Università degli Studi di Torino

Tesi di Dottorato di Ricerca in Scienze Biologiche
e Biotecnologie Applicate

PhD Thesis in Biological Sciences
and Applied Biotechnologies

Study of the molecular interactions in the rice-*Fusarium fujikuroi* pathosystem

Edoardo Piombo

Tutor: Davide Spadaro

XXXII Cycle: 2016 – 2019

Table of contents

1: INTRODUCTION

Bakanae

Bakanae, caused by the hemibiotrophic fungal pathogen *Fusarium fujikuroi* [teleomorph *Gibberella fujikuroi* (Sawada) Ito in Ito & K. Kimura], is a disease of rice originally observed in Japan in 1828 (Ito and Kimora, 1931), but now present in several countries in America, Europe, Asia, Oceania and Africa (Amatulli *et* al., 2010; Carter *et* al., 2008; Chen *et* al., 2014; Desjardins *et* al., 2000; Haq *et* al., 2010; Jeon *et al*., 2013; Karov *et* al., 2005; Khan *et al*, 2000; Kim *et al*., 2015; Zanudin *et al*., 2008). The name means "silly seedling", and it refers to the main symptom of the disease: the elongation and thinning of internodes, causing abnormal height and frail stems, together with thin leaves and grains entirely or partially empty.

The losses caused by bakanae can vary wildly depending on climate, rice cultivars and disease strain. In Thailand losses of 3-15% (Kanjanasoon *et al*., 1965) were registered, while in Macedonia they are between 2% and 20% (Karov *et* al., 2005) and in Japan they go from 20% to 50% (Ito and Kimura, 1931). In Iran losses of even 75% were reported (Saremi *et al*., 2008).

Fungi

Three fungal species have been found to be routinely associated with bakanae of rice: *Fusarium fujikuroi*, *Fusarium proliferatum* and *Fusarium verticilloides* (Amatulli *et* al., 2010; Desjardins *et al*., 2000). However, the major cause of the disease seems to be *F. fujikuroi*. Amatulli *et al* (2010) found that only *F. fujikuroi* strains were able to induce bakanae disease when inoculated on healthy seeds, suggesting the *F. proliferatum* and *F. verticilloides* may have a role as saprophytic or epiphytic. Bashyal *et al*. (2016b), conversely, found that some *F. proliferatum* and *F. verticilloides* isolates can induce the bakanae disease on rice, but with a lower severity. In a separate study (Bashyal *et al*., 2016a), the authors collected 126 *Fusarium* strains from bakanae-infected plants and verified that 99,2% of them

belonged to the species *F. fujikuroi*.

Symptomatology

Most *F. fujikuroi* strains induce symptoms as elongation and thinning of internodes, abnormal height, frail stems, thin leaves and grains entirely or partially empty. However, the phenotype of affected plants can vary wildly depending on the pathogen strain involved. There are reports of *F. fujikuroi* isolates inducing stunting, chlorosys, yellowing and root and crown rots. Rotting of seedlings prior to emergence is also common (Amoah *et al.*, 1995; Karov *et al.*, 2009, Gupta *et al*., 2015). Depending on the induction of either elongation or stunting, Niehaus *et al*. (2017) suggested a division of the species in two pathotypes according to genetic reasons: strains causing stunting have 1 more PKS (PKS51), whose deletion cause an even more severe stunting symtomatology, while they lack the gene Nrps31, necessary for the synthesis of apicidin F. Additionally, Niehaus *et al*. (2017) found out that elongation-inducing strains have a much higher expression of the gibberellin gene cluster, while stunting inducing ones are characterized by a greater production of fusaric acid and fumonisins, mycotoxins that facilitate virulence even if they are not necessary for it.

Disease cycle

An excellent overview of the rice-*F. fujikuroi* pathosystem has been given by Gupta and colleagues (2015). Being a seedborne pathogen, *F. fujikuroi* is mainly spread through infested seeds, surviving on them as hyphae for up to 4–10 months at room temperature and more than three years at 7°C in cold storage (Kanjansoon, 1965). Ascospores are produced by diseased plants during flowering and reach other plants by the wind. Conidia, on the contrary, are produced on diseased and dead plants and commonly contaminate seeds during harvesting, while another source of infection is the presence of spores and mycelium in the water used for soaking seeds to

stimulate germination (Karov *et al*., 2009).

Both ascospores and conidia infect seedlings through the roots and crown (Sun, 1975), allowing the fungus to invade the plant without producing visible symptoms, so that *F. fujikuroi* can be found in apparently healthy seeds.

Soilborne infection is possible as well, though *F. fujikuroi* can survive in infested soils only for a limited time (Kanjanasoon, 1965), with only 0.7% infection occurring 90 days after soil inoculation, and no infection resulting at 180 days since inoculation.

Routes of infection

A recent study (Sunani *et al*., 2019) showed how *F. fujikuroi* usually enters rice roots by penetration with infection hyphae, infection cushions and globose appressoria, colonising both the intracellular and intercellular spaces. The infection of the embryo can happen through 3 main routes: systemic infection by vascular system, stylar canal and growth of hyphae from outside the husk to the ovary, meaning that the pathogen can be both internally and externally seedborne, and the eradication of both internal and external inoculum are therefore necessary for the best management of the disease (Sunani *et al*., 2019).

Secondary metabolism

Gibberellins

The most studied secondary metabolite of *F. fujikuroi* is gibberellic acid, or gibberellin A3. Gibberellins are a class of plant hormones involved in the regulation of bolting, germination, elongation of roots and internodes (Hedden and Sponsel, 2015) by inducing relaxation of cell wall (Cosgrove and Sovonick-Dunford, 1989). They also have complex effects on flowering, with a promotive, neutral or inhibitory function depending on the studied plant species (Hedden and Sponsel, 2015; Pharis and King 1985; Zeevaart 1976). The history of the study of these phytohormones is closely related to the one of *F. fujikuroi*: the production of GA3 is the main cause

for the bakanae-like symptoms in *F. fujikuroi* (Niehaus *et al*., 2017), and gibberellins were initially considered to be just pathogenicity-related fungal metabolites, with the scientific community only later realizing their role as plant hormones. *F. fujikuroi* produces gibberellins in a different way from plants, with a cluster of 7 genes, not 9 (Bömke and Tudzynski, 2009), with one single gene (*cps/ks*) performing the functions of both ent-copalyl diphosphate synthase and ent-kaurene synthase. Moreover, 3β-hydroxylation is among the first steps of the synthesis in the fungus but the last step in plants, and the contrary is true for 13-hydroxylation (Hedden and Sponsel, 2015). Furthermore, fungal GA biosynthesis branches off from the general terpenoid biosynthetic pathway at the stage of farnesyl diphosphate (FDP), starting from isopentenyl diphosphate (IPP) produced via the mevalonic acid pathway (MVA pathway) (Bömke and Tudzynski, 2009), while in most plants the IPP necessary for biosynthesis is provided predominantly by the plastid-based methyl erythritol phosphate (MEP) pathway (Eisenreich et al., 2001; Kasahara et al., 2002; Sponsel, 2001).

Expression of the plant gibberellin gene cluster, as well as the fungal one, is important for virulence in *F. fujikuroi*: Matić and colleagues (2016) observed how the pathogen is able to induce overexpression of the plant cluster during the infection in the susceptible rice cultivar Galileo, but not on the resistant Selenio. However, the amount of gibberellins necessary for the induction of the symptoms is small: Michielse and colleagues (2014) verified that the silencing of the global regulator FfSge1 diminished gibberellin presence in inoculated plants to 55% of what was caused by the Wild Type isolates, but this reduced amount was still enough to cause full elongation and virulence.

Even if other *Fusarium spp*. contain all or most of the gibberellin biosynthetic cluster (Malonek *et al*., 2005a), only *F. fujikuroi* and a few strains of *F. proliferatum* are actually able to produce this class of molecules (Malonek *et al*., 2005b; Tsavkelova *et al*., 2008).

Fumonisins

Fumonisins are neurotoxic mycotoxins produced by *F. fujikuroi* (Desjardins *et al*., 1997; Wulff *et al*., 2010). The most studied fumonisin, FB1, is known to cause equine leucoencephalomalacia and porcine pulmonary oedema (Scott, 2012), and it has been associated with human oesophageal cancer (Chu and Li, 1994; Sydenham *et al.*, 1990) as well as both kidney and liver cancer in mice (Creppy, 2002). Most of the studies on fumonisin production by fungi were carried out on *Fusarium verticillioides* or *F. proliferatum*, where the gene cluster responsible for their production is composed of 17 genes, even if only 4 of them are absolutely necessary for the synthesis (FUM1, FUM6, FUM8, FUM21), and the knock-out of 4 of the others (FUM15-18) has no observable effect on fumonisin production (Alexander *et al*., 2009). The gene cluster is almost identical in *F. fujikuroi*, except for the absence of FUM20 (unknown function) and the presence of FUM17 (ceramide synthase) as a pseudogene (Wiemann *et al*., 2013). Despite the presence of a functioning cluster, elongation-inducing *F. fujikuroi* strains present a very low expression of fumonisin biosynthetic genes, while these mycotoxins are preferentially produced by stunting-inducing strains, in which the expression of the fumonisin gene cluster is a factor in virulence (Niehaus *et al*., 2017).

Fusaric acid

Fusaric acid is a myxotocin (Bacon *et al*., 1996), but it also has an acanthamoebicidal (Boonman et al. 2012), antioomycete (Son et al. 2008), and antimycobacterial activity (Pan et al. 2011). Furthermore, it is toxic for several plants, such as cotton and banana. It is produced by several Fusarium species (Bacon *et al*., 1996), including *F. oxysporum*, *F. verticilloides* and *F. proliferatum*, thanks to a cluster of 12 genes (Brown *et* al., 2015; Studt *et al*., 2016).

The role of fusaric acid in *F. fujikuroi* pathogenicity is similar to the one of fumonisins: elongation-inducing *F. fujikuroi* strains produce this mycotoxin in very low quantities, if any, while stunting-inducing ones present an higher biosynthesis of this metabolite, whose presence enhances virulence (Niehaus *et al*., 2017).

Others

F. fujikuroi is also known for producing a vast array of other secondary metabolites, including fusarin C (Barrero *et* al., 1991), fusarubins (Studt *et al*., 2012), moniliformin (Desjardins *et al*., 1997), beauvericin (Moretti *et al*., 1996), neurosporaxanthin (Avalos *et al*., 2012), apicidin F (Niehaus *et al*., 2014) and bikaverin (Balan *et al.*, 1970).

