in equal amounts of DNA to a final quantity of 1-1.5 µg and the TapeStation D5000 high sensitivity assay (Agilent Technologies, Cheadle, UK) was used as a quality control, to estimate size and quantify the

amplicons present in the libraries. A total of 7 libraries were prepared using the SQK-LSK109 sequencing kit (ONT), following the manufacturer's protocols and sequenced in different R10.3 MinION flow cells for 72 hours.

4.2.4 Nanopore sequence analysis

MinION raw reads in fast5 format were base called using Guppy v4.0.11, with the appropriate configuration file and demultiplexed using the Guppy barcoder. Nanoplot (version 1.28.1) was used to check the key features and statistics of the sequencing run (De Coster *et al.*, 2018).

MinION reads were mapped to the *Zymoseptoria tritici* IPO323 reference genome (RefSeq assembly accession: GCF_000219625.1) using Minimap2 (version 2.17) (Li, 2018) to separate the reads into the different amplicons (CYP51 and SdhA-d) using samtools (Li *et al.*, 2009). Reads from the different amplicons were filtered using Filtlong, minimum length and number of the highest quality bases were adjusted to the size of the target gene. Filtlong was used to select for the highest quality 10 million bases (CYP51, SdhA and SdhB), 2.58 million bases (SdhC) or 2.35 million bases (SdhD). All filtered reads were mapped again to the *Zymoseptoria tritici* IPO323 reference genome. Alignments were visualised using the IGV software (https://software.broadinstitute.org/software/igv/) (Robinson *et al.*, 2011; Thorvaldsdóttir *et al.*, 2013; Robinson *et al.*, 2017). Consensus sequences from the isolates were obtained with IGV and confirmed using MegaX (Kumar *et al.*, 2018). Existing haplotypes where numbered as detailed in chapter 2 whereas new haplotypes were named in alphabetical order, from "a" to "k".

4.2.5 Heatmap, PCA, NMDS analysis and table of frequencies

Each mutation was annotated by a letter, followed by a number and followed by a letter or a dash. The first letter indicates the amino acid present in the wild type sequence, followed by a number which indicates the position of the amino acid and if the number is followed by a letter the letter shows the amino acid present in the isolate and if followed by a dash indicate the deletion of the amino acid. As an example L50 shows that there is a Leucine (L) in position 50 of the amino acid of the wild type, while L50S indicates that the Leucine (L) in position 50 has been replaced by a Serine (S). A similar

notation was used for mutations present in the SDH, with a letter denoting the subunit followed by a dash and then the mutation. For example, B-N225T indicates the presence of mutation N225T in the subunit B, whilst C-N33T indicates the presence of mutation N33T in subunit C.

Data was initially visualised using heatmaps and in some cases with principal components analysis (PCA) and non-metric multidimensional scalling (NMDS) to provide an overview of the relationship between the samples and mutations. Heatmaps are a graphical presentation of data where numerical values are displayed by colors, highlighting the largest and smallest values in the data. The heatmaps were generated in R with the heatmaply function of the heatmaply package (Galili *et al.*, 2018). Data was analysed in two different sets, the first one a matrix with the data from collection 1 and a second matrix with the data of collection 2.

The PCA was developed to visualise patterns in the data and to be able to compare with the PCA performed in chapter 3 (section 3.3.2). Additionally, a NMDS was used to investigate in more detail the effect of the fungicide in the mutations pattern in more detail. PCA and NMDS were generated in R using the rda and metaNMDS functions of the vegan package, respectively (J Oksanen, 2020).

4.3 Results

4.3.1 Nanopore data analysis

The *CYP51* gene including its regulatory region and/or the genes of the SDH subunit a, subunit b and subunit c of 443 isolates were sequenced using ONT (Figure 4.2). However, only 327 out of 443 samples could be sequenced for SDH subunit D. The median read quality of the 7 libraries varied from a Phred score of 11.4 to 12.1 (accuracy 92.7 to 93.8%) (Table 4.3). Quality values were improved after the demultiplexing, dividing reads by amplicons and filtering the reads using Filtlong. After the filtering Phred scores for the median quality ranged from 12.2 to 13.2 (accuracy 94.0 to 95.2%).

Table 4.3 Main sequencing parameter obtained using Nanoplot

	SDH plate 1				SDH plate 2			
Target gene	SdhA	SdhB	SdhC	SdhD	SdhA	SdhB	SdhC	SdhD
Median read length	2,367	1,471	796	862	2,367	1,470	794	861
Median read quality	12.8	12.9	12.9	13.2	12.8	12.6	12.8	12.6
Median accuracy (%)	94.8	94.9	94.9	95.21	94.8	94.5	94.8	94.5
	SDH plate 3				SDH plate 4			
Target gene	SdhA	SdhB	SdhC	SdhD	SdhA	SdhB	SdhC	SdhD
Median read length	2,368	1,472	796	862	2,374	1,478	800	867
Median read quality	12.2	12.2	12.3	12.6	12.4	12.4	12.5	12.8
Median accuracy (%)	93.97	93.97	94.11	94.5	94.25	94.25	94.38	94.8
	Control plate				CYP51	Repeats plate		
target gene	SdhA	SdhB	SdhC	SdhD	CYP51	CYP51	SdhC	SdhD
Median read length	2,376	1,475	798	864	3,807	3,817	951	1,018
Median read quality	12.4	12.4	12.6	13.1	12.3	12.6	12.4	12.5
Median accuracy (%)	94.25	94.25	94.5	95.1	94.11	94.5	94.25	94.38

4.3.2 Analysis of the CYP51 gene and its promoter region

The *CYP51* gene and its regulatory region were successfully sequenced using ONT for 142 isolates, 74 from untreated plots (control plots), 33 from plots treated with prothioconazole and 35 from plots treated with the new triazole. Some of those isolates were collected prior to treatment (Collection 1). From these samples, 41 belonged to untreated plots, 8 from plots treated with protioconazole and 18 from plots treated with the new triazole (Table 4.4). From the rest of the isolates 33, 25 and 17 were from untreated, prothioconazole and triazole treated plots respectively as part of collection 2.

Firstly, heatmaps were generated to observe variations in mutation frequencies before and after treatment and between different plots/treatments within each collection. Figure 4.3 presents two heatmaps based on the number of times each *CYP51* mutation was present in each plot (new azole, Prothioconazole and Untreated) at the two

Table 4.4 Number of isolates tested in the current experiment

Fungicide group	Treatment	Collection 1	Collection 2
	Untreated	41	33
SDHI	Fluxapyroxad	33	19
	Bixafen	12	30
	Benzovindiflupyr	24	37
	Penthiopyrad	28	22
	New SDHI 1	13	23
	New SDHI 2	19	30
Triazole	Prothioconazole	8	25
	New Triazole	18	17

collection dates. In Figure 4.3 (top), isolates are from the first collection date before spraying (collection 1); hence, all these isolates came from untreated leaf samples but the heatmap rows are labelled according to the treatments that plots received later in the growing season. This heatmap shows that all three plots from collection 1 have broadly similar mutation patterns. More hetereogeneity between treatments were detected for samples from collection 2, after some plots had received fungicide applications (bottom of Figure 4.3). All mutations were present in both collections apart from Y159S which was only found in collection 1 at a low frequency. In general, both heatmaps show that mutation I381V is frequent regardless of when the sample was collected or treatment. Mutation A379G is present at a low frequency before and after treatment, with the exception of plots treated with the new azole in which case the mutation was more abundant. Untreated isolates from collection 2 and isolates treated with Prothioconazole have similar mutation patterns, while isolates from the plots treated with the new azole differ from the previous two. Generaly, the abundance of mutations in the plots treated with azoles are lower than untreated plots and plots treated with Prothioconazole.

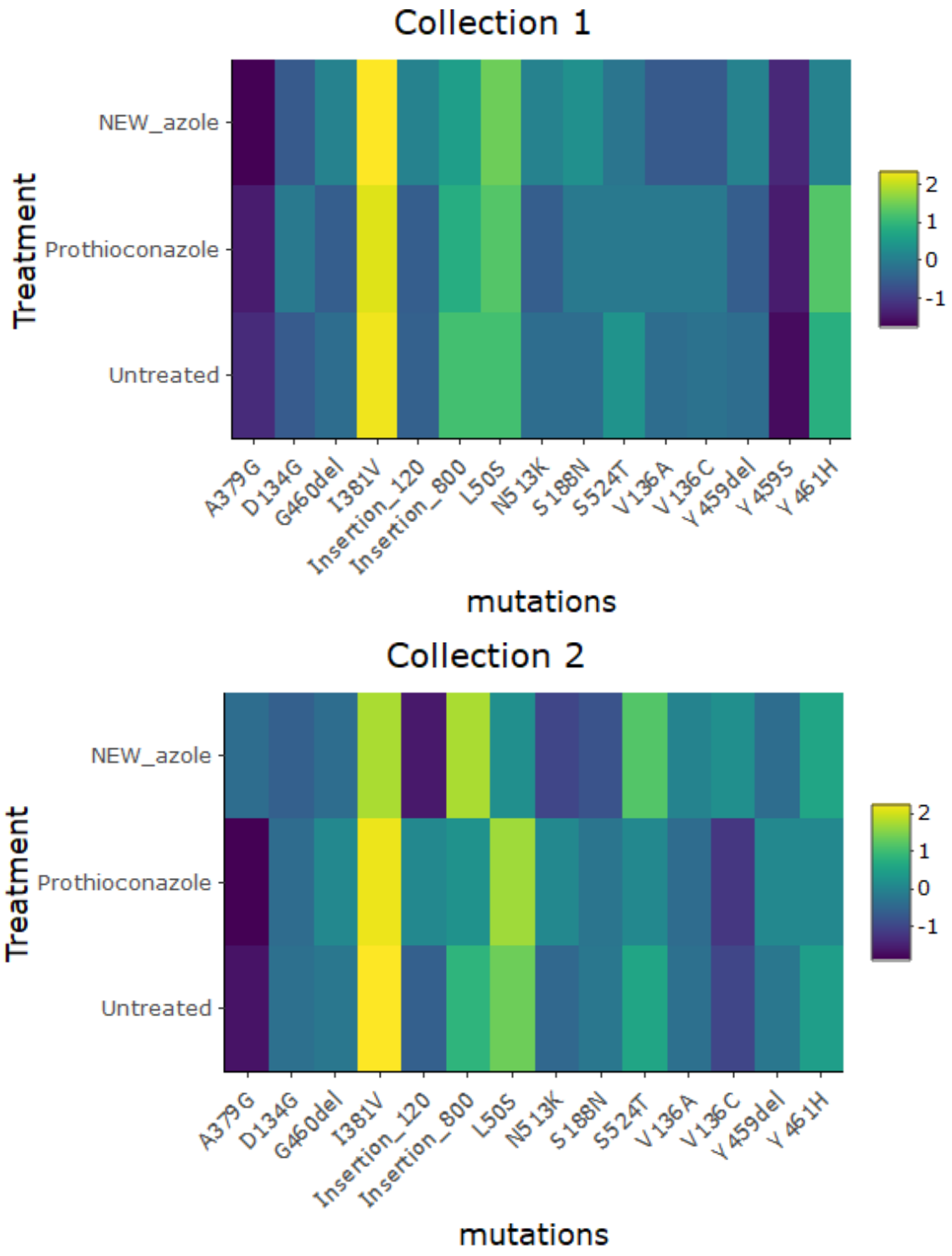


Figure 4.3 Heatmaps representing the abundance of CYP51 mutations in untreated plots and plots treated with prothioconazole or the new azole.

On the top, mutations from isolates collected before each treatment but classified by the treatment it will have in the future and on the bottom, mutations from treated plots.

Secondly, a PCA was generated to identify correlations within the CYP51 mutations found in the current experiment (Figure 4.4). All isolates from untreated plots and isolates from plots treated with azoles were included in the PCA. Mutations positioned in the origin of the PCA (0 value on both axes), indicate mutions which were ubiquotous or present at really low frequency. Mutation I381V is in the origin and A379G is close to the origin, suggesting a really low or really high frequency throughout the isolates. By comparison with the information given on the heatmaps (Figure 4.3it suggests that I381V is relatively ubiquitous across all the samples as it was found to be present at a high frequency in all plots for both sample collection times. Conversely, mutation A379G is found near the origina as it was present in a relatively low frequency in all plots for both sample collection times. There are 3 groups of mutations which indicates that those mutations normally occur together or in a similar frequencies. The first group, on the left side of the plot is formed by the 120 bp inserion, S188N, N513K and both deletions in Y459 and G460; the second, on the right side of the plot is formed by the 800 bp insertion, S524T and Y461H; and the last group is formed by D134G and V136A.