Fusarin C is a mycotoxin firstly identified in *F. verticilloides* by Wiebe and Bjeldanes (1981) and it has a mycoestrogenic and cytotoxic activity (Sondergaard *et al*., 2011). It also shows strain-dependant production in *F. proliferatum* (Niehaus *et al*., 2016). Fusarubins are pigments responsible for the coloration of *F. fujikuroi* perithecia (Studt *et al*., 2012), but specific types of fusarubins can have phytotoxic (Gopalakrishnan *et al*., 2005) and anticarcinogenic (Adorisio *et al*., 2019) activity. They are produced by *F. fujikuroi, F. verticilloides, F. mangiferae* and some strains of *F. proliferatum* (Niehaus *et al.,* 2016). Moniliformin is a mycotoxin proven to induce reductions of feed consumption and body weight gain, cardiotoxicity and/or alterations in serum biochemical analytes and hematologic values in several farm animal species (Fremy *et al*., 2019). It is produced by *F. fujikuroi*, *F. proliferatum*, *F. verticilloides* and *F. lactis* (Yang *et al*., 2011). Beauvericin is a carcinogenic mycotoxin (Taevernier et al., 2016), also produced by the entomopathogenic fungus *Beauveria bassiana* (Xu et al., 2008), and by certain strains of *F. proliferatum* as well (Liuzzi *et al.,* 2017). Neurosporaxanthin, produced by *F. fujikuroi*, *F. oxysporum* (Avalos *et al*., 2017), *F. verticilloides* (Ádám *et al*., 2011) and *F. graminearum* (Cambaza, 2018), is a carotenoid orange pigment (Avalos *et al*.,

2012).

The red pigment bikaverin is a vacuolation factor (Cornforth *et al*., 1971) with an antibiotic activity against several organisms, including the protozoon *Leishmania brasiliensis* (Balan *et al*., 1970), the oomycete *Phytophtora infestans* (*Son et al*., 2008) and the nematode *Bursaphelenchus xylophilus* (Kwon *et al.*, 2007). Moreover, it could have an application in cancer therapy due to its inhibition effect on human protein kinase CK2 (Haidar *et al*., 2019) and the cytotoxic activity on several types of cancer cell lines (Fuska *et al.*, 1975; Zhan *et al*., 2007). It can be produced by many *Fusarium spp.*, including *F. proliferatum, F. verticilloides*, *F. oxysporum* and *F. mangiferae* (Niehaus *et al.,* 2016).

Apicidin F, whose cluster has so far been found only in *F. fujikuroi* strains, is a cyclic tetrapeptide commonly only produced by *F. fujikuroi* at high nitrogen levels, and its biosynthetic gene cluster seem to be present only in elongation-inducing strains (Niehaus *et al*., 2017). Conversely, an unknown metabolite, whose base structure is formed by the protein PKS51, is normally present only in stunting-inducing strains (Niehaus *et al*., 2017), suggesting a genetic, phenotypic and metabolic separation between two populations of *F. fujikuroi* able to induce different symptoms.

Breeding resistant varieties

Numerous studies have investigated why some rice cultivars, such as the Italian Selenio and the Korean Nampyeong, are resistant to bakanae disease, identifying a number of QTLs in the genome of rice. Two genomic regions associated to bakanae resistance (qBK1_628091 and qBK4_31750955) were detected by Volante and colleagues (2017) with a genome wide association study. Ji and colleagues (2018) identified a QTL (qFfR1) containing at 95% probability 15 genes with putative functions related to plant defence, while another major QTL, qFfR9, was discovered at 30.1 centimorgan (cM) on rice chromosome 9 by Kang and colleagues (2019).

Lee and colleagues (2019) found out that the QTL qBK1 was able to confer resistance to the cultivar Shingwang, while QTL $qBK1^{WD}$ was responsible of 20.2% of the resistance in the japonica variety Wonseadaesoo (Lee *et al*., 2018). In addition, 13 novel QTLs were recently identified by Chen and colleagues (2019). Resistance-related genes have also been investigated by transcriptomics experiment, looking for genes upregulated in resistant cultivars after inoculation with *F. fujikuroi* (Ji *et al.*, 2016; Matić *et al*., 2016).

Overall, the efforts of the scientific community have brought to the identification of many putative resistance genes and genomic regions useful for the development of resistant cultivars, but a clear resistance mechanism has yet to be elucidated.

Control

The disease is mainly seedborne. In a study of Kanjanasoon (1965) it was proven that, even if artificially inoculated soil caused 93% infection immediately after inoculation, waiting 90 days before planting was sufficient to reduce the infection to 0.7%, with no infection occurring at 180 days since inoculation. Accordingly, only a small percentage of bakanae occurrences is believed to be derived from a soilborne infection.

For these reasons, the main form of control for bakanae is constituted by the chemical treatment of seeds. In Italy, procloraz was the most utilized product against bakanae up until 2015, when it was forbidden and replaced by fludioxonil. However, chemical treatments can induce the development of resistance in pathogens, or be harmful to beneficial microorganisms, and for this reasons academia and industry are looking for ways of treating seeds that do not involve pesticides (Munkvold, 2009). Matić and colleagues (2014) were able to show how *Pichia guilliermondii* R9 and *Metchnikowia pulcherrima* R23, combined with thermotherapy at 60 °C for 10 min, decreased the bakanae disease index below 5% while at the same time improving the seed germination rate, possibly because of a seed priming effect. The

incidence of bakanae on untreated seeds in the same study was 93%, while the disease index after application of procloraz was 4.3%.

Since an efficient screening of the seeds can dramatically reduce the incidence of the disease, several LAMP tests have recently been developed in order to check for the presence of *F. fujikuroi* in seed lots (Ortega *et al*., 2018; Sunani *et* al., 2019; Zhang *et al*., 2019).

The genome

F. fujikuroi is phylogenetically very close to *F. verticilloides* and *F. proliferatum*, despite having a shortened chromosome IV (Wiemann *et al*., 2013).

The presence of chromosome XII is strain specific in both *F. verticilloides* (Xu and Leslie, 1996) and *F. fujikuroi*, even if it seems to be always present in *F. proliferatum* (Niehaus *et al*., 2016) and this chromosome does not seem necessary for pathogenicity on rice (Wiemann *et al.*, 2013).

Only 5% of the genome is made of regions with more than 55% AT, with AT-rich regions being mostly limited to the telomeres (Wiemann *et al*., 2013). It cointains 1125 putative transcription factors, similarly to *F. verticilloides* (1136) and *F. proliferatum* (1116), but less than *F. oxysporum* (1644). The number of secreted proteins (1336) is similar to *F. verticilloides* and *F. proliferatum* as well, and 15% less than *F. oxysporum* (Niehaus *et* al., 2016; Wiemann *et al*., 2013). Transposable elements (TE) amount to 4.18% of the genome, while in *F. verticilloides* they are only the 1.07% of the total and in *F. proliferatum* they are 0.8%. This high amount of TE is, however, much less prominent than in *F. oxysporum*, where they can constiute 19.25% of the total genome (Niehaus *et al*., 2016).

Including both characterized and putative ones, *F. fujikuroi* has 45 secondary metabolite gene clusters, whose backbone genes are constituted of 17 type 1 poliketide synthases, 1 type 3 poliketide synthase, 15 non ribosomal peptide synthases, 2 dimethylallyl tryptophan synthases, 2 diterpene cyclases and 8

sesquiterpene cyclases.

Only 2.9% of *F. fujikuroi* genes do not have orthologs in other *Fusarium spp*. This percentage is similar in *F. proliferatum* (2.2%), but much higher in *F. verticilloides* (12.6%).

Description of the work

The second chapter of the thesis describes the comparison of three *F. fujikuroi* strains inducing different symptoms on rice, with the objective of studying the effect of temperature on bakanae disease and to link genomic polymorphisms with the different phenotypes shown by the plant hosts.

The third chapter details the sequencing of one avirulent strain and one lowly virulent one, which are then compared to a total of 12 virulent strains. The objective of the work is to identify genes putatively involved in *F. fujikuroi* pathogenicity.

The fourth chapter contains the work done with a *F. fujikuroi* isolate transformed with GFP (Green Fluorescent Protein), focusing on differences in *F. fujikuroi* infection of rootlets in rice cultivars either susceptible or resistant to bakanae. The efforts of the scientific community have been focused on the response of the adult plant, and little is known of the early rice-*F. fujikuroi* interactions.

The fifth chapter summarises and discusses the main conclusions of the thesis.

At the end of the thesis there are two appendices unrelated to the rice-*F. fujikuroi* pathosystem, which were included because they had an important role in the research activity performed during the PhD. Appendix A describes the sequencing and annotation of the first *Metschnikowia fructicola* genome, while Appendix B is about the study of griseofulvin production regulation in *Penicillium griseofulvum*, using knock-out mutants of key genes.

2: DIFFERENT PHENOTYPES, SIMILAR GENOMES: THREE NEWLY SEQUENCED *FUSARIUM FUJIKUROI* STRAINS INDUCE DIFFERENT SYMPTOMS IN RICE DEPENDING ON TEMPERATURE.

Edoardo Piombo[1,2], Pietro Bosio[1], Alberto Acquadro[1], Pamela Abbruscato[3], Davide Spadaro[1,2]

[1]Department of Agricultural, Forestry and Food Sciences (DISAFA), University of Torino, Grugliasco (TO), 10095, Italy.
[2]Centre of Competence for the Innovation in the Agroenvironmental Sector (AGROINNOVA), University of Torino, Grugliasco (TO), 10095, Italy.
[3]Parco Tecnologico Padano, Lodi (LO), 26900, Italy

Corresponding author: Davide Spadaro; E-mail: davide.spadaro@unito.it

Funding: This project has received funding from the European Union Horizon 2020 research and innovation program under grant agreement No. 634179 (EMPHASIS – "Effective Management of Pests and Harmful Alien Species – Integrated Solutions") and from AGER Foundation under Grant no. 2010–2369 ("RISINNOVA - Integrated genetic and genomic approaches for new Italian rice breeding strategies").

INTRODUCTION

Bakanae, caused by the hemibiotrophic fungal pathogen *Fusarium fujikuroi* [teleomorph *Gibberella fujikuroi* (Sawada) Ito in Ito & K. Kimura], is one of the most important diseases of rice (Carter *et al.*, 2008; Desjardins *et al.*, 1997). Crop losses due to bakanae are largely depending on climate and rice cultivars, varying from 3% to 75% in certain cases (Saremi *et al.*, 2008).

Originally observed in Japan in 1828, bakanae disease is now present in several countries in America, Europe, Asia and Africa. Researchers have used disparate approaches to investigate how rice plants can resist to the disease, from RNA sequencing to QTL mapping (Fiyaz *et al.*, 2016; Ji *et al.*, 2018; Matić *et al.* 2016), but, despite this, to date no rice cultivar showing a complete resistance to bakanae has been developed, and there is still a limited knowledge of the mechanisms of resistance (Bagga and Kumar, 2000; Desjardins *et al.*, 2000).

The pathogen commonly induces symptoms like abnormal height, thin leaves and grains entirely or partially empty, mainly due to the production of gibberellins (Niehaus *et al.*, 2017). *F. fujikuroi* is also able to increase the production of these phytohormones by the plant, with less susceptible cultivars showing less gibberellin production, and a reduced expression of their biosynthetic gene cluster, compared to highly susceptible cultivars (Kim *et al.*, 2018; Matić *et al.*, 2016; Siciliano *et al.*, 2015). Despite this, there are also reports of strains inducing stunted and chlorotic seedlings (Gupta *et al.*, 2015), often followed root and crown rots (Amoah *et al.*, 1995; Karov *et al.*, 2009). Due to these differences, *F. fujikuroi* strains have been recently divided in two pathotypes (Niehaus *et al.*, 2017).

Besides the ability to induce bakanae disease, some strains of *F. fujikuroi* are also known for the production of fumonisins, neurotoxic mycotoxins (Desjardins *et al.*, 1997; Wulff *et al.,* 2010). The most studied fumonisin, FB1, is known to cause equine leucoencephalomalacia and porcine pulmonary edema (Scott, 2012), and it has been associated with human esophageal cancer (Chu and Li, 1994; Sydenham

et al., 1990) and kidney and liver cancer in mouse (Creppy, 2002).

Fusarium fujikuroi has been sequenced for the first time in 2013 (Jeong *et al.*, 2013; Wiemann *et al.*, 2013), but many more strains have become available in recent years (Bashyal *et al.*, 2017; Chiara *et al.,* 2015; Niehaus *et al.*, 2017; Radwan *et al.*, 2018; Urbaniak *et al.*, 2018). This fungus has a genome of around 45 Mb, divided among 12 chromosomes, with repetitive elements constituting less than 1% of the total. The gene content on average varies between 13,000 and 15,000 genes, including around 1,200 genes encoding for secreted proteins.

Beside gibberellins, the fungus is able to produce a wide array of secondary metabolites, including both mycotoxins, such as fumonisins, fusaric acid, and fusarins (Bacon *et al.*, 1996; Barrero *et al.*, 1991; Desjardins *et al.*, 1997), and pigments, like bikaverin and fusarubins (Balan *et al*, 1970; Studt *et al.*, 2012). Forty-seven putative gene clusters for secondary metabolites were found in the reference genome of *F fujikuroi* (Wiemann *et al.*, 2013), and a number of these have been characterized in recent years (Janevska and Tudzynscki, 2018). A number of global and local regulators control the production of secondary metabolites, but many are also able to regulate gibberellin production, and therefore pathogenicity. These include the global nitrogen regulators *area* (Tudzynski *et al.*, 1999) and *areb* (Pfannmüller *et al.*, 2017) and the component of the velvet complex *lae1* (Niehaus *et al.*, 2018).

F. fujikuroi shows complete synteny in the fumonisin cluster with *F. verticillioides* and *F. oxysporum* (Wiemann *et al.*, 2013), despite a reduced production of these molecules (Stępień *et al.*, 2011; Wulff *et al.*, 2010) when compared with the above-mentioned species. However, in *F. fujikuroi* the ability to produce fumonisins, and the quantity produced, can vary significantly depending on the genotype and the environment (Matić *et al.*, 2013).

Temperature is one of the most important factors influencing both the virulence of *F. fujikuroi* strains and the production of fumonisins, but, notwithstanding, there are

few works investigating its effect on the rice-*F. fujikuroi* pathosystem (Saremi and Farrokhi, 2004; Matić *et al.*, 2017).

This work aims to use a combination of high-throughput sequencing, comparative genomics, chemical analyses and molecular biology to investigate the effect of temperature on virulence and secondary metabolite production in three *F. fujikuroi* strains showing different phenotype.

The considered strains are Augusto2, CSV1 and I1.3, all isolated from infected rice plants in northern Italy (Amatulli *et al.*, 2010).

MATERIALS AND METHODS

Microorganisms and seeds

The strains of *F. fujikuroi* named Augusto2, CSV1 and I1.3, previously isolated from diseased rice plants in Piedmont (Amatulli *et al.*, 2010) and maintained in the Agroinnova microorganism collection, were grown on sterile PDB for 10 days at 23°C. Afterwards, the suspensions were filtered through sterile gauze, centrifuged for 20 min at 6,000 RPM and resuspended in Ringer solution. The Burker chamber was then used to obtain concentrations of 10^5 conidia/ml. Rice seeds 'Galileo', susceptible to bakanae disease (Amatulli *et al.*, 2010) were thermally treated by dipping in water at 60°C for 5 min, immersed in a solution of 1% NaClO for 2 min and then washed three times with sterile water for 5 min. The seeds were then divided and immersed in the conidial suspension of the different strains and kept in agitation for 30 min.

Pathogenicity trials

After drying for 24 h on sterile paper, the seeds were sown in sterilized substrates (70% white peat and 30% clay, with pH between 5.5 and 6). The N content was between 110 mg/l and 190 mg/l, P_2O_5 was of 140-230 mg/l and K_2O was 170-280 mg/l. The plants were grown in two growth chambers: one was kept at 22°C and the

other at 31°C. Disease symptoms were monitored weekly starting one week post germination (wpg). A disease index was attributed, depending on the visible symptoms: 0: healthy plant; 1: reduced dimension, chlorotic leaves; 2: internode elongation, significant yellowing, significant dwarfism; 3: necrosis of the crown; 4: dead or not-germinated plant. Each strain was tested on 4 replicates of 30 plants. Four replicates of 30 uninoculated plants were used as control. The experiment was performed twice.

RNA extraction and qPCR

RNA was extracted from the basal half of the shoot of plants inoculated with each of the strains, as well as from control plants, by using the RNeasy kit (Qiagen, Hilden, Germany). The extracted RNA was quantified by Nanodrop (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and purified using the TURBO DNA-free kit (Ambion, Foster City, California, United States). The samples were then checked for DNA contamination by PCR. The gene used was the rice elongation factor 1-alpha. After verifying the sample purity, the RNA was used to obtain the cDNA, using the iScript cDNA synthesis kit (Biorad, Hercules, California, United States). The samples were then used in real time qPCR (Applied Biosystems StepOnePlus, Foster City, California, United States), with primers for *fum1* (fumonisin gene cluster polyketide synthase, *F. fujikuroi*), *fum21* (fumonisin gene cluster transcription factor, *F. fujikuroi*), *cps/ks* (gibberellin gene cluster ent-copalyl diphosphate synthase ent-kaurene synthase, *F. fujikuroi*), and *gib20ox1* (Gibberellin 20 oxidase 1, rice). The PCR mix were composed of 5 µl of Applied Biosystems SYBR Green Power Mix, 2 µl of cDNA, 0.15 µl of each primer (10 µM) and 2.4 µl of nuclease free water. The thermal cycler protocol was the following: 95°C for 10 min, followed by 40 cycles (95°C for 15 s; 60°C for 60 s) and 95°C for 15 s. The ubiquitin *F. fujikuroi* gene (Wiemann *et al.*, 2013) and the rice elongation factor 1-alpha (Manosalva *et al.*, 2009) were used as housekeeping genes,

respectively for fungal and plant genes. The sequences of the primers used are reported in **supplementary table 2.1**.
The efficiency of the primers used for *fum21* and *gib20ox1* amplification was tested with a standard curve built upon five serial dilutions (1:10).

***In vitro* assays**

Every strain was grown in PDB flasks (30 ml of medium) and YES Agar plates. The flasks and the plates were inoculated with 100 µl of a solution containing 10^5 conidia/ml, prepared following the same procedure used for the pathogenicity trials. The plates were kept at 24°C, with a 12:12 h light/dark photoperiod, a light intensity of 1 cd and a relative luminosity of 55 cd. During the fungal growth in YES Agar, the mycelial diameter was measured, and the color and texture were monitored.

Chemical extractions

Samples obtained by PDB flasks were filtered to separate the mycelium from the growth medium. Mycelium was weighed (500mg) and extracted with 1ml of methanol:water (8:2 v/v), during 1 hour in ultrasonic bath. Supernatant was centrifuged and filtered by 0.45 µm filters, after which it was placed in the vials for HPLC analysis. Regarding YES Agar plates, the extraction was carried out on the whole plates with 3ml of methanol. The solvent was spread on the whole surface and the mycelium was scratched and brought to suspension. The extract was then placed in tubes and concentrated with a Concentrator 5301 (Hamburg, Germany). The dried residue was dissolved in methanol:water (1:1 v/v) and placed in vials for HPLC analysis. Similarly to the procedure used for mycelia, 500 mg of *in vivo* sample were extracted with 1 ml of methanol:water (8:2 v/v) by ultrasonic bath for 1 hour. Supernatant was centrifuged and filtered with 0,45 µm filters, after which it

was placed in vials.

HPLC-MS/MS

Liquid chromatography was performed with Varian Model 212-LC micro pumps (Hansen Way, CA, USA) coupled with a Varian 126 autosampler Model 410 Prostar. A Synergi 4u Fusion-RP 80A (100 mm × 2.0 mm, Phenomenex, Castel Maggiore, Italy) analytical column was used coupled with Fusion-RP (4 × 2.0 mm) security guard for LC separation. The chromatographic conditions were: column temperature at 45 °C; mobile phase consisting of eluent A (HCOOH 0,05% in H2O) and eluent B (CH3CN). A gradient elution was applied as follows: 0 to 20% of B in 5 minutes, from 20% to 80% of B in 15 minutes, from 80% to 100% of B in 1 minute. Five minutes of post run were necessary for column conditioning before the subsequent injection. The injection volume was 20 µl, and the flow speed was of flow of 200 µl/min.

The triple quadrupole mass spectrometer (Varian 310-MS) was operated in the negative/positive electrospray ionization mode (ESI^-/ESI^+). To select the MS/MS parameters for the analysis of metabolites by multiple reaction monitoring (MRM). For the quantification of fumonisin B4 the calibration curve of fumonisin B1 was used, since fumonisin B4 currently lacks a specific commercial standard. Two transitions were selected for each compound: GA3: 345>214 (Collision energy (CE) 14 eV), 345>143 (CE 30 eV); FB1: 722>334 (CE 38 eV), 722>352 (CE 34 eV); FB2/FB3: 706>336 (CE 36 eV), 706>354 (CE 34 eV); FB4: 690>338 (CE 30 eV), 690>320 (CE 30 eV). The collision gas (Ar) pressure was set at 2 mbar for all of the experiments.