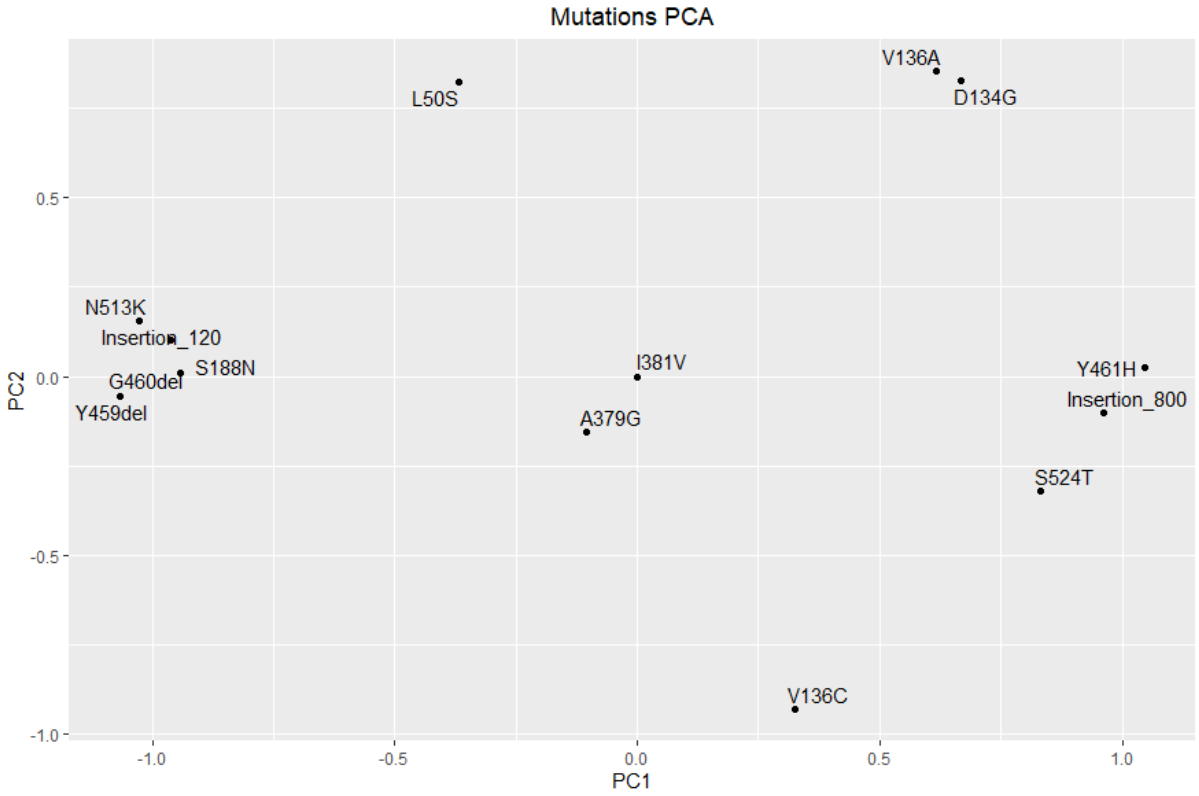


Figure 4.4 Principal Components Analysis representing the different CYP51 mutations found in the current experiment

As the last step, two non-metric multidimensional scalling or NMDS were used to investigate how the different communities are related to the different treatments. The NMDS for collection 1 (all untreated samples) shows that the plots (untreated, prothioconazole and new azole) are clustered together, suggesting that there is no difference within the treatments. However, the NMDS for collection 2 (treated samples) shows that isolates treated with prothioconazole and isolates treated with the new azole differ from each other, having a section in common, while the untreated isolates share similarities with both treatments.

The frequency of each mutation was calculated (Table 4.5) for all the untreated isolates collected before each treatment (collection 1) and isolates from the collection 2, isolates after both sprays (T1 and T2) and each by treatment (Untreated, Prothioconazole and new azole). Both untreated groups (collection 1 and 2) have similar amino acid frequencies, regardless of when the samples were collected. When comparing SNP frequencies of samples from collection 2, untreated isolates and isolates obtained from prothioconazole plots are quite similar while isolates from plots treated with the new azole differ significantly. The most frequent mutation in each treatment was I381V, which was present in every single isolate. Out of all the mutations two were only present in the untreated samples of the collection 1, Y459S and the 3,000 bp insertion (page 29). In general, the frequency of each mutation is lower in the plots treated with the new azole, apart from S524T which is present in 82.35% of the isolates and V136, the reference nucleotide (IPO323), which was not found and was divided similarly between the mutation V136A and V136C.

Regarding the insertion in the promoter region, 66% of the untreated samples had the 800 bp insertion, while 34% carried the 120 bp insertion. Both 120 bp and 800 bp were equally divided in prothioconazole plot while in the plot sprayed with the new azole just the 800 bp insertion was found.

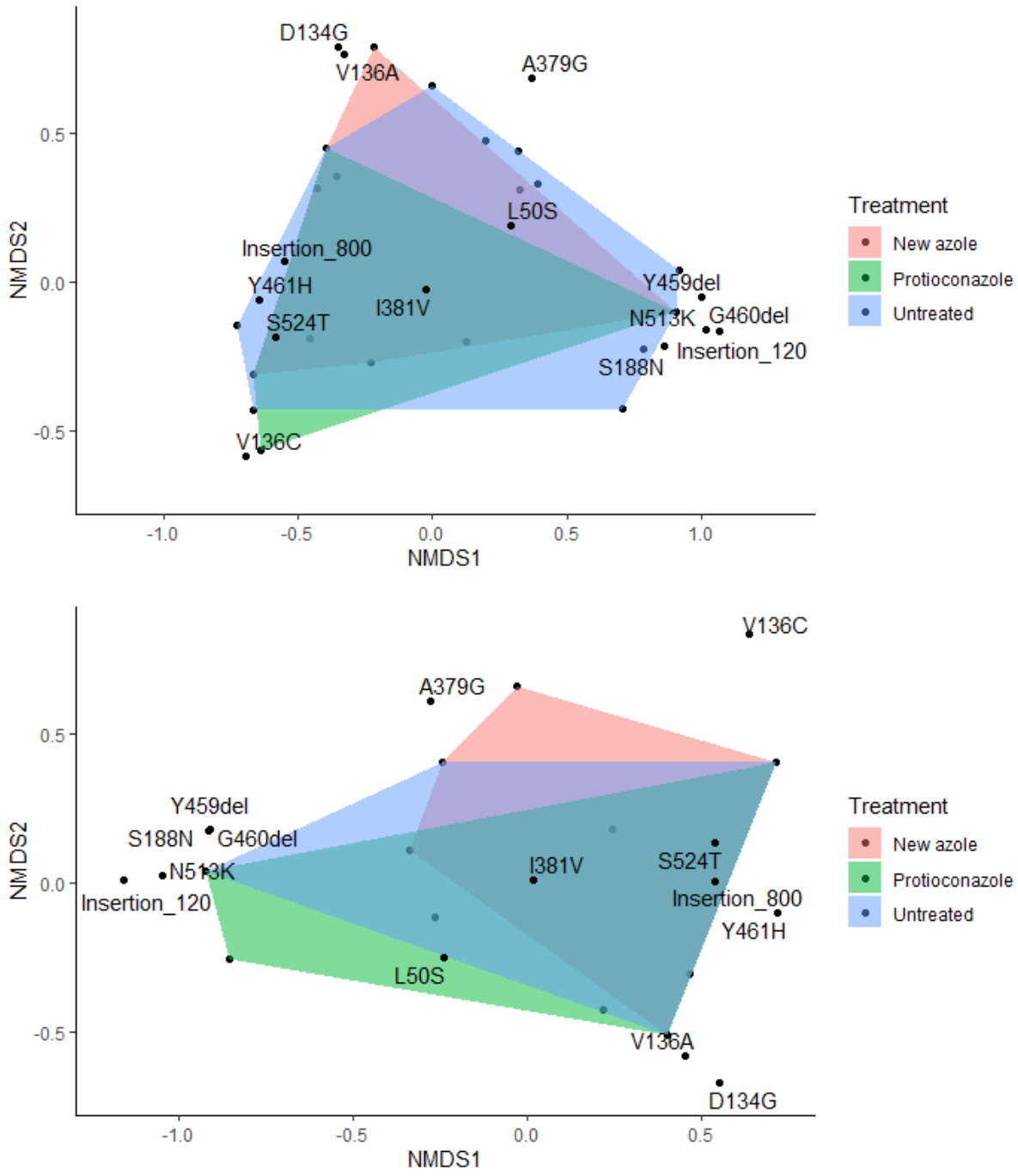


Figure 4.5 Non-metric multi-dimensional scaling (NMDS) plots displaying *CYP51* mutation detected in *Zymoseptoria tritici* isolates. The different colours indicate the plot where isolates were collected from. On the top, the NMDS from isolates collected prior treatment (stress value = 0.04866) and on the bottom isolates collected after treatment was applied (stress value = 0.01929).

Table 4.5 Frequencies of the CYP51 mutations present in the samples from Bishop Burton .

Each mutation was annotated by a letter, followed by a number and in some cases followed by a letter or a dash. The first letter indicates the amino acid in the reference isolate (IPO323), followed by the position of the amino acid and in some cases followed by the aminoacid change, a dash indicates the deletion of the amino acid.

	Collection 1	Collection 2		
Mutation	Untreated	Untreated	Prothioconazole	New azole
L50	26.87	21.21	12.00	47.06
L50S	73.13	78.79	88.00	52.94
D134	71.64	60.61	64.00	70.59
D134G	28.36	39.39	36.00	29.41
V136	32.84	36.36	48.00	0.00
V136A	32.84	39.39	36.00	47.06
V136C	34.33	24.24	16.00	52.94
S188	61.19	57.58	60.00	76.47
S188N	38.81	42.42	40.00	23.53
A379	94.03	90.91	100.00	64.71
A379G	5.97	9.09	0.00	35.29
I381V	100.00	100.00	100.00	100.00
Y459	61.19	57.58	52.00	64.71
Y459-	35.82	42.42	48.00	35.29
Y459S	2.99	0.00	0.00	0.00
G460	64.18	57.58	52.00	64.71
G460-	35.82	42.42	48.00	35.29
Y461	40.30	42.42	52.00	35.29
Y461H	59.70	57.58	48.00	64.71
N513	64.18	63.64	52.00	82.35
N513K	35.82	36.36	48.00	17.65
S524	53.73	39.39	52.00	17.65
S524T	46.27	60.61	48.00	82.35
120	32.84	33.33	48.00	0.00
800	65.67	66.67	52.00	100.00
3000	1.49	0.00	0.00	0.00

A total of 22 different haplotypes were found in the population (Table 4.6). All haplotypes had between 3 and 10 mutations. Some of the types shared the same aminoacid changes but a different insertion size in the promoter region; for instance, haplotypes 9a and 9b all contain mutations V136C, I381V, Y461H and S524T but type 9a has an 800 bp insertion and type 9b has 3,000 bp insertion in the promoter region.

Similarly, haplotypes 1a and 1b share the mutations L50S, D134G, V136A, I381V and Y461H but type 1a has an insertion in the promoter region of 120 bp while type 1b has an insertion of 800 bp. The haplotypes with the fewest mutations, are haplotype “a” with the 800 bp insertion and mutations V136C, I381V and Y461H, and haplotype “c” with the 800 bp insertion and mutations I381V, Y461H, S524T. The haplotype with most mutations was haplotype 34 with with the 800 bp insertion and 10 mutations L50S, D134G, V136A, S188N, A379G, I381V, Y459-, G460-, N513K, S524T.

Table 4.6 *Zymoseptoria tritici* CYP51 haplotypes used in this study.

The first column refers to the name of the isolates or haplotype; the second row represents the amino-acid position and the third row the reference amino acid (IPO323). No letter means no change, a dash refers to the deletion of the amino acid and a capital letter refers to the amino acid observed in that position.