Sequencing, assembly and analysis

The *F. fujikuroi* strains Augusto2, CSV1 and I1.3 were sequenced by Parco Tecnologico Padano using a next generation Illumina MiSeq sequencer. For each

strain, a paired end library was generated using the Nextera XT DNA preparation kit (Illumina, San Diego, California, United States). For strain I1.3, a mate-pair library was also generated using the Nextera Mate Pair kit (Illumina, San Diego, California, United States), following the protocols provided by the manufacturer. Libraries were purified by AMPure XP beads and normalized to ensure equal library representation in the pools. Equal volumes of libraries were diluted in the hybridization buffer, heat denatured and sequenced. Standard phi X control library (Illumina) was spiked into the denatured HCT 116 library. The libraries and phi X mixture were finally loaded into a MiSeq 250 and MiSeq 300-Cycle v2 Reagent Kit (Illumina). Base calling was performed using the Illumina pipeline software. Demultiplexing was done using an Illumina provided software. Trimming of adapters and removal of ambiguous bases was done using Trimgalore (*https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/*), and the resulting cleaned reads were checked with fastqc (*https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*) for remaining contamination. For the I1.3 reads, the program "Scythe" (*https://github.com/vsbuffalo/scythe*) was also used to remove remaining adapters. Initially, *de novo* assembly was performed, using SPAdes version 3.7.1 (Bankevich *et al.*, 2012), and the obtained assembly was used in a reference guided approach with IMR-DENOM (*http://mtweb.cs.ucl.ac.uk/mus/www/19genomes/IMR-DENOM/*), since the low sequencing coverage of Augusto2 and CSV1 made it impossible to obtain a good purely *de novo* assembly (**Supplementary table 2.2**). The selected mapper used in IMR-DENOM was bwa (Li and Durbin, 2009).

Gene prediction

Gene prediction was conducted using the version 2.31.8 of MAKER (Cantarel *et al.*, 2008). Both predictors augustus v.2.5.5 (Stanke and Waack, 2003) and SNAP v.2006-07-28 (*http://korflab.ucdavis.edu/software.html*) were used. augustus used

the "--fusarium" option for gene prediction, while SNAP was trained to obtain a file.hmm specific for the three genomes. The necessary repeat libraries were constructed using the basic procedure (*http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/Repeat_Library_Construction--Basic*). The external data provided to MAKER, and used for the training of SNAP, were all the EST, protein sequences and transcript sequences of *F. fujikuroi* available on NCBI. To launch MAKER, the option "-fix_nucleotides" was used, in order to allow the program to work with degenerate nucleotides present in the external data. The option "correct_est_fusion" was also activated in the control files. After the analysis, introns shorter than 10 bp, predicted by SNAP, were removed, and, when this caused a frameshift mutation, the prediction of the gene splicing sites was repeated with augustus v.2.5.5 (Stanke and Waack, 2003).

SNP mining

The clean paired end reads of each of the three strains of interest were mapped on the reference genome of *F. fujikuroi* strain IMI 58289 (Wiemann *et al.*, 2013), using bwa v.0.7.12-r1039 (http://bio-bwa.sourceforge.net) with default options. The resulting sam files were converted to sort.bam by samtools v.0.1.19-96b5f2294a (*http://samtools.sourceforge.net/*), and they were used for SNP mining with the following pipeline:

'samtools mpileup -guf reference.fa augusto2.sort.bam CSV1.sort.bam I1.3.sort.bam | bcftools view -cg - | vcfutils.pl varFilter -D 200 -Q 20 - > file.vcf'

Heterozygous SNPs were assumed to be derived from sequencing errors and were excluded from the analysis. The SNPs were mapped on the genome with the program CircosVCF (Drori *et al.*, 2017). The reads of the strains were also mapped, using the same pipeline, on the scaffold 005 of the *F. fujikuroi* strain B14 (Genbank: FMSL01000005.1), and Tablet (Milne *et al.*, 2013) was used to check if any reads

mapped on the gene FFB14_06372.

Analysis of polymorphisms

The SnpEff program v. 4.2 (Cingolani *et al.,* 2012) was used to evaluate the impact of the SNPs/indels identified with the SNP mining, after building a database for IMI 58289 following the manual instructions (http://snpeff.sourceforge.net/SnpEff_manual.html#databases). Afterwards, we checked if the strains Augusto2, CSV1 and I1.3 presented missense or nonsense polymorphisms in the gibberellin and fumonisin gene clusters, or in other genes involved in the biosynthesis regulation of these metabolites (see **supplementary table 2.3** for references). The presence of these polymorphisms was then checked in the sort.bam files with the viewer Tablet (Milne *et al.*, 2013). EffectorP 1.0 and 2.0 (Sperschneider *et al.*, 2016) were used on the secreted portion of the *F. fujikuroi* proteome (Wiemann *et al.*, 2013) to predict putative effector genes, and these genes were also checked for polymorphisms. The impact of polymorphisms of interest was predicted with Provean Protein (Choi and Chan, 2015). All the genes presenting putatively MODERATE and HIGH impact polymorphisms in either Augusto2 or CSV1, but not in both, were identified, according to the evaluation of SNPeff. These genes were annotated with BLAST2GO with default parameters, and, when they presented GO terms related to regulation of transcription, pathogenesis or metabolism, the impact of their polymorphisms was predicted with Provean Protein (Choi and Chan, 2015).

Structural variant analysis

The software BreakDancer v1.3.6 (Fan *et al*., 2014) was used to identify structural variants in the genome. Variants with a score lower than 80 were removed, and an original python script (**supplementary file 1**) was used to identify genes localized in the regions affected by the remaining variations. Genes present in an area

involved in a deletion are considered to be affected, as are genes that have the edge of an inversion or a translocation inside their sequence. The script only works on variations involving only one scaffold, and therefore structural variants affecting different chromosomes were checked manually.

Phylogenetic analysis

OrthoFinder v. 2.3.3 (Emms and Kelly, 2015) was used with the option "-M msa" to obtain a genome-wise phylogenetic tree based on single-copy genes, comparing the *F. fujikuroi* strains Augusto2, CSV1 and I1.3 to several other annotated isolates of the same species. The strains used for this analysis were: B20 (GenBank: GCA_900096605.1), C1995 (GenBank: GCA_900096645.1), E282 (GenBank: GCA_900096705.1), FGSC_8932 (GenBank: GCA_001023045.1), FSU48 (GenBank: GCA_900096685.1), IMI58289 (GenBank: GCA_900079805.1), KSU3368 (GenBank: GCA_001023065.1), KSU X-10626 (GenBank: GCA_001023035.1), m567 (GenBank: GCA_900096615.1), MRC2276 (GenBank: GCA_900096635.1) and NCIM1100 (Genbank: GCA_900096625.1), with *Fusarium oxysporum* f. sp. *lycopersici* 4287 used as outgroup (GenBank: GCA_000149955.2). STAG (Emms and Kelly, 2018) was used to generate an unrooted species tree, and the root was placed with MEGA (Kumar *et al*., 1994) between *F. oxypsorum* and the *F. fujikuroi* strains.

Comparing Augusto2 and CSV1

Proteinortho v. 5.16 (Lechner *et al.*, 2011) was used to identify genes present either in CSV1 or in Augusto2, but not in both. Following this, genes unique to CSV1 were blasted against the genome of Augusto2, and vice versa. Genes with a good blast hit were excluded from the analysis, since they could be actually present in both strains, their absence in one derived by an error from the gene predictor.

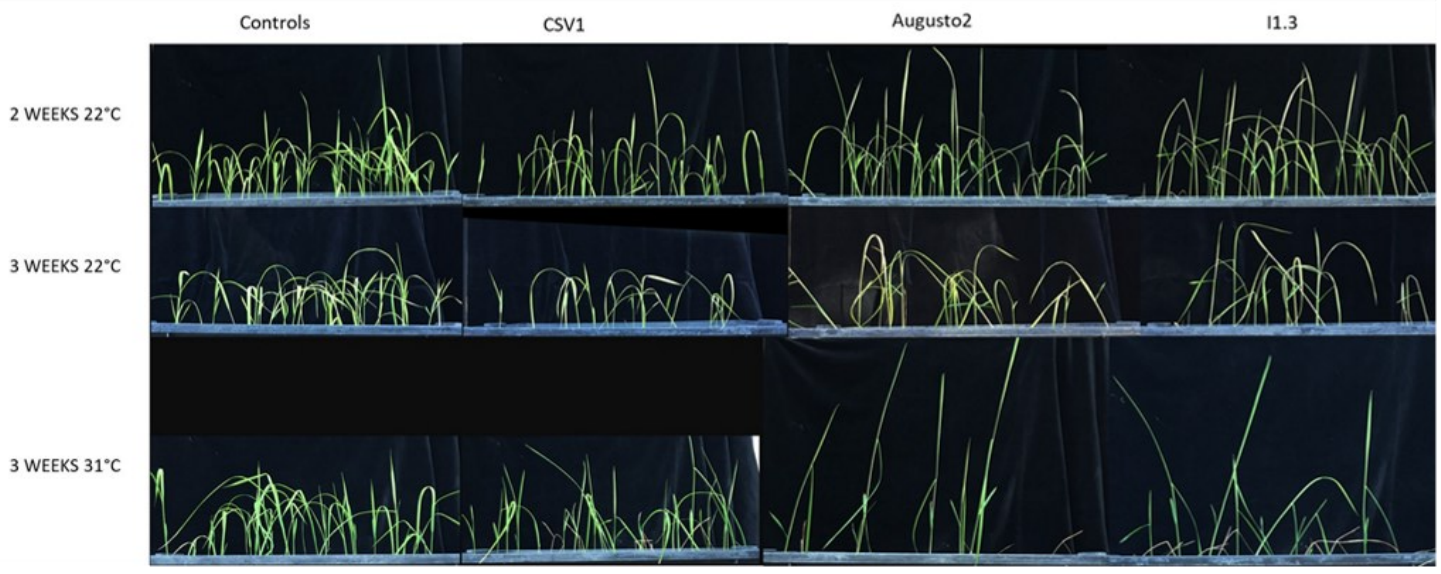

Figure 2.1: Rice plants (cv. Galileo) inoculated with *F. fujikuroi* strains CSV1, Augusto2 or I1.3.

RESULTS

Pathogenicity trials

Plants inoculated with the strains Augusto2, CSV1 and I1.3 presented widely different symptoms (**figure 2.1**). At 22°C and 2 wpg, the symptoms of all the strains were mixed between those associated to the two pathotypes identified by Niehaus *et al*. (2017) some plants showed elongation, while others were stunted. At 3 wpg, CSV1 induced stunting and withering, while I1.3 tended to induce more elongation and plants inoculated with Augusto2 could present both types of symptoms. Disease indexes are reported in **figure 2.2**. While all the strains showed a similar virulence at 22°C, at 31°C Augusto2 and I1.3 were much more virulent, and nearly all the plants died at 2 wpg, with the remaining ones showing extreme elongation.

In vitro trials

The three strains on YES Agar produced mycelia of different color, dimension and texture (**supplementary figure 2.1**). CSV1 and I1.3 mycelia are characterized by a reverse red-orange color, not present in Augusto2. This color is also present in the front view of the CSV1 mycelia. Growth speed was not uniform as well (**supplementary figure 2.2**). CSV1 grew faster at the beginning of the trial, reaching a diameter of around 40 mm at 5 days after the inoculation. Afterwards, Augusto2 and I1.3 started growing faster, reaching, after 14 days of inoculation, average diameters of 80 and 65 mm, against the 50 mm of CSV1.

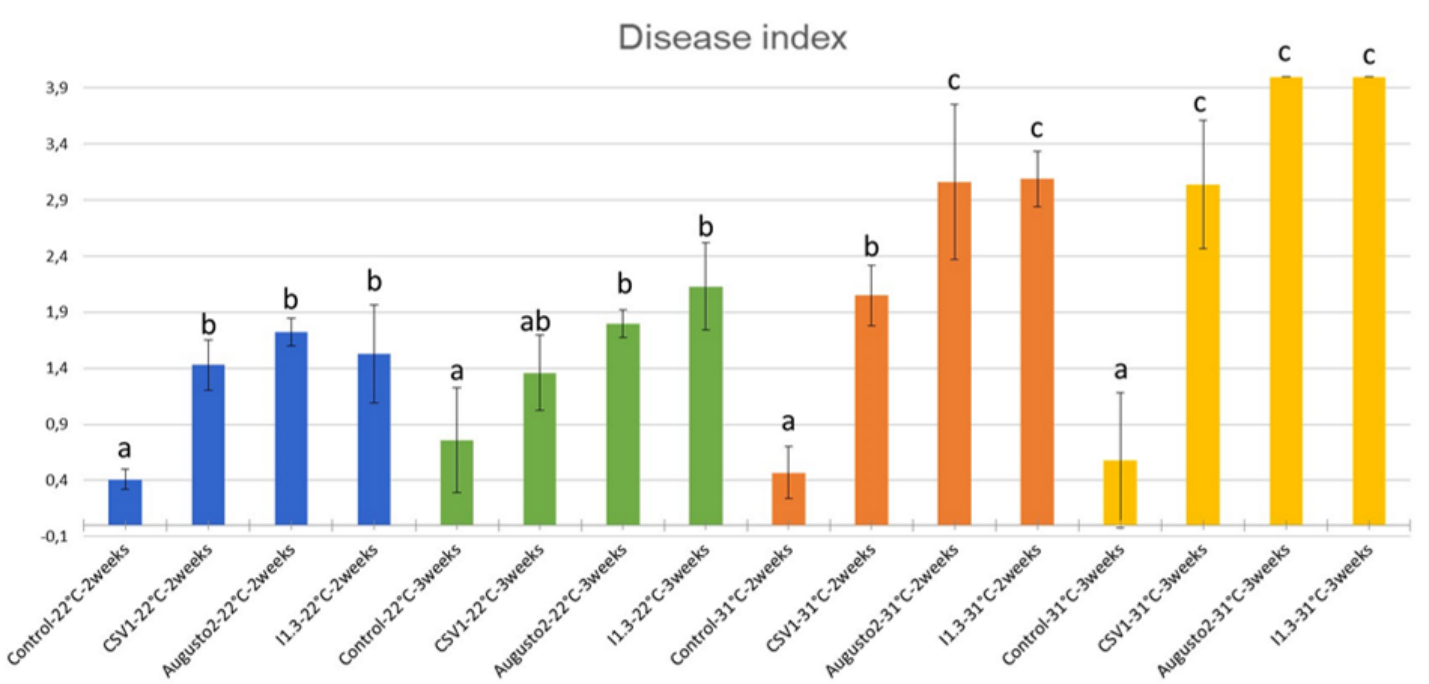

Figure 2.2: Disease indexes of rice plants (cv. Galileo) inoculated with the 3 studied strains of *F. fujikuroi*. Values with the same letter are not statistically different according to the Duncan's test ($P<0.05$).

Chemical analyses

The results of the *in vivo* quantification of GA3 are presented in **figure 2.3**. At 22°C and 2 wpg, plants inoculated with Augusto2 contained slightly more GA3 than those inoculated with the other strains. However, one week later the highest amount of GA3 was found in I1.3-inoculated plants. At 31°C and 2 wpg, the highest quantity of GA3 was found in plants inoculated with Augusto2 and I1.3, while plants affected by strain CSV1 had a GA3 quantity slightly but not significantly higher than control plants. At 3 wpg, the quantity of GA3 in CSV1-inoculated plants decreased still, probably because many plants died. Fumonisins were not present *in vivo* at a detectable level in any of the plant samples.

Besides, GA3 and fumonisin production by the three strains were tested *in vitro* **(supplementary table 2.4)**. CSV1 did not produce GA3 at detectable levels *in vitro* on YES Agar, but it was the highest producer on PDB (10,676.7 ppb). Fumonisins were produced *in vitro* by Augusto2 (181,052 ppb of FB1 on average), and much less by CSV1 (22 ppb of FB1), while I1.3 did not produce these mycotoxins in any situation.

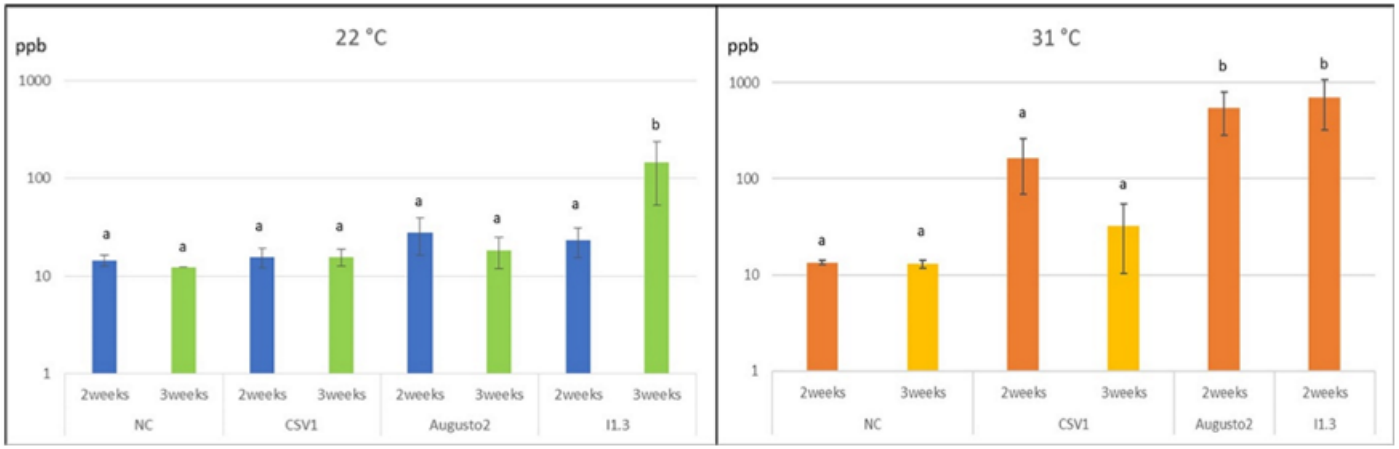

Figure 2.3: GA3 quantity in rice plants (cv. Galileo) inoculated with the 3 studied strains of *F. fujikuroi*. Analysis done at 2 or 3 weeks since germination, at 22°C and 31°C. The error bars represent the standard deviation. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$). This test was executed in an independent manner for the samples at 22 °C and those at 31 °C.

Real time RT-PCRs

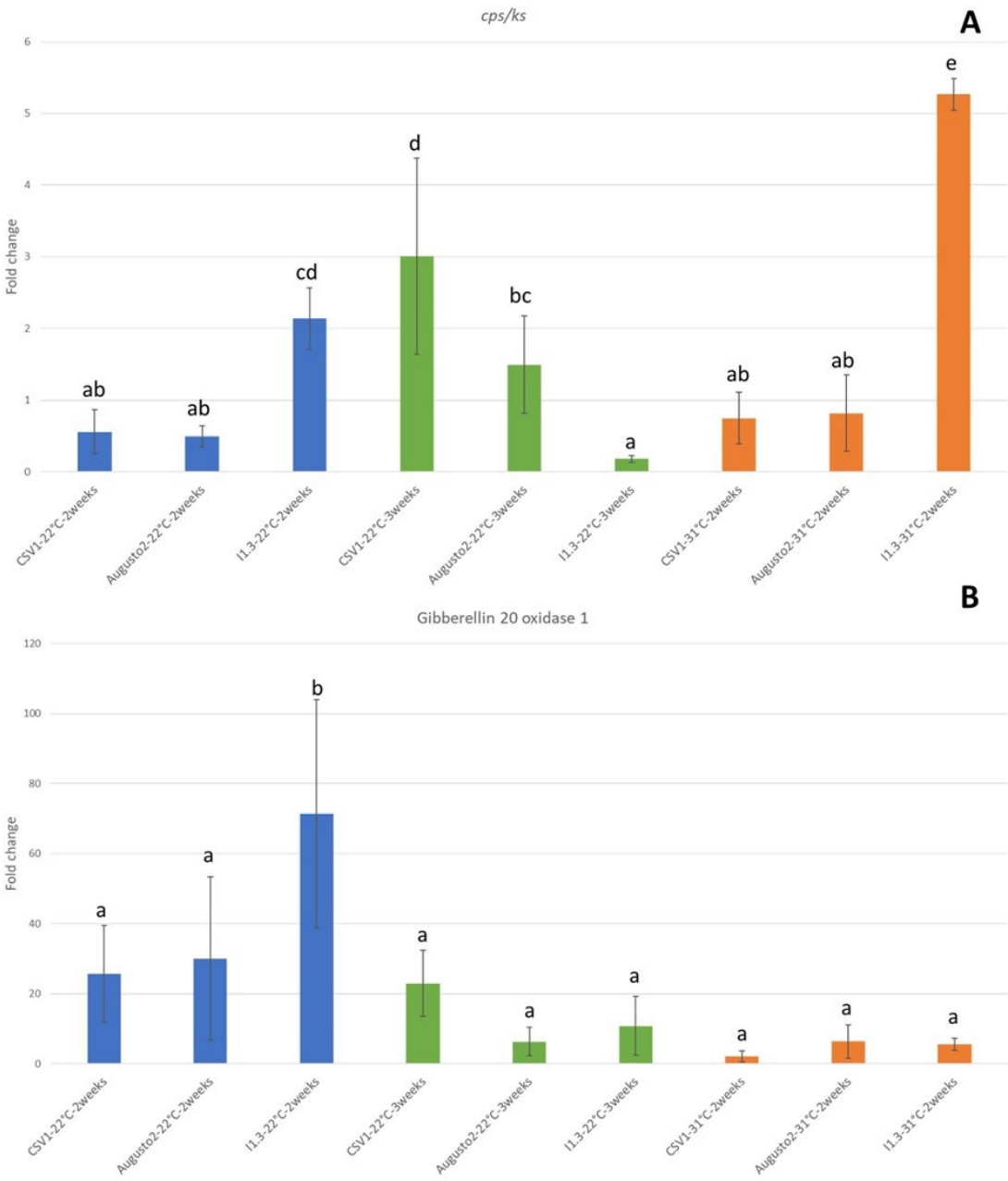

Figure 2.4: Gene expression of *cps/ks*, a gene of *F. fujikuroi* gibberellin cluster (A), and of gibberellin 20 oxidase 1, a gene of rice gibberellin cluster (B). Data obtained by reverse transcriptase real time PCR. The error bar is the standard deviation. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

The gene expression of *cps/ks* and Gibberellin 20 oxidase 1 at various time points is presented in **figure 2.4**. At 2 wpg, *cps/ks*, a key gene of the fungal gibberellin gene cluster, was mainly expressed in I1.3-inoculated plants, both at 22°C and 31°C. On the contrary, at 22°C and 3 wpg, this gene was mostly expressed in strain CSV1. In Augusto2, the level of expression did not change significantly in the three examined conditions.