		Amino acid changes										
Type	Insertion	50	134	136	188	379	381	459	460	461	513	524
wt	0	L	D	V	S	A	I	Y	G	Y	N	S
a	800			C			V			H		
9a	800			C			V			H		T
9b	3000			C			V			H		T
b	800			C		G	V	-	-			T
c	800						V			H		T
d	800	S		A			V			H		
3	800	S		A	N	G	V	-	-			T
14	800	S		A	N	G	V	-	-		K	T
e	120	S		C	N		V	-	-		K	
22	800	S		C	N		V			H		T
f	800	S		C	N	G	V	-	-			T
g	800	S		C			V			H		T
h	120	S			N		V	-	-		K	
i	800	S			N		V			H	K	
2	120	S			N		V				K	
j	120	S					V	-	-		K	
34	800	S	G	A	N	G	V	-	-		K	T
1a	120	S	G	A			V			H		
1b	800	S	G	A			V			H		
4	800	S	G	A			V			H		T
27	800	S	G	A		G	V	-	-		K	T
39	800	S	G	A		G	V	-	-			T
70	800	S	G	A			V	S				
k	800	S	G	A			V					T

The frequency of each haplotype was calculated according to the collection and treatment. As for the mutations analysis, all the isolates from collection 1 were grouped together as these isolates were not treated. A total of 20 different haplotypes were found on the untreated samples from collection 1 whilst in collection 2, between 6 and 7 haplotypes where identified in each of the treatments. The most common haplotypes from collection 1 were also the most common haplotypes in the untreated plots from collection 1, while most of the haplotypes present in low frequencies in collection 1 were not found in the untreated plot from collection 2. Haplotype 4 was identified on 8.96% of the samples in collection 1, while the frequency on the untreated samples of collection 1 was higher, at 30.30%.

Table 4.7 Frequencies of the CYP51 haplotypes present in the Bishop Burton samples.

	Collection 1	Collection 2		
Type	Untreated (%)	Untreated (%)	Prothioconazole (%)	New azole (%)
a	2.99	0.00	0.00	0.00
9a	20.90	18.18	12.00	35.29
9b	1.49	0.00	0.00	0.00
b	0.00	0.00	0.00	11.76
c	1.49	3.03	0.00	0.00
d	2.99	0.00	0.00	0.00
3	1.49	0.00	0.00	5.88
14	0.00	0.00	0.00	17.65
e	1.49	0.00	0.00	0.00
22	4.48	0.00	4.00	0.00
f	0.00	6.06	0.00	0.00
g	2.99	0.00	0.00	0.00
h	26.87	33.33	36.00	0.00
i	1.49	0.00	0.00	0.00
2	1.49	0.00	0.00	0.00
j	1.49	0.00	12.00	0.00
34	1.49	3.03	0.00	0.00
1a	1.49	0.00	0.00	0.00
1b	10.45	6.06	4.00	17.65
4	8.96	30.30	28.00	11.76
27	1.49	0.00	0.00	0.00
39	1.49	0.00	0.00	0.00
70	2.99	0.00	0.00	0.00
k	0.00	0.00	4.00	0.00

From collection 2, similar frequency patterns were found when comparing untreated and prothioconazole plots (types 9a, h, 1b and 4). The frequency of haplotype 9a in the untreated plot was 18.18%, while the frequency decreased in the plots treated with prothioconazole and increased in the plots treated with the new azole (35.29%). Some of the types were only found in one of the plots of collection 2, such as types b, c, or d between many others. Types b and 14 were present in plots treated with the new azole in more that 10% of the isolates.

4.3.3 Analysis of the four subunits of the SDH gene

The four SDH subunits (SdhA, SdhB, SdhC and SdhD) from 327 isolates were successfully sequenced, 146 untreated samples from collection 1, prior to the treatment and 181 isolates from collection 2, after treatments were applied (Table 4.4). From the isolates of collection 2, 20 belonged to untreated plots,19 from plots treated with Fluxapyroxad, 30 from plots treated with Benzovindiflupyr, 37 from plots treated with Penthiopyrad, 23 from plots treated with the new SDHI 1 and 30 from plots treated with the new SDHI 2. However, the SdhD sequence of 33 isolates could not be determined.

A total of 51 different mutations were identified within the 4 subunits, 8 belong to subunit A, 4 to subunit B, 26 to subunit C and 13 to subunit D. From the 51 mutations, 24 mutations had not been previously reported: A-S3P, A-V6A, A-V6 M, A-K49T, A-I217S, A-D480V, A-A484S and A-E487D in subunit A; no new mutations were found in subunit B; mutations C-Q42E, C-S49C, C-I78M, C-T79A, C-T79H, C-N86T, C-V91S, C-A118V, C-V128M, C-I137F, C-R151G and C-H152L in subunit C and mutations D-Q13P, D-T19A, D-T20I, D-I50L in subunit D. From the 24 new mutations, 6 mutations where found in amino acid positions where another mutation has been reported, such as C-R151G which has not been previously reported while C-R151 has been previously reported. These mutations are C-T79A, C-T79H, C-N86T, C-V91S, C-R151G and C-H152L. Those isolates which failed the sequencing of any of the subunits were discarded and they were not considered in subsequent analysis.

Following the same pipeline as detailed in 4.3.2, heatmaps were generated to provide an overview of the data. In this case, four different heatmaps were created, the first two heatmaps using the mutations which were only present before the treatment (collection 1) and another one with the mutations which were only present on isolates

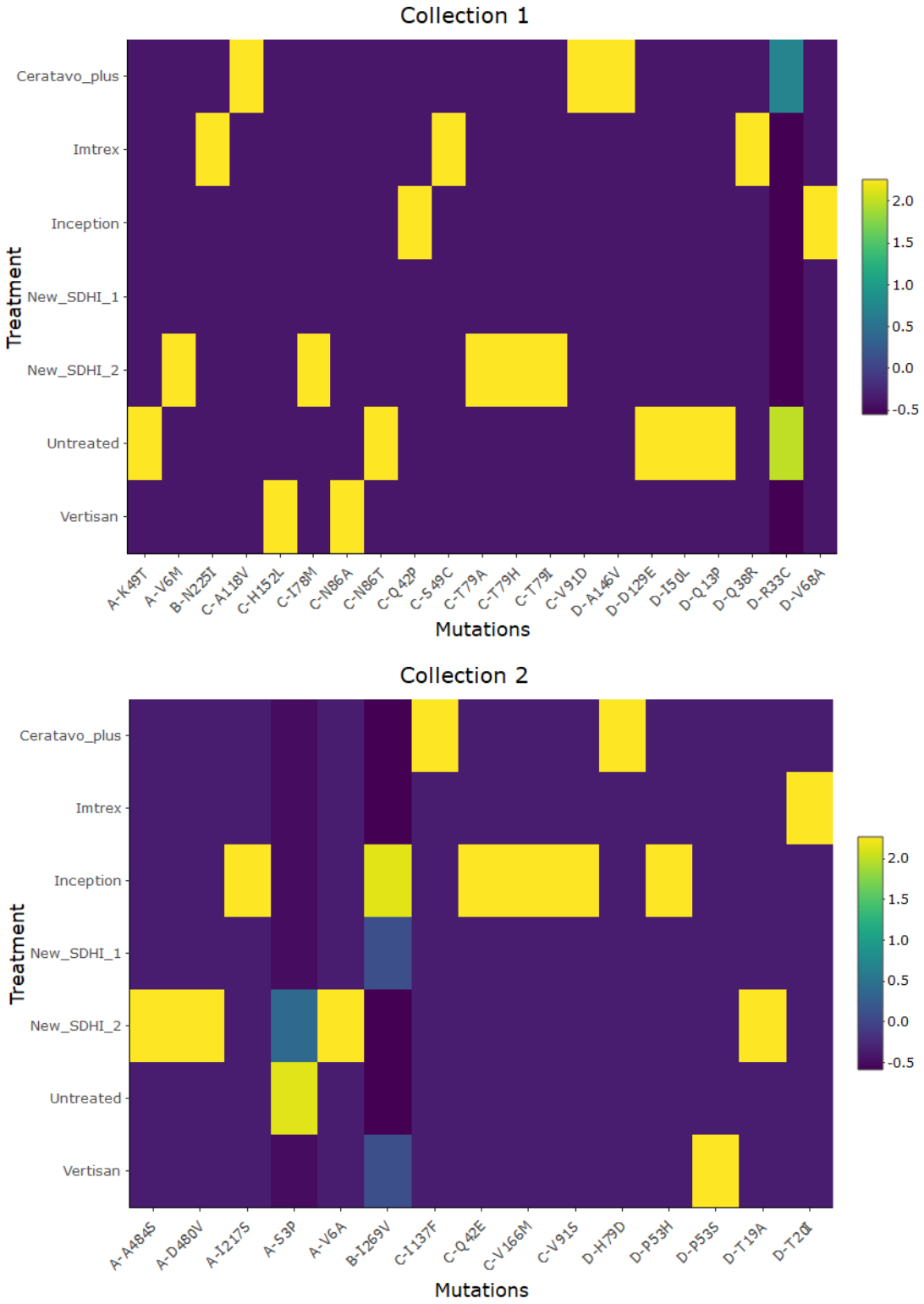


Figure 4.6 Heatmaps representing the abundance of SDH mutations in untreated plots and plots treated with ceratavo-plus, Imtrex, Inception, New SDHI 1, New SDHI 2.

On the top, mutations found only in isolates collected before each treatment but classified by the treatment it will have in the future (collection 1) and on the bottom, mutations just present from treated plots (collection 2).

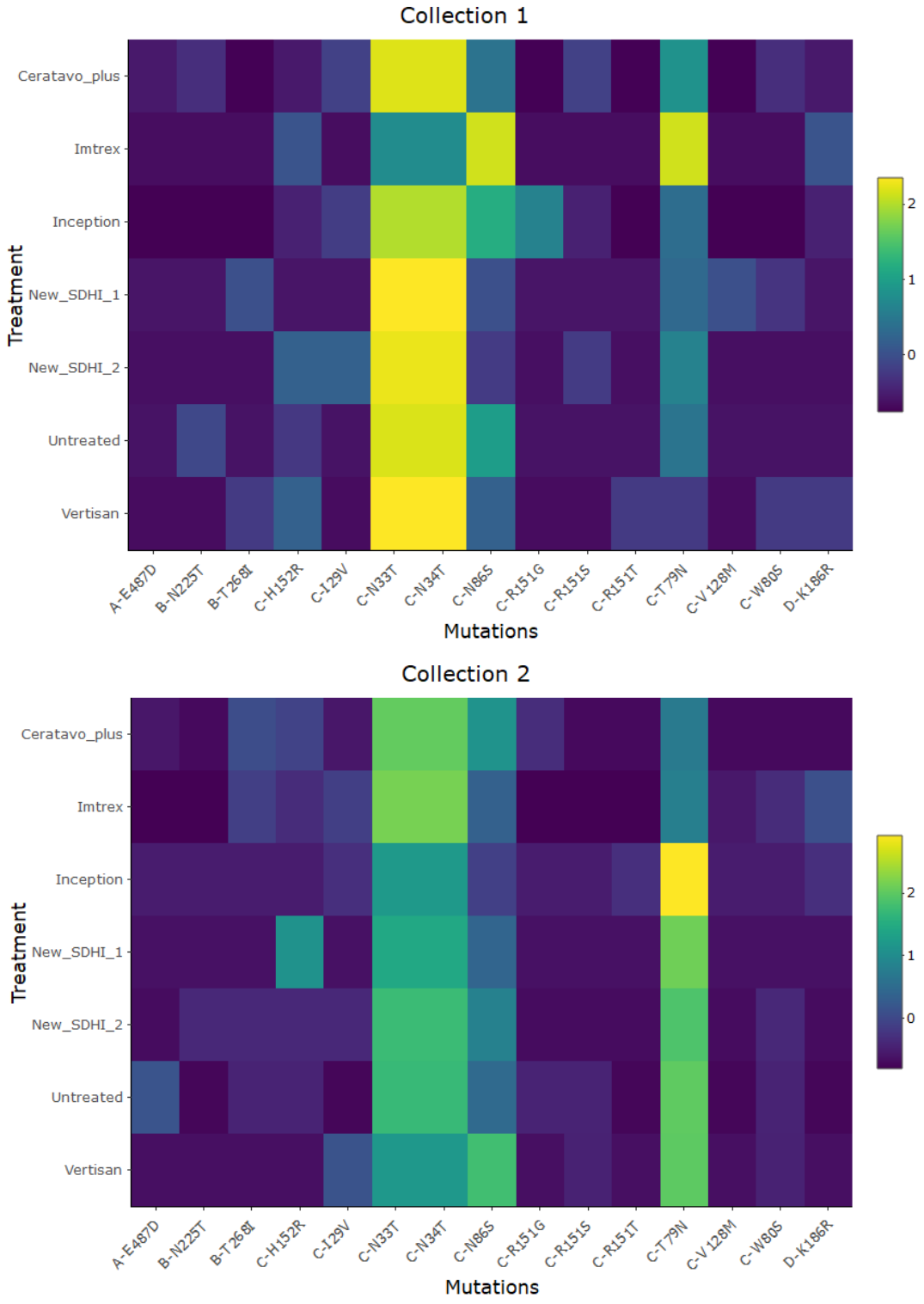


Figure 4.7 Heatmaps representing the abundance of the SDH common mutations between untreated plots and plots treated with prothioconazole or the new azole.

On the top, mutations from isolates collected before each treatment but classified by the treatment it will have in the future and on the right, mutations from treated plots.

collected from treated plots (collection 2) (Figure 4.6). The other two heatmaps were created using the mutations which were shared between collection 1 and collection 2 (Figure 4.7). In general, most of the mutations found only present before or after the treatment are only present in one of the treatments. The exceptions to this are D-R33C which was present in two different plots of collection 1 and mutations A-S3P and B-I269V which were present in two different treatments each.

Heatmaps from the mutations in common (Figure 4.7) before and after the treatment indicates mutations C-N33T, C-N33T, C-N86S and C-T79N are fairly common independent of the treatment. All mutations appear to be similarly frequent before and after the treatment, exept from C-T79N which seems more frequent in collection 2.

For a better understanding of the SDH mutations, the frequency of each were calculated according to the collection and treatment (Table 4.8). All the isolates from collection 1 were grouped together as these isolates were not treated. A total of 35 different mutations were found on the untreated samples from the collection 1 while in collection 2, between 6 and 14 mutations were identified within each treatment. Most common mutations from collection 1 (C-N33T, C-N34T, C-T79N and C-N86S) were also most common in the untreated plots from collection 2. These mutations were present in similar frequencies apart from C-T79N which increased from 19.86% in collection 1 to 45% in collection 2. Mutations in low frequency in collection 1 (A-E487D, B-T268I, C-W80S, C-R151G and C-R151S), were slightly increased in the untreated samples of collection 2. However, most of the mutations present in low frequency in collection 2 were not present in collection 2.

For collection 2, the frequency of mutations C-N33T and C-N34T increased in plots treated with Imtrex (68.42%), Ceratavo plus (56.76%), new SDHI 1 (52.17%) and new SDHI 2 (53.33%), while it decreased in plots treated with Inception and had similar values in plots treated with Vertisan (40.91%). The frequency of mutation C-H152R was considerably increased in plots treated with the new SDHI 1 (43.48%), while this was not present in plots treated with Inception or Vertisan. Similarly, the frequency of D-K186R was present at 21.05% in the isolates from Imtrex treated plots and at very low frequency or absent in the rest of the plots.

Table 4.8 Frequencies of the SDH mutations present in this study

The first row shows if the samples were collected before or after the treatment (collection 1 and collection 2 respectively). The second row show the fungicides used in the treatment and the first column indicate the amino acid change (mutation). The first letter of each mutation indicates the SDH subunit, followed by a dash and followed by a letter which indicates the amino acid in the reference isolate (IPO323), followed by the position of the amino acid and followed by the aminoacid change.

	Collection 1	Collection 2						
Mutation	Untreated	Untreated	Imtrex	Inception	Ceratavo plus	Vertisan	New_SDHI1	New_SDHI2
A-S3P	0%	15%	0%	0%	0%	0%	0%	3.33%
A-V6A	0%	0%	0%	0%	0%	0%	0%	3.33%
A-V6M	2.05%	0%	0%	0%	0%	0%	0%	0%
A-K49T	0.68%	0%	0%	0%	0%	0%	0%	0%
A-I217S	0%	0%	0%	3.33%	0%	0%	0%	0%
A-D480V	0%	0%	0%	0%	0%	0%	0%	3.33%
A-A484S	0%	0%	0%	0%	0%	0%	0%	6.67%
A-E487D	0.68%	15%	0%	0%	2.70%	0%	0%	0%
B-N225I	2.05%	0%	0%	0%	0%	0%	0%	0%
B-N225T	3.42%	0%	0%	0%	0%	0%	0%	6.67%
B-T268I	2.05%	5%	15.79%	0%	16.22%	0%	0%	6.67%
B-I269V	0%	0%	0%	13.33%	0%	4.55%	4.35%	0%
C-I29V	4.79%	0%	15.79%	3.33%	2.7%	18.18%	0%	6.67%
C-N33T	45.89%	40%	68.42%	26.67%	56.76%	40.91%	52.17%	53.33%
C-N34T	45.89%	40%	68.42%	26.67%	56.76%	40.91%	52.17%	53.33%
C-Q42E	0%	0%	0%	3.33%	0%	0%	0%	0%
C-Q42P	0.68%	0%	0%	0%	0%	0%	0%	0%
C-S49C	2.05%	0%	0%	0%	0%	0%	0%	0%
C-I78M	0.68%	0%	0%	0%	0%	0%	0%	0%
C-T79A	0.68%	0%	0%	0%	0%	0%	0%	0%
C-T79H	0.68%	0%	0%	0%	0%	0%	0%	0%
C-T79I	1.37%	0%	0%	0%	0%	0%	0%	0%
C-T79N	19.86%	45%	36.84%	53.33%	29.73%	59.09%	69.57%	56.67%
C-W80S	2.74%	5%	10.53%	0%	0%	4.55%	0%	7%
C-N86A	2.05%	0%	0%	0%	0%	0%	0%	0%
C-N86S	21.23%	20%	26.32%	6.67%	37.84%	54.55%	26.09%	33.33%
C-N86T	0.68%	0%	0%	0%	0%	0%	0%	0%
C-V91D	0.68%	0%	0%	0%	0%	0%	0%	0%
C-V91S	0%	0%	0%	6.67%	0%	0%	0%	0%
C-A118V	0.68%	0%	0%	0%	0%	0%	0%	0%
C-V128M	1.37%	0%	5.26%	0%	0%	0%	0%	0%
C-I137F	0%	0%	0%	0%	2.70%	0%	0%	0%
C-R151G	3.42%	5%	0%	0%	8.11%	0%	0%	0%
C-R151S	3.42%	5%	0%	0%	0%	4.55%	0%	0%
C-R151T	0.68%	0%	0%	3.33%	0%	0%	0%	0%
C-H152L	1.37%	0%	0%	0%	0%	0%	0%	0%
C-H152R	6.16%	5%	10.53%	0%	13.51%	0%	43.48%	6.67%
C-V166M	0%	0%	0%	10%	0%	0%	0%	0%
D-Q13P	2.05%	0%	0%	0%	0%	0%	0%	0%
D-T19A	0%	0%	0%	0%	0%	0%	0%	6.67%
D.T20	0%	0%	0%	0%	0%	0%	0%	0%
D-T20I	0%	0%	5.26%	0%	0%	0%	0%	0%
D.Q38	0%	0%	0%	0%	0%	0%	0%	0%
D-Q38P	0%	5%	0%	0%	0%	0%	0%	0%
D-Q38R	2.05%	0%	0%	0%	0%	0%	0%	0%
D-I50L	0.68%	0%	0%	0%	0%	0%	0%	0%
D-P53H	0%	0%	0%	3.33%	0%	0%	0%	0%
D-P53S	0%	0%	0%	0%	0%	4.55%	0%	0%
D-V68A	0.68%	0%	0%	0%	0%	0%	0%	0%
D-H79D	0%	0%	0%	0%	10.81%	0%	0%	0%
D-D129E	0.68%	0%	0%	0%	0%	0%	0%	0%
D-A146V	0.68%	0%	0%	0%	0%	0%	0%	0%
D-K186R	2.74%	0%	21.05%	3.33%	0%	0%	0%	0%

The isolates could be attributed to a total of 85 different haplotypes. Most of the haplotypes were present in less than 5% of the population with only 4 haplotypes present in more than 5% of the population. The most abundant haplotype was ID3, the same haplotype as the wild type which was present in 13.5% of the isolates, followed by ID15 with mutation C-T79N which was present in 12% of the isolates, ID35 with mutations C-N33T, C-N34T and C-W80S which was present in 10.7% of the isolates and ID13 with mutations C-N33T and C-N34T which was present in 7.6% of the isolates.

4.3.4 Effect of the use of a single fungicide on yield

Control plots, which were treated only with Corbel to control rust and powdery mildew, yielded 7.4 t/ha, which is considerably less than the yields attained by treated plots ($P < 0.0001$, Figure 4.8), which were between 0.2 and 1.6 t/ha higher. The highest yield increased were attained on plots treated with the new fungicides (New SDH1, New SDH2 and New Triazole) with a yield improvement of 1.6, 1.3 and 1.1 t/ha, respectively. The lowest yield improvement of just 0.2 t/ha, was associated with plots treated with Fluxapyroxad. The remaining treatments resulted in similar intermediate yield improvements, with an improvement of 0.6 t/ha when sprayed with Penthiopyrad, 0.7 t/ha when treated with Protihoconazole or Bixafen and 0.9 t/ha when treated with Benzovindiflupyr.

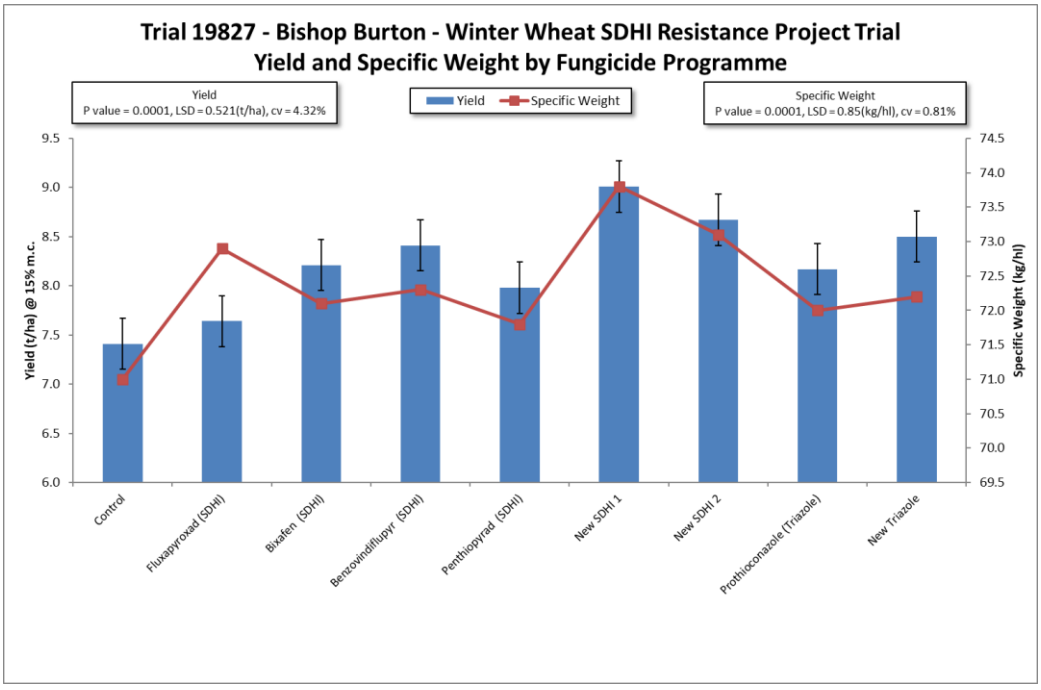


Figure 4.8 Yield and Specific Weight by Fungicide Programme

4.4 Discussion

MinION was successfully used to sequence the target genes of over 400 *Z.tritici* isolates, including *CYP51* and three out of the four subunits from the *SDH* (SdhA, SdhB and SdhC). However, the sequence of SDH-d subunit (~100 isolates) could not be obtained, due to the design of non-specific primers, which also amplified accessory copies of the gene present on different chromosomes. (Stukenbrock *et al.*, 2010) reported the presence of genes in dispensable chromosomes homologues to the core chromosomes.

The experiments presented here help to understand how *Z. tritici* populations evolves and how they are affected by the use of single fungicides. The analysis of *CYP51* shows that mutation I381V was ubiquitous, being present in each of the isolates independently of the plot treatment. The high frequency of I381V is supported by previous studies which affirm that mutation I381V is frequent in the UK (Jørgensen *et al.*, 2020a).The amino acid frequencies within the different plots from collection 1, prior to treatment were almost identical. Similarly, aminoacid frequencies obtained in collection 1 were also similar to the untreated samples from collection 2, indicating that the population structure within the untreated plots were similar at the beginning and at the end of the season. However, the population structure of the samples from collection 2, after the treatment were more diverse. Moreover, the frequency of the aminoacid S524 was higher in isolates from collection 1 when compared with the untreated samples from collection 2.

The mutation patterns between untreated plots and plots treated with prothioconazole-desthio were also similar, suggesting that prothioconazole-desthio does not have a great impact on the current *CYP51* population structure of *Z. tritici* isolates. Generally, the frequencies of the reference aminoacid (aminoacid present in the reference genome, IPO323) was decreased in the treated samples. In other words, the frequency of the reference aminoacid were lower on treated samples, indicating that *Z. tritici* population is adapting to the fungicide treatments. However, the distribution of these mutations varied on the isolates treated with the New Azole, showing higher frequencies on the reference aminoacid and suggesting there is higher STB control, caused by the lower frequency of mutations which provides resistance to the fungal pathogen. However, monitoring is still recommended to track possible shifts in the

mutations patterns, which highlights the importance of understanding the haplotype-phenotype relationship for a better disease management.