Regarding the gene Gibberellin 20 oxidase 1, belonging to the plant gibberellin gene cluster, the expression was higher in strain I1.3 at 22°C and 2 wpg, with CSV1 and Augusto2 showing similar expression. At 31°C, the RT-PCR results were similar in every strain, showing low expression, while at 22°C and 3 wpg, CSV1 induced the highest level of expression, followed by I1.3 and Augusto2.

Neither *fum1* nor *fum21* expression was detected in plants at any time point, confirming the results obtained by the chemical analyses.

Sequencing, assembly and bioinformatic analysis

The results of the genome sequencing are presented in **table 2.1**. MiSeq Illumina sequencing produced respectively 3.76, 3.9 and 2.12 millions of raw paired ends reads for Augusto2, CSV1 and I1.3. For I1.3, 16.57 million reads of mate pair reads were also obtained.

The estimated coverage, based on the 43.65 Mb length of the reference genome of strain IMI 58289 (Wiemann *et al.*, 2013), was 9.68X, 10.16X and 62.34X, for Augusto2, CSV1 and I1.3, respectively. Starting from these data, the reference guided approach with IMR/DENOM allowed to reconstruct the 12 chromosomes of the three *F. fujikuroi* strains. Using MAKER, it was possible to predict 13563, 13578 and 13690 proteins for Augusto2, CSV1 and I1.3. The assemblies and their annotations were deposited in GenBank: I1.3 (Accession number: CP023101 -

CP023112); Augusto2 (CP023089 - CP023100); CSV1 (CP023077 - CP023088). The genes present either in Augusto2 or in CSV1, but not in both, are listed in **supplementary table 2.5**.

Table 2.1: Data regarding the genome reference-guided assembly and annotation of strains Augusto2, CSV1 and I1.3 of *Fusarium fujikuroi*.

	Augusto2	CSV1	I1.3
Genome Size	~43.7 Mb	~43.7 Mb	~45.6Mb
Sequencing coverage	9.7X	10.2X	62.3X
Number of contigs	12	12	12
Number of Large contigs (>100 Kb)	12	12	12
N50 (base pairs)	4,218,434	4,212,448	4,426,414
GC content	47.49%	47.51%	47.2%
Number of genes	13,563	13,578	13,690
Annotated genes	10,073	10,080	9,838

SNP mining

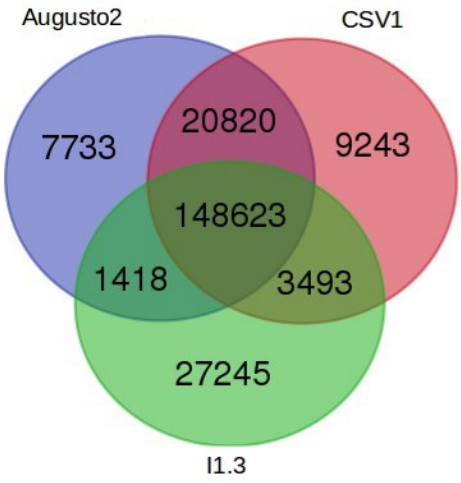

Figure 2.5: Venn graphic showing the distribution of polymorphisms among the *F. fujikuroi* strains Augusto2, CSV1 and I1.3. The genome of strain IMI 58289 was used as reference in the SNP calling. Image obtained with the software at the following link: *http://bioinformatics.psb.ugent.be/webtools/Venn/*

Compared to the reference genome of *F. fujikuroi* strain IMI 58289, 178,594, 182,179 and 180,779 SNPs/indels were found in Augusto2, CSV1 and I1.3, respectively (**supplementary file 2.2**). The distribution of these polymorphisms in the three strains is shown in **figure 2.5**, and their position on the reference genome is presented in **figure 2.6**. The vast majority of the polymorphisms (over 80%; 148,623 SNPs/indels) are shared by the three strains. The differences between the analysed strains and the reference strain were evaluated with SNPeff (Cingolani *et*

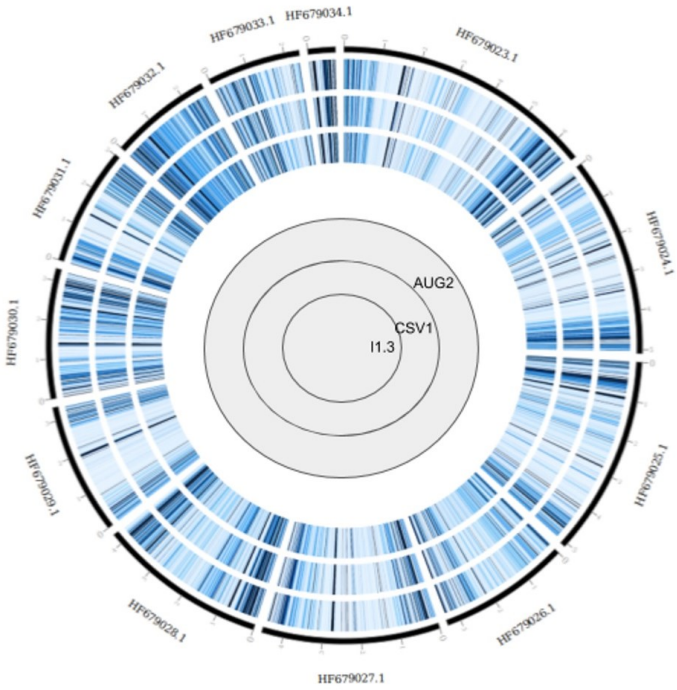

Figure 2.6: The figure shows the localization of polymorphisms in the *F. fujikuroi* strains Augusto2, CSV1 and I1.3 on the reference genome of strain IMI 58289. The external ring shows the polymorphisms of strain Augusto2, the central one shows those of CSV1 and the internal one shows the ones of I1.3. Image obtained with CircosVCF (Drori *et al.*, 2017).

al., 2012), and the results of this analysis are presented in **table 2.2**. None of the strains presented a unique polymorphism in the gibberellin gene cluster, not even at intergenic level (data not shown). With "unique polymorphism", a mutation not common to all the three strains is meant. In the fumonisin gene clusters, on the other hand, there were a number of polymorphisms upstream and downstream the genes, together with various unique missense polymorphisms. In particular, in *I1.3* strain, there were 4 missense polymorphisms in the transcription factor *fum21* and 2 in the

polyketide synthase *fum1*. One missense polymorphism in each of these two genes was also present in the strains CSV1 and Augusto2.

Table 2.2: Number and putative effect of polymorphisms detected in the strains CSV1, Augusto2 and I1.3 of *Fusarium fujikuroi*. The reference used for the SNP calling was the genome of *F. fujikuroi* strain IMI 58289.

Strain	**CSV1**	**Augusto2**	**I1.3**
Number of polymorphisms	**182,179**	**178,594**	**180,779**
SNPs	176,34	172,722	174,933
Insertions	3,061	3,08	3,082
Deletions	2,778	2,778	2,764
Variant rate	1/241 bases	1/245 bases	1/242 bases
Predicted polymorphism effect			
Silent	41,476	41,457	41,375
Missense	27,404	27,385	27,356
Nonsense	376	373	373
Polymorphism impact			
High	973	973	949
Moderate	27,589	27,565	27,551
Low	43,318	43,301	43,228
Negligible	714,739	708,481	712,012

Regarding the fusaric acid gene cluster, there were some intron and intergenic polymorphisms, but no missense or nonsense polymorphisms. Unique missense and nonsense SNPs in the regulators were also searched (**supplementary table 2.3)**, and 1 missense SNP in the sequence of the global regulator *vea* was identified in the strains Augusto2 and I1.3. All these polymorphisms were analysed with PROVEAN PROTEIN (Choi and Chan, 2015), and two SNPs observed in the strain I1.3, one in the polyketide synthase *fum1* and one in the transcription factor *fum21* were predicted to have a deleterious effect on the function of their protein. CSV1 and Augusto2, despite their differences in the phenotype, had most polymorphisms in common. Only 138 reference genes have a missense, nonsense or frameshift polymorphism not common to both CSV1 and Augusto2. Of this subset, 34 genes had some GO terms related to pathogenicity, metabolism or regulation of transcription, and only eight had stop, frameshift or missense mutations predicted to be deleterious by PROVEAN Protein (**supplementary table 2.5**). By mapping the reads on the scaffold 005 of *F. fujikuroi* strain 005 (Genbank: FMSL01000005.1), the gene FFB14_06372, encoding PKS51, a protein involved in causing stunting and withering in hosts, was not covered in reads in any of the strains, suggesting its absence in the analysed genomes.

Structural variant analysis

The results of breakdancer are presented in **supplementary file 2.3.** 107 deletions, 21 inversions and 3 intra-chromosomal translocations were identified, putatively affecting the function of 66 genes in at least one of the strains.

Phylogenetic analysis

The analysis with OrthoFinder identified 14,699 orthogroups among the considered proteomes. In the resulting phylogenetic tree, *F. fujikuroi* strains Augusto2 and CSV1 seem to be closer to each other than to I1.3, although it must be noted that the

support values of the tree tend to be low (**supplementary figure 2.3**).

Effector prediction and analysis

A list of 323 putative effectors was obtained by running EffectorP 1.0 and 2.0 on the *F. fujikuroi* secretome (Wiemann *et al.*, 2013) and cross-referencing results (Sperschneider *et al.*, 2016). Two of the identified genes had missense polymorphisms present in CSV1 but not in Augusto2: FFUJ_01956 and FFUJ_11601. Analysis with PROVEAN Protein, however, predicted that these two polymorphisms did not have an effect on the protein function.

DISCUSSION

Temperature effect

This work investigates the effect of temperature in the rice-*F. fujikuroi* pathosystem: three newly sequenced *F. fujikuroi* strains induce different symptoms in rice depending on the temperature. Pathogenicity trials were conducted with three strains of different virulence, and in every case, the pathogen was much more virulent at 31°C than at 22°C. In fact, at 31°C the strains I1.3 and Augusto2 were so virulent that most of the plants were dead after 2 wpg.