The high frequency of A379, reference amino acid, on the plot treated with prothioconazole-desthio indicates that A379G is sensitive to this fungicide. Similarly the high frequency of the 800 bp insertion indicates that isolates with the 120 bp insertion are more sensitive to the New Azole. This is interesting given that the 120 bp insertion in the promoter region has previously been shown to be linked to an increase in the resistance levels of the isolates to the azoles (references). This pattern seems to have been broken with the new azole.

A total of 22 different CYP51 haplotypes were identified, independently of the promoter region. If the promoter region were to be included the number of haplotypes would increase to 24, considering that the insertion in the promoter region can affect the sentivity to azoles (Leroux and Walker, 2011; Cools and Fraaije, 2013; Kildea *et al.*, 2019). Out of all the haplotypes, half of them were previously reported in Section 3.3.2 and some were also previously reported by other authors (Huf *et al.*, 2018; Jørgensen *et al.*, 2020a), but some new haplotypes were also identified in this study. The high number of haplotypes and the discovery of new haplotypes which are emerging over time highlights the potential of Z. tritici to adapt to the environment and the complexity of the CYP51 haplotypes (Hellin *et al.*, 2020). A large diversity of haplotypes were obtained in collection 1 (untreated) when compared to collection 2. Haplotypes a, 9b, d, e, g, I, 2 1a, 27, 39 and 70 were only present in collection 1 and in less than 3% of the population, which could be caused by a high frequency of resistant isolates after spraying, these resistant isolates are more likely to survive and consequently results in a lower diversity of haplotypes after spraying. However, it could also be partly an artefact due to the larger number of isolates sequenced from collection 1 when compared to collection 2. Additionally, some haplotypes were only present in collection 2; haplotype f was only present in the untreated plots (6,06%), haplotype k only present in plots treated with prothioconazole (4%) and haplotypes b and 14 in plots treated with the new azole (11.76% and 17.65% respectively). All these haplotypes were present at a low frequency with the exception of haplotypes b and 14 from prothioconazole treated plots which was present in over 10% of the population, indicating that the new azole has a low effectivity on those types. Haplotypes 9a, 1b, 4 and h were the most common haplotypes from the untreated plots which were also frequent on

prothioconazole treated plots and plots treated with the new azole, with the exception of haplotype h which was not present on the plots treated with the new azole. Taking into account the haplotypes along with the yield, the highest yield was obtained in the plot treated with the new azole followed by the plots treated with prothioconazole and lastly the untreated ones. This indicates that the new azole is more effective against septoria that prothioconazole. However, prothioconazole still provides protection against septoria.

The analysis of the four subunits of the SDH shows that there are several polymorphisms which occur in the different subunits and which translate into 85 different haplotypes, associated with 50 mutations. From the total of 50 mutations, 24 have not been previously reported and most of them are present at low frequencies, which is suggestive of these being new emerging mutations, some of which may not survive into the new season. From the mutations which were not previously reported, 8 belong to mutations in SDH-A, which is a subunit which is not normally analysed as it does not form part of the binding pocket. However, it was decided to include the analysis of SDH-A in this experiment since mutations in this subunit could be involved with changes in the protein structure.

The most common mutations shared between all the treatments were present in subunit SDH-C, mutations C-N33T, C-N34T, C-T79N and C-N86S. Mutations C-T33T, C-T34T and C-N86S were also the most common mutations in Slovenia (Kiiker 2020) and in the work developed by Hellin et al. (2020) where they reported mutations C-T79N/I and C-N86S to be fairly common (around 21-23%) in the population in Ireland and UK and which are linked to a moderate resistance to SDHIs, However, in this study the frequency of these mutations were relatively high with the frequency of mutation C-T79N ranging between 29% and 69% in treated samples and the frequency of mutation C-N86S ranging between 20% and 54.55% in treated samples with the exception of the plot treated with Inception, where the frequency was 6.67%. Moreover, Hellin reported a high C-H152R frequency which has been reported as a mutation involved in a high resistance of Z. tritici to all SDHIs. This mutation was not present in plots treated with Inception and Vertisan, present in 5% of the untreated samples and in the plot treated with the new SDHI 2, around 10% of plots treated with Imtrex and Ceratavo plus and a high frequency (over 40%) on plots treated with the new SDHI 1. Conversely, the high frequency of C-H152R in the plot treated with the new SDHI 1

does not correlate with experiments where the presence of C-H152R was reported to be rare in the field and linked to a high resistance to fungicides, given that in our experiments this treatment obtained the highest yield. However, it may be that the associated levels of resistance are expressed at the haplotype rather than the single mutation level and that other mechanisms such as the overexpression of the efflux transporters are at play.

From the total of 83 different SDH haplotypes, only 3 haplotypes were present in over 5% of the population. However, the main cause of the high variety of haplotypes could be caused by the high number of mutations, the higher the number of mutations the higher the probability of creating new haplotypes. The haplotype carrying mutations C-T79N, C-W80S, C-N86S, B-N225T and B-T268I is considered to have a reduced sensitivity to SDHIs (Kiiker *et al.*, 2021) but was not found in the current experiment.

In summary, the analysis of *CYP51* and the four subunits of *SDH* after the use of single azoles or SDHIs indicate that *Z. tritici* populations are being shaped by fungicide treatments. It reveals the adaptation of the pathogen through the selection pressure caused by the application of fungicides. Therefore, is important to track the population composition to be able to provide more targeted spray programmes and therewith better longer lasting disease control due to reduced resistance development.

4.5 Future Research

The analysis of the SdhD subunit for some isolates failed caused by the use of non specific primers which amplify accessory copies present in different genes. As a future experiment it would be recommended to design a new set of primers and test it with a higher number of samples. Moreover, it would be interesting to sequence a higher number of treated samples to have a better overview of the treatments.

Chapter 5. General discussion

Septoria tritici blotch is the most damaging foliar disease of UK wheat. The use of fungicides is key to control this disease which is caused by the fungal pathogen *Zymoseptoria tritici.* In recent decades there has been an increase in the resistance levels of *Z. tritici* to azoles and SDHIs, the two main groups of fungicides used to control this disease. The overall aim of the work was to explore how we could determine the overal resistance status of populations of *Z. tritici* using third generation sequencing (Minion – Oxford Nanpore) by analysisng the target genes of azoles and SDHIs without the need to isolate the fungal pathogen. Our hypothesis was that a more thorough understanding of the population would help farmers to choose a more targeted spray programme in real-time to achieve better long-term control of STB, which would reduce the waste of using an inappropriate fungicide, contribute to decreased yield losses in wheat and improve resistance management.

Three experimental chapters were designed and developed to achieve this main goal. However, each of the chapters has its own, more specific goal. Chapter 2 focused on the design and optimisation of nanopore sequencing to analyse septoria from individual isolates, mock communities and field populations. Chapter 3 focused on the understanding of the link between haplotype and resistance phenotype, to understand how variations in the target genes affect the resistance levels of the fungal pathogen to the fungicide. Chapter 4 focused on the analysis of the target genes of field samples to understand how the genetic structure of the population, changes after the application of solo fungicides.

The research in Chapter 2 used nanopore sequencing to analyse isolates and also mock communities that comprised two or three isolates. However, we demonstrated that we were only able to determine the most abundant isolates of more complex populations, due to the sequencing error rates. Nevertheless, we estimated that we were able to identify haplotypes that were present at a frequency of at least 5% of the population. Below this threshold haplotypes could be confused with artifact sequences generated by random and in particular recurrent errors. This chapter explored the use of nanopore sequencing to resolve the full length sequence of *CYP51* and the four subunits of *SDH* within populations of *Z.tritici*.

The relationship between the *CYP51* haplotype and its phenotype is not fully understood. Differences between resistance levels in the laboratory tested isolates when compared to the resistance levels in the field are fairly common. This could be caused by the isolation process, *Z. tritici* might not behave in the same way when isolated as when being part of a community with other isolates present in the same leaf.

In Chapter 3 we explored the *Z. tritici CYP51* haplotype-phenotype conundrum by analysing the three main resistance mechanisms, single nucleotide polymorphisms (SNPs), insertions in the promoter region and upregulation of the efflux transporters. This chapter demonstrates the importance of analysing haplotypes rather than mutations and to take into account all three mechanisms for a better understanding of the septoria community. However, uncertainty remains in the exact relationship between the haplotype and its phenotype, and additional research is still required. For example, isolates with the same haplotype, with the same combination of mutations in the target gene, could have different resistance values to the same fungicides. Some of these changes can be explained by having a different insertion in the promoter region or by the upregulation of the transporters, while some of the changes remain unclear due to the undetermined result for the transporter upregulation test. Moreover, this experiment was limited by the number of isolates tested, and thus we recommend that future work should expand the experiment with a higher number of isolates and the confirmation of the overexpression of the transporters to provide a better understanding of the haplotype-phenotype relationship.

In Chapter 4 we used nanopore sequencing to analyse the target genes of azoles and SDHIs. In this case since *SDH* is formed by four subunits, isolates were sequenced instead of DNA extracted from full leaves, where the probability of finding several isolates in the same leaf is really high. To obtain the haplotype of each isolate the four *SDH* subunits were considered, providing understanding of which subunit belongs to which isolate. Several isolates can occur within the same leaf, consequently different genotypes for the different subunits can be found in the same leaf, making it impossible to know which SDH combination was from the same isolate as the four subunits are coded on different chromosomes. However, sequencing populations from leaves could be achieved for *CYP51* and its regulatiory region as both are in the same region of the genome, in the same chromosome, just a few bases apart. Most experimental spray

trials are focused on which fungicide or combination of fungicides lead to the greatest yield. Conversely, in this experiment we investigated how the population changes after fungicide use. The ultimate aim was to explore if generating population level haplotype data might be useful for changing spray decisions or understanding the selection pressure exerted by the different products. This chapter shows a high number of haplotypes for both groups of fungicides, some of these haplotypes were previously reported and some of them were not, showing the high potential of *Z. tritici* to adapt rapidly to fungicides or how the population is being shaped by the use of fungicides.

Due to the pandemic caused by SARS-CoV-2 some work could not be completed. Experiments were designed at the beginning of the PhD and optimised during the following two years and a half to investigate the fitness costs for *Z. tritici* carrying mutations involved in fungicide resistance. The hypothesis behind this chapter was that by understanding the nature of the fitness costs we might improve disease and resistance management. Currently there is no evidence of fitness penalties, though most of the work has been performed in-vitro exploring characteristics of single isolates. Competition experiments were optimised in control environment (CE) rooms, to explore differences in competitiveness between pairs or small artificial populations of sensitive and resistant isolates when inoculated onto susceptible wheat varieties. To evaluate the progression of the populations in terms of presence and frequency of the isolates used to inoculate the plants we had originally planned to use digital-droplet PCR (ddPCR) to accurately quantify changes in the populations and assess competition/fitness variation between isolates. We recommend that this type of experiment is undertaken in future as it would indicate how different isolates compete between each other in mixed infections.

This research described in this thesis highlights the importance of analysing all the resistance mechanisms, including mutations in target genes, insertions in the promoter region and multi drug resistance (MDR) to understand the phenotype of individual isolates within a population. The importance of looking at the combinations of the mutations instead of single mutations as different combinations will change the structure of the binding pocket. By analysing the haplotype, the mutation combinations, can help to choose a spray programme and protect fungicides to avoid resistance development. However, this work highlights the difficulty of controlling STB due to the complexity of the field populations where farmers have to deal with several strains or

isolates with different *Z. tritici* haplotypes and adaptability of fungicide resistance in *Z. tritici.*

Nanopore sequencing can help with quick decision making and a better understanding of the population that is being controlled in a crop. This technology can be used to monitor infection even when there are no visual symptoms. This is especially important as *Z. tritici* has a long and variable latent period influenced by environmental factors, since most fungicides have low curative activity they are normally applied in a preventative programme and being able to understand the population prior to symptoms may be important for deploying effective control measures.

Long-term monitoring of *Zymoseptoria tritici* populations is highly recommended for early detection of changes in the sensitivity of the pathogen to fungicides. Monitoring is useful at a national level to help inform the industry of the changing situation. However, monitoring on its own will be insufficient due to the complexity of *Z. tritici* life-history and genome. Cosequently, researchers, growers and even the government should collaborate to ensure sustainable control of STB. In Scandinavian countries the government have limited the availability of an active ingredient across the industry. Consequently, the industry is forced to work together to manage how and when they will use the available products to better manage disease, resistance and environmental impacts.