Niehaus *et al.* (2017) showed how there are at least two pathotypes of *F. fujikuroi*, one associated with bakanae-like symptoms and gibberellin production, the other inducing withering and stunting. The pathotypes are thought to be diverse from a phylogenetic, symptomatic and metabolomic point of view. However, in the current study, strains phylogenetically close to each other were capable of inducing both types of symptoms. At 22°C and 2 wpg, the symptoms of all the strains are mixed, with some plants showing elongation and some stunting (**figure 2.1**). However, at this time point, I1.3 showed a high expression of both *cps/ks* and Gibberellin 20 Oxidase 1 (**figure 2.4**), belonging respectively to the fungal and plant gibberellin gene clusters, and this induced one week later “bakanae-like” symptoms in all the

plants. Conversely, CSV1 and Augusto2 did not express strongly the gibberellin gene clusters, and therefore the induced symptoms were mostly dwarfism, with no elongation in CSV1 and slight elongation in Augusto2 (**figure 2.1**). This is corroborated by the HPLC-MS analysis: plants inoculated with strain I1.3 contained a higher concentration of GA3 at 3 wpg, while one week before the quantities were similar for every strain (**figure 2.3**). On the other hand, at 31°C, the surviving plants inoculated with Augusto2 and I1.3 showed elongation (**figure 2.1**), and they had a very high content in GA3 (**figure 2.3**), while CSV1 mostly induced stunting, and contained less GA3. In addition, the expression of *cps/ks* and Gibberellin 20 oxidase 1 was low in CSV1, and one week later the GA3 level was even less. The very low number of surviving plants did not permit to perform analysis at 31°C and 3 wpg for Augusto2 and I1.3, but they both showed a low expression of Gibberellin 20 oxidase 1 at 2 wpg, though a significant expression of *cps/ks* was measured in I1.3. Even at 2 wpg, most of the plants were dead, so the significantly greater expression of *cps/ks* in I1.3 at 2 wpg is due to the survival of few plants which showed a high expression level.

In conclusion, it seems that, despite their proximity from a phylogenetical point of view (**figures 2.5** and **2.7**), I1.3 induces a "bakanae-like" phenotype at all temperatures, and CSV1 is characterized by low GA3 production and stunting, while Augusto2 is actually capable of changing the induced symptoms depending on the temperature, being closer to I1.3 at 31°C and a mix of both phenotypes at 22°C.

The gene encoding PKS51, associated with the *F. fujikuroi* pathotype causing stunting and withering, was not present in the three examined strains.

Fumonisin production

Fumonisins are mycotoxins whose consumptions produces a vast array of effects on animals, including nephrotoxicity and hepatotoxicity (Bolger *et al*., 2001), as well as neurotoxicity and cardiotoxicity (Scott, 2012). Fumonisins or fumonisin

transcripts were not detected *in vivo*, neither with HPLC-MS nor with real time PCRs, but this was expected, given the fact that this pathogen produces minimal amounts of these metabolites (Wiemann *et al.*, 2013). However, fumonisins were detected *in vitro* for strains Augusto2 and CSV1. I1.3 did not produce fumonisins at a detectable level neither *in vivo* nor *in vitro*, likely as an effect of the putatively important polymorphisms that this strain has in the transcription factor *fum21* and the polyketide synthase *fum1*, since both genes are essential for the correct functioning of the gene cluster (Alexander *et al.*, 2009).

A study of Cruz *et al*. (2013) found no relationship in *F. fujikuroi* between pathogenicity and the ability to produce fumonisins (Cruz *et al*., 2013), while Niehaus *et al*. (2017) observed that the deletion of the fumonisin PKS caused a reduction in virulence, but only in stunting-inducing strains. Our data correlate well with these studies: none of our strains produced fumonisins in detectable quantity *in vivo* and, while this has probably a negligible effect on the virulence of I1.3 and Augusto2, it may impact the virulence of CSV1, which induced stunting at both the tested temperatures and presented a lower virulence than the other two isolates.

In the work of Matić *et al.* (2013) the fumonisin synthesis of the same three strains was analysed, with similar results: Augusto2 produced by far the highest quantity of these mycotoxins, followed by CSV1. Interestingly, in the conditions tested in that work, strain I1.3 was able to produce a small amount of fumonisin B1.

Different phenotypes, similar genomes

The three sequenced *F. fujikuroi* strains were isolated from the same geographic area, but their phenotype was very different. CSV1 and Augusto2 are particularly close from an evolutionary point of view (**supplementary figure 2.3**), with only 21,887 SNPs between them, but they differ in virulence, reaction to temperature, induced symptoms, colony morphology and color, growth speed, fumonisin and gibberellin production. Given the low sequencing coverage used, the amount of

SNPs was probably underestimated, but the high percentage of shared polymorphisms (93% of the total for CSV1 and 95% for Augusto2) is a further proof of the low evolutionary distance between the two strains.

The most common genes involved in pathogenesis and gibberellin production were checked for polymorphisms, but no SNPs that could explain these variations were found. Even if Augusto2 and I1.3 had a missense SNP in *vea*, a regulator of secondary metabolism associated to fumonisin and fusarin production, there is currently no evidence linking fusarins to the development of the disease. Fumonisin production is believed to have no relationship with pathogenicity of elongation-inducing *F. fujikuroi* as well (Cruz *et al*., 2013; Niehaus *et al*., 2017), and none of the considered strains produced fumonisins at a detectable level *in vivo*. Two putative effectors presented missense polymorphisms in CSV1 and not in the other strains, but a prediction analysis with PROVEAN Protein showed that it is unlikely for these differences to have an impact on the protein function.

The differences between the genomes of Augusto2 and CSV1 were further investigated by checking missense, frameshift or nonsense SNPs present in either CSV1 or Augusto2, but not both. The genes presenting these polymorphisms were filtered by checking for GO terms related to metabolism, pathogenicity or gene regulation, and the missense SNPs were evaluated with PROVEAN Protein, discarding those with a putative neutral effect. The remaining genes are listed in **supplementary table 2.5**, which contains also the genes putatively affected by a structural variant either in Augusto2 or in CSV1, but not in both. An other source of phenotype variation could be the absence or presence of certain genes in the genomes, though only 14 genes were present either in Augusto2 or CSV1, but not in both (**supplementary table 2.5**).

The genes with predicted function-affecting polymorphisms do not seem directly correlated to the observed differences in the phenotype, and neither do the genes present in only one of the genomes. However, the protein CCT62922.1, a pisatin

demethylase, was putatively affected by an inversion in CSV1, and this class of proteins is known to be a factor of virulence in both *F. oxysporum* and *F. solani* (Rocha *et al*., 2015; Wasmann and VanEtten, 1996). Conversely, CCT63174.1, an endo polygalacturonase, a virulence factor in *F. graminearum* (Paccanaro *et al*., 2017), was removed by a deletion in Augusto2 and I1.3, but not in CSV1. Another protein putatively not functioning in Augusto2 was CCT73390.1, an integral membrane protein, and some proteins of this class are factors of virulence for plant pathogens, such as integral membrane protein PTH11, which is required for pathogenicity and appressorium formation in *Magnaporthe grisea* and it exhibits host-preferential expression in *F. graminearum* (DeZwaan *et al*., 1999; Harris *et al*., 2016). Finally, CCT74990.1, related to a fructosyl amino acid oxidase, was predicted to be affected by an inversion in CSV1, but this protein was proven to be dispensable for development and growth in *Aspergillus nidulans*, whose null mutant for this gene grew normally and developed as many conidia and sexual structures as the wild-type (Jeong *et al*., 2002).

Besides these proteins, a number of the genes presented in **supplementary table 2.5** are currently uncharacterized, and their activity could contribute to the differences observed between Augusto2 and CSV1

CONCLUSIONS

This study presents a comparative genomics analysis of three *F. fujikuroi* strains isolated in northern Italy, the largest European production area of rice. The strains showed remarkable difference in the phenotype, despite being very close from an evolutionary point of view, suggesting that a few key mutations in a small number of genes can dramatically alter the phenotype induced by the pathogen. A few candidate genes that may explain these phenotypic differences were identified.

The species *F. fujikuroi* was recently divided in two phylogenetically separated pathotypes (Niehaus *et al.*, 2017), which induce respectively bakanae symptoms or

stunting and withering. However, in this study it has been observed that minimal genetic differences can induce symptom modifications, and some strains may be able to induce both types of phenotypes, depending on environmental factors such as temperature.

Finally, it was observed that the considered *F. fujikuroi* strains became much more virulent at higher temperatures. This observation could be linked to the effect of occurring climatic changes. The rise of average temperatures in spring may affect rice production not only with increasing losses induced by abiotic stresses, but also with the average increase of virulence of *F. fujikuroi*. While the danger posed to rice by climate change favoring abiotic stresses is known (Mohammed and Tarpley, 2009), and efforts are underway to obtain climate-resilient cultivars (Sreenivasulu *et al.*, 2015), there is little knowledge over the impact of increased temperatures on the interactions between rice and fungal pathogens.

ACKNOWLEDGMENTS

The Authors thank Dr. Ilenia Siciliano for performing part of the chemical analyses and Dr. Houda Banani for measuring the *in vitro* growth of the microorganisms.

SUPPLEMENTARIES

Supplementary figure 2.1: Front and reverse view of *F. fujikuroi* strains Augusto2, CSV1 and I1.3, growing on YES Agar plates at different time points. The inoculation on the plates was done with 100 µl of a suspension of 5*10^5 conidia/ml

Supplementary figure 2.2: The graphic shows the increasing in diameter of colonies of *F. fujikuroi* strains Augusto2, CSV1 and I1.3, growing on YES Agar plates. The inoculation on the plates was done with 100 µl of a suspension of 5*10^5 conidia/ml.

Supplementary figure 2.3: The tree describes the phylogeny of the strains Augusto2, CSV1 and I1.3 of *F. fujikuroi*, in relation to other strains of the same

species. *Fusarium oxysporum* f. sp. *lycopersici* 4287 was used as outgroup (GenBank: GCA_000149955.2). The tree was obtained by using the programs OrthoFinder 2.3.3 (Emms and Kelly, 2015) and STAG (Emms and Kelly, 2018). The root was placed with MEGA (Kumar *et al*., 1994) between *F. oxypsorum* and the *F. fujikuroi* strains.

Supplementary table 2.1. Primers used for the reverse transcription real time PCRs.

Supplementary table 2.2. Data regarding the *de novo* assembly of strains Augusto2, CSV1 and I1.3 of *Fusarium fujikuroi*.

Supplementary table 2.3. Genes of the gibberellin and fumonisin gene clusters, and regulators checked for polymorphisms in the three strains.

Supplementary table 2.4. HPLC-MS quantification of GA3, fumonisin B1, fumonisin B2, fumonisin B3 and fumonisin B4 in *F. fujikuroi* strains growing on PDB and YES Agar media.

Supplementary table 2.5. Sheet 1 contains genes present either in strain CSV1 or Augusto2, but not both. The putative function of their closest blast hit is included in the table, as is their absence or presence in the strain I1.3 genome. Sheet 2 contains all genes presenting a deleterious missense polymorphism or a nonsense or frameshift polymorphism either in Augusto2 or CSV1, but not in both. Only genes described by at least a GO term related to regulation of transcription, pathogenesis or metabolism were included. The position of the polymorphism, the amino acid commonly found in that position and the protein putative function are indicated. Sheet 3 contain information about the genes putatively affected by structural variants identified with BreakDancer (Fan *et al*., 2014). Genes were considered to be putatively affected when they had a deletion inside their sequence, or when the extremity of an inversion or translocation was localized inside their sequence.