In recent years new tools have become available which have the potential to improve global food security and sustainable agriculture. Tools such as genome editing technologies, which have progressed and become powerful genetic tools for increasing pathogen resistance in plants (Chen *et al.*, 2019). One of the most used gene editing tools is clustered regularly interspaced short palindromic repeats/CRISPR-associated protein 9, also known as CRISPR/Cas system. Improvements in CRISPR/Cas systems, have made genome editing a widely used, low-cost, easy-to-use targeted genetic manipulation tool that has been applied to many crops. CRISPR-Cas has been used to knock out *eIF4E* in cucumbers, *SWEET14* in rice, *CsLOB1* in citrus, and *DMR6-1* or *JAZ2* in tomatoes (Yin and Qiu, 2019). These tools can be used by plant breeders for rapid improvement of agricultural crops and could be used to help contolling *Z. tritici*.

The work presented here on fungicide resistance highlights how highly adaptable *Z. tritici* is, a factor that makes STB particularly difficult to control. Even if gene editing is really successful it would be dangerous to rely on one edited variety due to the high plasticity and adaptability of *Z. tritici* to changes in its pathosystem. Consequently, to control STB a sustainable management solution should combine aspects of control from genetic (conventional and gene editing) through agronomic practices, the use of biological control measures as well as synthetic fungicide use in an IPM approach. We believe that including strategies such as monitoring and diagnostics in an IPM scheme will lead to the smarter use of the available fungicides. This will result in selection of the best product to control the tested population, this will lead to reduced selection pressure that leads to wide spread resistance and potentially an overall reduction in the prophylactic use of fungicides which is common in the industry. Collectively these measures will safeguard active ingredients for the future and lead to improved environmental stewardship, which may help relieve the pressure to withdraw products from the market.

References

Adams, M.D., Myers, E.W., Ballew, R.M., Wortman, J.R., Kodira, C.D., Skupski, M., Broder, S., Zinder, N., Simon, M., Fasulo, D., Flanigan, M., Hannenhalli, S., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Bonazzi, V., Chandramouliswaran, I., Charlab, R., di Francesco, V., Ge, W., Gong, F., Ji, R.R., Lai, Z., Merkulov, G.V., Milshina, N., Nusskern, D., Salzberg, S., Wang, J., Wei, M.H., Xiao, C., Zhan, M., Zhang, W., Zhong, F., Zhu, S.C., Zhao, S., Gilbert, D., Cravchik, A., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Busam, D., Carver, A., Center, A., Lai Cheng, M., Curry, L., Desilets, R., Ferriera, S., Hladun, S., Kalush, F., Kline, L., Love, A., Mann, F., Moy, M., Moy, L., Murphy, B., Rogers, Y.-H., Romblad, D., Stewart, E., Strong, R., Thomas, R., Tse, S., Wetter, J., Williams, S., Williams, M., Wolfe, K., Campbell, M.J., Sjolander, K.V., Mi, H., Guo, N., Lippert, R., Schwartz, R., Yooseph, S., Basu, A., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y.H., Deslattes Mays, A., Dombroski, M., Fosler, C., Glasser, K., Glodek, A., Gorokhov, M., Henderson, S., Hoover, J., Jennings, D., Kagan, L., Lewis, M., McDaniel, J., Newman, M., Rowe, W., Simpson, M., Turner, R., Venter, E., Xia, A. and Zhu, X. (2001) 'The Sequence of the Human Genome', Science (American Association for the Advancement of Science), 291(5507), pp. 1304-1351.

Allioui, N., Siah, A., Brinis, L., Reignault, P. and Halama, P. (2016) 'Identification of QoI fungicide-resistant genotypes of the wheat pathogen Zymoseptoria tritici in Algeria', Phytopathologia mediterranea, 55(1), p. 89.

Bajgai, Y., Yeshey, Y., De Mastro, G., Ghimiray, M., Chhogyel, N., Tshewang, S. and Alhajj Ali, S. (2019) 'Influence of nitrogen application on wheat crop performance, soil properties, greenhouse gas emissions and carbon footprint in central Bhutan', Environmental development, 32, p. 100469.

Ben M'Barek, S., Karisto, P., Abdedayem, W., Laribi, M., Fakhfakh, M., Kouki, H., Mikaberidze, A. and Yahyaoui, A. (2020) 'Improved control of septoria tritici blotch in durum wheat using cultivar mixtures', Plant pathology, 69(9), pp. 1655-1665.

Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Fitzhugh, W., Gage, D., Harris, K., Heaford, A., Lehoczky, J., McKernan, K., Mesirov, J.P., Miranda, C., Naylor, J., Rosetti, M., Santos, R., Sougnez, C., Stojanovic, N., Subramanian, A., Wyman, D., Ainscough, R., Burton, J., Carter, N., Dunham, I., French, L., Hunt, A., Jones, M., Lloyd, C., Mungall, A., Wilson, R.K., Marra, M.A., Mardis, E.R., Pepin, K.H., Miner, T.L., Minx, P.J., Clifton, S.W., Branscomb, E., Wenning, S., Doggett, N., Cheng, J.-F., Uberbacher, E., Frazier, M., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Metzker, M.L., Naylor, S.L., Nelson, D.L., Sakaki, Y., Fujiyama, A., Toyoda, A., Itoh, T., Kawagoe, C., Taylor, T., Artiguenave, F., Bruls, T., Nyakatura, G., Rump, A., Weinstock, K., Hong, M.L., Yang, H., Yu, J., Huang, G., Rowen, L., Qin, S., Davis, R.W., Federspiel, N.A., Proctor, M.J., Chen, F., De La Bastide, M., Dedhia, N., Bateman, A., Birney, E., Brown, D.G., Clamp, M., Copley, R.R., Eichler, E.E., Haussler, D., Hermjakob, H., Hokamp, K., Kennedy, S., Kulp, D., Lowe, T.M., Mikkelsen, T., Moran, J.V., Pollara, V.J., Slater, G., Stupka, E., Thierry-Mieg, D., Wagner, L., Wallis, J., Wolf, Y.I., Yang, S.-P., Shimizu, N., Kawasaki, K.,

Athanasiou, M., Schultz, R. and Patrinos, A. (2001) 'Initial sequencing and analysis of the human genome', Nature (London), 409(6822), pp. 860-921.

Blake, J.J., Gosling, P., Fraaije, B.A., Burnett, F.J., Knight, S.M., Kildea, S. and Paveley, N.D. (2017) '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.

Borg, J., Kiær, L.P., Lecarpentier, C., Goldringer, I., Gauffreteau, A., Saint-Jean, S., Barot, S. and Enjalbert, J. (2018) 'Unfolding the potential of wheat cultivar mixtures: A meta-analysis perspective and identification of knowledge gaps', Field crops research, 221, pp. 298-313.

Bragg, L.M., Stone, G., Butler, M.K., Hugenholtz, P. and Tyson, G.W. (2013) 'Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data', PLoS computational biology, 9(4), pp. e1003031-e1003031.

Brown, J.K.M. (2002) 'Yield penalties of disease resistance in crops', Current Opinion in Plant Biology, 5(4), pp. 339-344.

Brown, J.K.M., Chartrain, L., Lasserre-Zuber, P. and Saintenac, C. (2015) 'Genetics of resistance to Zymoseptoria tritici and applications to wheat breeding', Fungal genetics and biology, 79, pp. 33-41.

Burke, J.J. and Dunne, B. (2008) 'Field testing of six decision support systems for scheduling fungicide applications to control Mycosphaerella graminicola on winter wheat crops in Ireland', The Journal of Agricultural Science, 146(4), pp. 415-428.

Cao, F., Gao, L., Wei, L., Chen, Z., Wang, Y., Ran, X., Meng, X. and Tao, J. (2017) 'Performance Comparison of NextSeq and Ion Proton Platforms for Molecular Diagnosis of Clinical Oncology', Tumori, 103(3), pp. 223-230.

Chen, K., Wang, Y., Zhang, R., Zhang, H. and Gao, C. (2019) 'CRISPR Cas Genome Editing and Precision Plant Breeding in Agriculture', Annual review of plant biology, 70(1), pp. 667-697.

Clarner, P., Lau, S.K., Chowdhury, T., Guilmette, E., Trapa, P., Lo, S.-C. and Shen, S. (2021) 'Development of a One-Step RT-ddPCR Method to Determine the Expression and Potency of AAV Vectors', Molecular therapy. Methods & clinical development.

Cools, H.J., Bayon, C., Atkins, S., Lucas, J.A. and Fraaije, B.A. (2012) 'Overexpression of the sterol 14α-demethylase gene (MgCYP51) in Mycosphaerella graminicola isolates confers a novel azole fungicide sensitivity phenotype', Pest management science, 68(7), pp. 1034-1040.

Cools, H.J. and Fraaije, B.A. (2013) 'Update on mechanisms of azole resistance inMycosphaerella graminicolaand implications for future control', Pest Management Science, 69(2), pp. 150-155.

Curvers, K., Pycke, B., Kyndt, T., Vanrompay, D., Haesaert, G. and Gheysen, G. (2015) 'A high-resolution melt (HRM) assay to characterize CYP51 haplotypes of the wheat pathogen Mycosphaerella graminicola', Crop Protection, 71, pp. 12-18.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M. and Van Broeckhoven, C. (2018) 'NanoPack: visualizing and processing long-read sequencing data', Bioinformatics, 34(15), pp. 2666-2669.

de Waard, M.A. and van Nistelrooy, J.G.M. (1979) 'Mechanism of resistance to fenarimol in Aspergillus nidulans', Pesticide Biochemistry and Physiology, 10(2), pp. 219-229.

Del Sorbo, G., Schoonbeek, H.-J. and De Waard, M.A. (2000) 'Fungal Transporters Involved in Efflux of Natural Toxic Compounds and Fungicides', Fungal Genetics and Biology, 30(1), pp. 1-15.

delaGorgendiere, K., Gatulis, J., Brissette, J., Scioletti, J., Bishop, B., Piermarini, P., Dhakal, R. and Yu, N. (2018) 'P022NGS homopolymer detection error resulted A03:01 mistyped as A03:21n', Human immunology, 79, pp. 76-76.

Dooley, H. (2015) 'Fungicide-resistance management tactics: impacts on Zymoseptoria tritici populations', PhD Thesis. University of Reading, UK.

Dooley, H., Shaw, M.W., Mehenni-Ciz, J., Spink, J. and Kildea, S. (2016) 'Detection of Zymoseptoria tritici SDHI-insensitive field isolates carrying the SdhC-H152R and SdhD-R47W substitutions', Pest Manag Sci, 72(12), pp. 2203-2207.

Estep, L.K., Torriani, S.F.F., Zala, M., Anderson, N.P., Flowers, M.D., McDonald, B.A., Mundt, C.C. and Brunner, P.C. (2015) 'Emergence and early evolution of fungicide resistance in North American populations of Zymoseptoria tritici', Plant Pathology, 64(4), pp. 961-971.

FAO (2018) 'The future of food and agriculture - Alternative pathways to 2050.', Rome, 224pp. Licence: CC BY-NC-SA 3.3 IGO.

Fones, H. and Gurr, S. (2015) 'The impact of Septoria tritici Blotch disease on wheat: An EU perspective', Fungal Genetics and Biology, 79, pp. 3-7.

Fraaije, B. (2014a) Use of loop-mediated isothermal amplification assays to detect azole-insensitive CYP51-overexpressing strains of Zymoseptoria tritici.

Fraaije, B.A. (2014b) 'Use of loop-mediated isothermal amplification assays to detect azole-insensitive CYP51-overexpressing strains of Zymoseptoria tritici', Phytopathology, 104(11), pp. 41-42.

Fraaije, B.A., Cools, H.J., Kim, S.H., Motteram, J., Clark, W.S. and Lucas, J.A. (2007) 'A novel substitution I381V in the sterol 14?-demethylase (CYP51) of Mycosphaerella graminicola is differentially selected by azole fungicides', Molecular Plant Pathology, 8(3), pp. 245-254.

Fraaije, B.A., Lucas, J. A., Clark, W. S. and Burnett, F. J. (2003) ' QoI resistance development in populations of cereal pathogens in the UK.', Proceedings BCPC International Congress - Crop Science and Technology. Glasgow,, 10-12 November 2003. pp. 689-694.

Frith, M.C., Wan, R. and Horton, P. (2010) 'Incorporating sequence quality data into alignment improves DNA read mapping', Nucleic acids research, 38(7), pp. e100-e100.

Galili, T., O'Callaghan, A., Sidi, J. and Sievert, C. (2018) 'heatmaply: an R package for creating interactive cluster heatmaps for online publishing', Bioinformatics, 34(9), pp. 1600-1602.

Gladders, P., Paveley, N.D., Barrie, I.A., Hardwick, N.V., Hims, M.J., Langton, S. and Taylor, M.C. (2001) 'Agronomic and meteorological factors affecting the severity of leaf blotch caused by Mycosphaerella graminicola in commercial wheat crops in England', Annals of applied biology, 138(3), pp. 301-311.

Goodwin, S.B. (2012) Resistance in wheat to septoria diseases caused by mycosphaerella graminicola (septoria tritici) and phaeosphaeria (stagonospora) nodorum.

Gosling P, N.C., Watts J (2016) 'HGCA Wheat disease management guide ', 28.

Griffin MJ, F.N. (1985) 'Laboratory studies of benzimidazole resistance in Septoria tritici', EPPO Bulletin, 15, p. 7.

Halberstadt, A.L., Klein, L.M., Chatha, M., Valenzuela, L.B., Stratford, A., Wallach, J., Nichols, D.E. and Brandt, S.D. (2019) 'Pharmacological characterization of the LSD analog N-ethyl-N-cyclopropyl lysergamide (ECPLA)', Psychopharmacology (Berl), 236(2), pp. 799-808.

Hellin, P., Duvivier, M., Clinckemaillie, A., Bataille, C., Legrève, A., Heick, T.M., Jørgensen, L.N., Andersson, B., Samils, B., Rodemann, B., Berg, G. and Kildea, S. (2020) 'Multiplex qPCR assay for simultaneous quantification of CYP51-S524T and SdhC-H152R substitutions in European populations of Zymoseptoria tritici', Plant pathology.

Hennessy, D.A. (2006) 'On Monoculture and the Structure of Crop Rotations', American journal of agricultural economics, 88(4), pp. 900-914.

Holland-Letz, T. and Kopp-Schneider, A. (2015) 'Optimal experimental designs for dose-response studies with continuous endpoints', Arch Toxicol, 89(11), pp. 2059-68.

Hu, M. and Chen, S. (2021) 'Non-target site mechanisms of fungicide resistance in crop pathogens: A review', Microorganisms (Basel), 9(3), pp. 1-19.

Huf, A., Rehfus, A., Lorenz, K.H., Bryson, R., Voegele, R.T. and Stammler, G. (2018) 'Proposal for a new nomenclature for CYP51 haplotypes in Zymoseptoria tritici and analysis of their distribution in Europe', Plant pathology, 67(8), pp. 1706-1712.

J Oksanen, F.G.B., Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R Minchin, RB O'Hara, GL Simpson, P Solymos, MHH Stevens, E Szoecs, H Wagner (2020) 'vegan: Community Ecology Package. R package version 2.5-6. 2019'.

Jess, S., Kildea, S., Moody, A., Rennick, G., Murchie, A.K. and Cooke, L.R. 70 (2014) 'European Union policy on pesticides: implications for agriculture in Ireland'. Chichester, UK, pp. 1646-1654.

Jonathan Blake, S.W., Claude Maumene, and Lise Nistrup Jørgensen (2011) 'Evaluation of the benefits provided by the azole class of compounds in wheat, and the effect of losing all azoles on wheat and potato production in Denmark, France and the UK.', Impact of the loss of all azoles. ADAS, p. 23.

Jørgensen, L.N., Hovmøller, M.S., Hansen, J.G., Lassen, P., Clark, B., Bayles, R., Rodemann, B., Flath, K., Jahn, M., Goral, T., Jerzy Czembor, J., Cheyron, P., Maumene, C., De Pope, C., Ban, R., Nielsen, G.C. and Berg, G. (2014) 'IPM Strategies and Their Dilemmas Including an Introduction to www.eurowheat.org', Journal of Integrative Agriculture, 13(2), pp. 265-281.

Jørgensen, L.N., Matzen, N., Hansen, J.G., Semaskiene, R., Korbas, M., Danielewicz, J., Glazek, M., Maumene, C., Rodemann, B., Weigand, S., Hess, M., Blake, J., Clark, B., Kildea, S., Batailles, C., Ban, R., Havis, N. and Treikale, O. (2018) 'Four azoles' profile in the control of Septoria, yellow rust and brown rust in wheat across Europe', Crop Protection, 105, pp. 16-27.

Jørgensen, L.N., Matzen, N., Heick, T.M., Havis, N., Holdgate, S., Clark, B., Blake, J., Glazek, M., Korbas, M., Danielewicz, J., Maumene, C., Rodemann, B., Weigand, S., Kildea, S., Bataille, C., Brauna-Morževska, E., Gulbis, K., Ban, R., Berg, G., Semaskiene, R. and Stammler, G. (2020a) 'Decreasing azole sensitivity of Z. tritici in Europe contributes to reduced and varying field efficacy', Journal of plant diseases and protection (2006), 128(1), pp. 287-301.

Jørgensen, L.N., Matzen, N., Heick, T.M., Havis, N., Holdgate, S., Clark, B., Blake, J., Glazek, M., Korbas, M., Danielewicz, J., Maumene, C., Rodemann, B., Weigand, S., Kildea, S., Bataille, C., Brauna-Morževska, E., Gulbis, K., Ban, R., Berg, G., Semaskiene, R. and Stammler, G. (2020b) 'Decreasing azole sensitivity of Z. tritici in Europe contributes to reduced and varying field efficacy', Journal of plant diseases and protection (2006).

Jørgensen, L.N., van den Bosch, F., Oliver, R.P., Heick, T.M. and Paveley, N.D. (2017) 'Targeting Fungicide Inputs According to Need', Annual Review of Phytopathology, 55(1), pp. 181-203.

Katoh, K., Kuma, K.-i., Toh, H. and Miyata, T. (2005) 'MAFFT version 5: improvement in accuracy of multiple sequence alignment', Nucleic acids research, 33(2), pp. 511-518.

Kema, G.H.J., Verstappen, E.C.P., Todorova, M. and Waalwijk, C. (1996a) 'Successful crosses and molecular tetrad and progeny analyses demonstrate heterothallism in Mycosphaerella graminicola', Lower Eukaryotes and Organelles, 30(3), pp. 251-258.

Kema, G.H.J., Yu, D., Rijkenberg, F.H.J., Shaw, M.W. and Baayen, R.P. (1996b) 'Histology of the pathogenesis of Mycosphaerella graminicola in wheat', Phytopathology, 86(7), pp. 777-786.

Keon, J., Antoniw, J., Carzaniga, R., Deller, S., Ward, J.L., Baker, J.M., Beale, M.H., Hammond-Kosack, K. and Rudd, J.J. (2007) 'Transcriptional adaptation of Mycosphaerella graminicola to programmed cell death (PCD) of its susceptible wheat host', Molecular plant-microbe interactions, 20(2), pp. 178-193.

Kiiker, R., Juurik, M., Heick, T.M. and Mäe, A. (2021) 'Changes in DMI, SDHI, and QoI Fungicide Sensitivity in the Estonian Zymoseptoria tritici Population between 2019 and 2020', Microorganisms (Basel), 9(4), p. 814.

Kildea, S. (2009) 'Fungicide resistance in wheat pathogen Mycosphaerella graminicola', Belfast: The Queen's University of Belfast.

Kildea, S., Marten-Heick, T., Grant, J., Mehenni-Ciz, J. and Dooley, H. (2019) 'A combination of target-site alterations, overexpression and enhanced efflux activity contribute to reduced azole sensitivity present in the Irish Zymoseptoria tritici population', European Journal of Plant Pathology, 154(3), pp. 529-540.

Kim, Y.M., Park, K., Joo, G.J., Jeong, E.M., Kim, J.E. and Rhee, I.K. (2004) 'Glutathione-dependent biotransformation of the fungicide chlorothalonil', J Agric Food Chem, 52(13), pp. 4192-6.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H. and Phillippy, A.M. (2017) 'Canu: scalable and accurate long-read assembly via adaptive -mer weighting and repeat separation', Genome research, 27(5), p. 722.

Kristoffersen, R., Jørgensen, L.N., Eriksen, L.B., Nielsen, G.C. and Kiær, L.P. (2020) 'Control of Septoria tritici blotch by winter wheat cultivar mixtures: Meta-analysis of 19 years of cultivar trials', Field Crops Research, 249.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms', Molecular biology and evolution, 35(6), pp. 1547-1549.

Laver, T., Harrison, J., O'neill, P.A., Moore, K., Farbos, A., Paszkiewicz, K. and Studholme, D.J. (2015) 'Assessing the performance of the Oxford Nanopore Technologies MinION', Biomolecular Detection and Quantification, 3(C), pp. 1-8.

Leguia, M., Vila-Sanjurjo, A., Chain, P.S.G., Berry, I.M., Jarman, R.G. and Pollett, S. (2020) 'Precision Medicine and Precision Public Health in the Era of Pathogen Next-Generation Sequencing', The Journal of infectious diseases, 221(Supplement_3), pp. S289-S291.

Leroux, P., Albertini, C., Gautier, A., Gredt, M. and Walker, A.-S. (2007) 'Mutations in theCYP51 gene correlated with changes in sensitivity to sterol 14α-demethylation inhibitors in field isolates ofMycosphaerella graminicola', Pest Management Science, 63(7), pp. 688-698.

Leroux P, G.M., Walker AS, Moinard JM, Caron D (2005) 'Resistance of the wheat leaf blotch pathogen Septoria tritici to fungicides in France', Proceedings of the Modern Fungicides and Antifungal compounds, IV, p. 9.

Leroux, P. and Walker, A.-S. (2011) 'Multiple mechanisms account for resistance to sterol 14α-demethylation inhibitors in field isolates of Mycosphaerella graminicola', Pest Management Science, 67(1), pp. 44-59.

Li, H. (2018) 'Minimap2: pairwise alignment for nucleotide sequences', Bioinformatics (Oxford, England), 34(18), p. 3094.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) 'The Sequence Alignment Map format and SAMtools', Bioinformatics, 25(16), pp. 2078-2079.

Loch, J.A., Li, H., Walter, K., Liao, A.K., Chris Pinkard, D., Ivanov, D.V., O'Neill, M.J., Aniebo, I.C., Haudenschild, C.D., Ewan, M.R., Earnshaw, D.J., Ernest Sohna Sohna, J., Spence, E.J., Worsley, G.J., Goddard, C.P., Bailey, D.M.D., Ling Ng, B., Moon, K.W., Carter, R.J., Khrebtukova, I., Mullikin, J.C., Parkinson, M.L., Banerjee, S., Harnish, K., Sizto, N.L., Etchin, S.S., Smith, M.E., Shaw, R.J.,

Novo, S.M., Ning, Z., Diakoumakos, K.D., Alam, M.D., Bridgham, J.A., Sluis, J.P., Mammen, R.M., Lok, M., Egbujor, U.C., Pickering, L., Scott Furey, W., Kumar, A., Mehta, P., Barbour, S.G., Brown, C.G., Rae, V.H., Luo, S., Zuerlein, M., Torrance, P.E., Hoschler, K., Rogers, J., Smith, V.P., Elmore, D.W., Ostadan, O., Sabot, A., Neil Cooley, R., Evers, D.J., Fedurco, M., Heyer, N.I., Karbelashvili, M.S., Black, P.J., Brown, A.A., Romieu, A., Rawlings, S.R., Cox, A.J., Bignell, H.R., Castillo, N., Virk, S.M., McCooke, N.J., Durbin, R., Ellis, D.J., Pliskin, D.P., Ruediger, S.T., Boutell, J.M., vandeVondele, S., Schenker, M.R., Rasolonjatovo, I.M.J., Hims, M.M., Benson, K.F., Yan, J., Seoane, J.M., Walcott, G.C., Irving, L.J., Rogers, J., Granieri, P.A., Hurwitz, S., Milton, J., Smith, G.P., Rogert Bacigalupo, M.C., McCauley, P.G., Sutton, N., Obradovic, B., Tzonev, S.S., Stevens, K., Buermann, D.H., Quijano, V.J., Anastasi, C., Rigatti, R., Benoit, V.A., Crake, N.R., Bundu, A.A. and Scally, A. (2008) 'Accurate whole human genome sequencing using reversible terminator chemistry', Nature (London), 456(7218), pp. 53-59.

Lu, F.F., Liu, J.T., Zhang, N., Chen, Z.J. and Yang, H. (2020) 'OsPAL as a key salicylic acid synthetic component is a critical factor involved in mediation of isoproturon degradation in a paddy crop', Journal of cleaner production, 262, p. 121476.

M. Nordberg, J.D., D. M. Templeton (2004) 'Glossary of terms used in toxicokinetics (IUPAC Reccomendations 2003)', Pure and Applied Chemistry, 76(5), pp. 1033-1082.

Malmberg, M.M., Spangenberg, G.C., Daetwyler, H.D. and Cogan, N.O.I. (2019) 'Assessment of low-coverage nanopore long read sequencing for SNP genotyping in doubled haploid canola (Brassica napus L.)', Scientific reports, 9(1), pp. 8688-12.

Maxam, A.M. and Gilbert, W. (1977) 'A New Method for Sequencing DNA', Proceedings of the National Academy of Sciences - PNAS, 74(2), pp. 560-564.

McDonald, M.C., Renkin, M., Spackman, M., Orchard, B., Croll, D., Solomon, P.S. and Milgate, A. (2019) 'Rapid Parallel Evolution of Azole Fungicide Resistance in Australian Populations of the Wheat Pathogen Zymoseptoria tritici', Applied and environmental microbiology, 85(4).

McDonald, M.C., Williams, A.H., Milgate, A., Pattemore, J.A., Solomon, P.S. and Hane, J.K. (2015) 'Next-generation re-sequencing as a tool for rapid bioinformatic screening of presence and absence of genes and accessory chromosomes across isolates of Zymoseptoria tritici', Fungal Genet Biol, 79, pp. 71-5.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. and Lanfear, R. (2020) 'IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era', Molecular biology and evolution, 37(5), pp. 1530-1534.

Mitra, A., Sarkar, M. and Chatterjee, C. (2017) 'Modulation of Immune Response by Organophosphate Pesticides: Mammals as Potential Model', Proceedings of the Zoological Society, 72(1), pp. 13-24.

Mullins, J.G.L., Parker, J.E., Cools, H.J., Togawa, R.C., Lucas, J.A., Fraaije, B.A., Kelly, D.E., Kelly, S.L. and Tuite, M.F. (2011) 'Molecular Modelling of the

Emergence of Azole Resistance in Mycosphaerella graminicola', PLoS ONE, 6(6).

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N. and Hase, T. (2000) 'Loop-mediated isothermal amplification of DNA', Nucleic acids research, 28(12), pp. E63-63.

O'Donnell, D.R., Parigi, A., Fish, J.A., Dworkin, I. and Wagner, A.P. (2014) 'The roles of standing genetic variation and evolutionary history in determining the evolvability of anti-predator strategies', PloS one, 9(6), pp. e100163-e100163.

Oerke, E.C. (2005) 'Crop losses to pests', The Journal of Agricultural Science, 144(1), pp. 31-43.

Omrane, S., Sghyer, H., Audéon, C., Lanen, C., Duplaix, C., Walker, A.-S. and Fillinger, S. (2015) 'Fungicide efflux and the MgMFS1 transporter contribute to the multidrug resistance phenotype in Zymoseptoria tritici field isolates', Environmental microbiology, 17(8), pp. 2805-2823.

Paulsen, I.T., Brown, M.H. and Skurray, R.A. (1996) 'Proton-dependent multidrug efflux systems', Microbiological reviews, 60(4), pp. 575-608.

Pieczul, K. and Wąsowska, A. (2017) 'The application of next-generation sequencing (NGS) for monitoring of Zymoseptoria tritici QoI resistance', Crop Protection, 92, pp. 143-147.

Pinheiro, C., Bretz, F., Branson M (2006) 'Analysis of Dose-Response Studies-Modeling approaches', Dose finding in drug development, pp. 146-171.

Price, C.L., Parker, J.E., Warrilow, A.G.S., Kelly, D.E. and Kelly, S.L. (2015) 'Azole fungicides - understanding resistance mechanisms in agricultural fungal pathogens', Pest Management Science, 71(8), pp. 1054-1058.

Radhakrishnan, G.V., Cook, N.M., Bueno-Sancho, V., Lewis, C.M., Persoons, A., Mitiku, A.D., Heaton, M., Davey, P.E., Abeyo, B., Alemayehu, Y., Badebo, A., Barnett, M., Bryant, R., Chatelain, J., Chen, X., Dong, S., Henriksson, T., Holdgate, S., Justesen, A.F., Kalous, J., Kang, Z., Laczny, S., Legoff, J.-P., Lesch, D., Richards, T., Randhawa, H.S., Thach, T., Wang, M., Hovmaller, M.S., Hodson, D.P. and Saunders, D.G.O. (2019) 'MARPLE, a point-of-care, strain-level disease diagnostics and surveillance tool for complex fungal pathogens', BMC Biology, 17(1).

Ritz, C., Baty, F., Streibig, J.C. and Gerhard, D. (2015) 'Dose-Response Analysis Using R', PLoS One, 10(12), p. e0146021.

Ritz, C. and Streibig, J.C. (2005) 'Bioassay Analysis using R', Journal of Statistical Software, 12(5).

Robinson, J.T., Thorvaldsdóttir, H., Wenger, A.M., Zehir, A. and Mesirov, J.P. (2017) 'Variant Review with the Integrative Genomics Viewer', Cancer research, 77(21), p. e31.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) 'Integrative genomics viewer', Nature Biotechnology, 29(1), pp. 24-26.

Roohparvar, R., De Waard, M.A., Kema, G.H.J. and 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(5), pp. 378-388.

Rothfels, C.J., Pryer, K.M. and Li, F.W. (2017) 'Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing', New Phytologist, 213(1), pp. 413-429.

Rudd, J.J., Kanyuka, K., Hassani-Pak, K., Derbyshire, M., Andongabo, A., Devonshire, J., Lysenko, A., Saqi, M., Desai, N.M., Powers, S.J., Hooper, J., Ambroso, L., Bharti, A., Farmer, A., Hammond-Kosack, K.E., Dietrich, R.A. and Courbot, M. (2015) 'Transcriptome and Metabolite Profiling of the Infection Cycle ofZymoseptoria triticion Wheat Reveals a Biphasic Interaction with Plant Immunity Involving Differential Pathogen Chromosomal Contributions and a Variation on the Hemibiotrophic Lifestyle Definition', Plant Physiology, 167(3), pp. 1158-1185.

Russell, P.E. (2005) 'A century of fungicide evolution', The Journal of Agricultural Science, 143(1), pp. 11-25.

Samils, B., Andersson, B., Edin, E., Elfstrand, M., Rönneburg, T., Bucur, D., Hutton, F., Heick, T.M., Hellin, P. and Kildea, S. (2021) 'Development of a PacBio Long-Read Sequencing Assay for High Throughput Detection of Fungicide Resistance in Zymoseptoria tritici', Frontiers in microbiology, 12, pp. 692845-692845.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) 'DNA Sequencing with Chain-Terminating Inhibitors', Proceedings of the National Academy of Sciences - PNAS, 74(12), pp. 5463-5467.

Sanssené, J., Selim, S., Roisin-Fichter, C., Djerroud, L., Deweer, C. and Halama, P. (2011) 'Protective and curative efficacy of prothioconazole against isolates of Mycosphaerella graminicola differing in their in vitro sensitivity to DMI fungicides', Pest Management Science, 67(9), pp. 1134-1140.

Scalliet, G., Bowler, J., Luksch, T., Kirchhofer-Allan, L., Steinhauer, D., Ward, K., Niklaus, M., Verras, A., Csukai, M., Daina, A. and Fonné-Pfister, R. (2012) 'Mutagenesis and Functional Studies with Succinate Dehydrogenase Inhibitors in the Wheat Pathogen Mycosphaerella graminicola (Fungal Target Resistance to Carboxamides)', PLoS ONE, 7(4), p. e35429.

Smith, D. (2013) 'Using Fungicides on Wheat', Integrated Pest and Crop Management.

Stammler, G., Carstensen, M., Koch, A., Semar, M., Strobel, D. and Schlehuber, S. (2008) 'Frequency of different CYP51-haplotypes of Mycosphaerella graminicola and their impact on epoxiconazole-sensitivity and -field efficacy', Crop Protection, 27(11), pp. 1448-1456.

Stammler, G. and Semar, M. (2011) 'Sensitivity of Mycosphaerella graminicola (anamorph: Septoria tritici) to DMI fungicides across Europe and impact on field performance', EPPO Bulletin, 41(2), pp. 149-155.

Steven Kildea, L.S., Mladen Cucak, Fiona Hutton (2021) 'Detection of virulence to septoria tritici blotch (STB) resistance conferred by the winter wheat cultivar Cougar in the Irish Zymoseptoria tritici population and potential implications for STB control', Plant Pathology.

Stukenbrock, E.H., Jørgensen, F.G., Zala, M., Hansen, T.T., McDonald, B.A. and Schierup, M.H. (2010) 'Whole-genome and chromosome evolution associated with host adaptation and speciation of the wheat pathogen Mycosphaerella graminicola', PLoS genetics, 6(12), pp. e1001189-e1001189.

Suffert, F., Sache, I. and Lannou, C. 60 (2011) 'Early stages of septoria tritici blotch epidemics of winter wheat: build-up, overseasoning, and release of primary inoculum' Plant Pathol., pp. 166-177.

Sundin, D.R., Bockus, W.W. and Eversmeyer, M.G. (1999) 'Triazole seed treatments suppress spore production by Puccinia recondita, Septoria tritici, and Stagonospora nodorum from wheat leaves', Plant disease, 83(4), pp. 328-332.

Taher, K., Graf, S., Fakhfakh, M.M., Salah, H.B.H., Yahyaoui, A., Rezgui, S., Nasraoui, B. and Stammler, G. (2014) 'Sensitivity of Zymoseptoria tritici Isolates from Tunisia to Pyraclostrobin, Fluxapyroxad, Epoxiconazole, Metconazole, Prochloraz and Tebuconazole', Journal of phytopathology, 162(7-8), pp. 442-448.

Thorvaldsdóttir, H., Robinson, J.T. and Mesirov, J.P. (2013) 'Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration', Briefings in Bioinformatics, 14(2), pp. 178-192.

Tytgat, O., Gansemans, Y., Weymaere, J., Rubben, K., Deforce, D. and Van Nieuwerburgh, F. (2020) 'Nanopore Sequencing of a Forensic STR Multiplex Reveals Loci Suitable for Single-Contributor STR Profiling', Genes, 11(4), p. 381.

Vagndorf, N., Heick, T.M., Justesen, A.F., Andersen, J.R., Jahoor, A., Jørgensen, L.N. and Orabi, J. (2018) 'Population structure and frequency differences of CYP51 mutations in Zymoseptoria tritici populations in the Nordic and Baltic regions', European Journal of Plant Pathology, 152(2), pp. 327-341.

van den Berg, F., van den Bosch, F. and Paveley, N.D. (2013) 'Optimal fungicide application timings for disease control are also an effective anti-resistance strategy: a case study for Zymoseptoria tritici (Mycosphaerella graminicola) on wheat', Phytopathology, 103(12), pp. 1209-19.

van den Bosch, F., Paveley, N., van den Berg, F., Hobbelen, P. and Oliver, R. (2014) 'Mixtures as a Fungicide Resistance Management Tactic', Phytopathology, 104(12), pp. 1264-1273.

Wick, R.R., Judd, L.M. and Holt, K.E. (2019) 'Performance of neural network basecalling tools for Oxford Nanopore sequencing', Genome biology, 20(1), pp. 129-129.

Wieczorek, T., Berg, G., Semaškienė, R., Mehl, A., Sierotzki, H., Stammler, G., Justesen, A. and Jørgensen, L. (2015) 'Impact of DMI and SDHI fungicides on disease control and CYP51 mutations in populations of Zymoseptoria tritici from Northern Europe', European Journal of Plant Pathology, 143(4), pp. 861-871.

Yamashita, M. and Fraaije, B. (2018) 'Non-target site SDHI resistance is present as standing genetic variation in field populations of Zymoseptoria tritici', Pest management science, 74(3), pp. 672-681.

Yin, K. and Qiu, J.-L. (2019) 'Genome editing for plant disease resistance: applications and perspectives', Philosophical transactions. Biological sciences, 374(1767), pp. 20180322-20180322.

Zadoks, J.C., Chang, T.T. and Konzak, C.F. (1974) 'A decimal code for the growth stages of cereals', Weed research, 14(6), pp. 415-421.

Zeng, F., Jiang, R. and Chen, T. (2013) 'PyroHMMsnp: an SNP caller for Ion Torrent and 454 sequencing data', Nucleic acids research, 41(13), pp. e136-e136.

Zhan, J., Mundt, C.C. and McDonald, B.A. (1998) 'Measuring Immigration and Sexual Reproduction in Field Populations of Mycosphaerella graminicola', Phytopathology, 88(12), p. 1330.