Supplementary file 2.1. Python3 script used to identify genes putatively affected by structural variations identified with BreakDancer (Fan *et al*., 2014).

Supplementary file 2.2. Vcf file obtained from the SNP calling of the reference genome of *F. fujikuroi* strain IMI 58289.

Supplementary file 2.3. Results of BreakDancer (Fan *et al*., 2014), run with default parameters.

All the supplementary material is available online at: https://drive.google.com/drive/folders/1MAF2Vag3B5sAgS9BPtPI4AOFZWItRAkB?usp=sharing

3: SEQUENCING OF AVIRULENT STRAINS OF *FUSARIUM FUJIKUROI* REVEALS GENES PUTATIVELY INVOLVED IN BAKANAE DISEASE OF RICE

Edoardo Piombo[1,2], Marco Rosati[2], M. Lodovica Gullino[1,2] and Davide Spadaro[1,2]

[1]AGROINNOVA – Centre of Competence for the Innovation in the Agro-environmental Sector, University of Turin, Largo Braccini 2, 10095 Grugliasco, TO, Italy.
[2]Department of Agricultural, Forestry and Food Sciences (DISAFA), University of Turin, Largo Braccini 2, 10095 Grugliasco, TO, Italy.

Corresponding author: Davide Spadaro. Email: davide.spadaro@unito.it

Funding: This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 634179 "Effective Management of Pests and Harmful Alien Species - Integrated Solutions" (EMPHASIS) and from AGER Foundation under Grant no. 2010–2369 ("RISINNOVA - Integrated genetic and genomic approaches for new Italian rice breeding strategies").

INTRODUCTION

The hemibiotrophic pathogen *Fusarium fujikuroi* [teleomorph *Gibberella fujikuroi* (Sawada) Ito in Ito and K. Kimura] (Carter *et al.*, 2008; Desjardins *et al.*, 1997) is the causal agent of bakanae, a disease that can cause up to 75% of crop losses on rice (Saremi *et al*., 2008).

This disease was described for the first time in Japan in 1828, but today bakanae is present in several countries of America, Europe, Asia and Africa. The disease commonly induces symptoms like abnormal height, thin leaves and grains entirely or partially empty, symptoms mainly due to the production of gibberellins, both by the pathogen itself and by inducing gibberellin synthesis in the plant (Matić *et al.*, 2016; Niehaus *et al.*, 2017).

No available rice cultivar is showing complete resistance to the pathogen, even if the rice-*F. fujikuroi* pathosystem has been studied with a variety of approaches, from RNA sequencing to QTL mapping (Fiyaz *et al.*, 2016; Ji *et al.*, 2018; Matić *et al.* 2016). Less susceptible cultivars, like 'Selenio', show less gibberellin production, and a reduced expression of their biosynthetic gene cluster, compared to highly susceptible cultivars like 'Dorella' (Matić *et al.*, 2016; Siciliano *et al.*, 2015). However, certain strains have a virulence mechanism not based on gibberellins and internodes elongation, inducing instead stunting and root and crown roots (Amoah *et al*., 1995; Gupta *et al*., 2015; Karov *et al*., 2009). This difference in behaviour has led to the division of *F. fujikuroi* in two pathotypes, depending on the ability to cause either elongation or stunting (Niehaus *et al*., 2017). A recent work outlined the key role played by temperature, besides genetics, in the development of different symptoms (Piombo et al., 2020).

The metabolomic capacity of *F. fujikuroi* is not limited to gibberellins, with the fungus being able to synthetize fumonisins, fusaric acid, fusarins, bikaverin and fusarubins (Bacon *et al.*, 1996; Balan *et al.,* 2016; Barrero *et al.*, 1991; Desjardins *et al.*, 1997; Studt *et al*., 2012). The biosynthesis of secondary metabolites is

regulated by several global and local regulators, including the nitrogen regulators AreA (Tudzynski *et al.*, 1999) and AreB (Pfannmüller *et al.*, 2017b), and the component of the velvet complex Lae1 (Niehaus *et al.*, 2018).

F. fujikuroi has been sequenced for the first time in 2013 (Jeong *et al.*, 2013; Wiemann *et al.*, 2013), with more strains becoming available in recent years (Bashyal *et al.*, 2017; Chiara *et al.,* 2015; Niehaus *et al.*, 2017; Radwan *et al.*, 2018; Urbaniak *et al.*, 2018; Piombo *et al*., 2020). The genome size of the fungus is around 45 Mb, divided into 12 chromosomes, with repetitive elements constituting less than 1% of the total. The number of genes varies between 13,000 and 15,000 genes, depending on strain, with around 1,200 genes encoding for the secretome. Despite the number of sequenced strains, there has not been to date a study comparing virulent and avirulent isolates of *F. fujikuroi*. Previously, several strains of *F. fujikuroi* were compared for their virulence on rice, by performing pathogenicity tests (Amatulli *et al.,* 2010). As it has recently been demonstrated that the virulence of *F. fujikuroi* strains depends on the temperature (Piombo *et al.,* 2020), some putatively avirulent strains of *F. fujikuroi* were inoculated in susceptible rice plants at two temperatures: 22°C and 28°C. *F. fujikuroi* strains showed greater virulence at higher temperatures.

In this work, we aimed to demonstrate the avirulence of SG4 and C2S, two *F. fujikuroi* strains, and to compare their genome with that of a highly virulent strain I1.3 (Matic *et al*., 2013). The objective was to identify, through a comparative genomic approach, potential genes of *F. fujikuroi* involved in the pathogenicity process, present in the virulent strain and not in the avirulent ones.

MATERIALS AND METHODS

Microorganisms.

The *F. fujikuroi* strains named SG4, C2S and I1.3, previously isolated from rice plants in northern Italy (Amatulli *et al*., 2010) and maintained as glycerol stock at -

80°C in the Agroinnova microorganism collection, were grown on PDB (Potato Dextrose Broth, Sigma Aldrich, St. Louis, Missouri, USA) for 10 days, at room temperature (23°C±1°C) on a rotary shaker (124 rpm). At the end of this period, the conidial suspensions were filtered through sterile gauze, centrifuged for 20 minutes at 6,000 rpm, and resuspended in Ringer solution (Merck Millipore, Burlington, Massachusetts, USA). The Burker chamber was then used to obtain concentrations of 10^5 conidia/ml.

Pathogenicity trials.

The rice cultivar used for the pathogenicity tests was 'Galileo', susceptible to bakanae disease (Amatulli *et al.,* 2010). The seeds were put in water heated at 60°C for 5 minutes, immersed in a solution of 1% NaClO for two 2 minutes and then washed three times with sterile water for 5 minutes. The seeds were then divided in groups of 240 and shaken with 40 ml of the prepared conidial suspensions for 30 minutes on a horizontal shaker.

After 24 h drying on sterile paper, the seeds were sown in sterilized soil (70% white peat and 30% clay, with pH between 5.5 and 6). The N content was 110-190 mg/l, P_2O_5 was of 140-230 mg/l and K_2O was 170-280 mg/l. The plants were grown in growth chamber: 120 rice per strain were planted, plus 120 as control, all divided in biological replicates of 30 plants.

At 2 and 3 weeks after germination, every plant was assigned a disease index, depending on the visible symptoms: 0: healthy plant; 1: reduced dimension, chlorotic leaves; 2: internode elongation, significant yellowing, significant dwarfism; 3: necrosis of the crown; 4: dead or not-germinated plant.

Two trials were conducted: one at 22°C, with a photoperiod of 12 hours of dark and 12 of light, the other at 28°C, with a photoperiod of 8 hours of dark and 16 of light. Each trial was performed twice. For each isolate, the final disease index was calculated as the mean of three biological replicates in each of the two repetitions.

Data analysis was performed using SPSS software (SPSS Inc., version 21.0). Statistical significance was judged at $P < 0.05$. When the analysis of variance was statistically significant, Duncan's test was used to compare the average disease indexes.

***In vitro* growth assays.**

The *F. fujikuroi* strains I1.3, SG4 and C2S were inoculated on PDB (Potato Dextrose Broth, Sigma Aldrich) for 7 days, at room temperature (23°C±1°C) with shaking (124 rpm). At the end of this period, the conidial suspensions were filtered through sterile gauze, centrifuged for 20 minutes at 6,000 rpm, and resuspended in Ringer solution (Merck Millipore). The Burker chamber was then used to obtain concentrations of 10^5 conidia/ml. Petri dishes of YES Agar solid medium (yeast extract 20 g/l, saccharose 150 g/l, $MgSO_4$ 0.5 g/l, trace elements 1 ml/l, agar 20 g/l) were inoculated with 20 µl of conidial suspension placed in the middle of the plate, and left in the dark at 22°C and 28°C for 9 days. The diameter of the colonies was measured every three days during the assay.

DNA extraction and sequencing.

F. fujikuroi strains C2S and SG4 were grown on YES (yeast extract 20 g/l, saccharose 150 g/l, $MgSO_4$ 0.5 g/l, trace elements 1 ml/l) liquid medium for one week, the micelium was isolated with a vacuum pump and liophilyzed (liophilizer FD1.0, De Mori, Milan, Italy). 0.05 g of liophilyzed tissue were used for the DNA extraction with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany), following the instructions of the producer, after 2 minutes shearing with a Tissue Lyser (Qiagen) at maximum speed.

DNA integrity was checked with Nanodrop (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and with Qubit dsDNA assay kit (Thermo Fisher Scientific), standardized at 5 ng/µl. DNA was further observed on 1% agarose gel

(data not shown). Library preparation (Illumina PE 150, San Diego, USA) and MiSeq HighThroughput Illumina sequencing were performed by Novogene. Sequencing of *F. fujikuroi* strain I1.3 was performed by Parco Tecnologico Padano, and is described in Piombo *et al*. (2020).

Assembly and gene prediction.

Adapter removal and read cleaning was executed with Trimmomatic v.036 (Bolger *et al.*, 2014) using default parameters, and the results were checked with fastqc (*https://www.bioinformatics.babraham.ac.uk/projects/fastqc/*). The clean reads were assembled with SPAdes version 3.7.1 (Bankevich *et al*., 2012), and abyss-sealer 2.1.0 (Paulino *et al.*, 2015) was used to perform the gap-closing.

Gene prediction was conducted using MAKER version 2.31.8 (Cantarel *et al.*, 2008) with predictors augustus v.2.5.5 (Stanke and Waack, 2003) and SNAP v.2006-07-28 (*http://korflab.ucdavis.edu/software.html*) used for the *de novo* prediction step. Augustus was run with the "--fusarium" option for gene prediction, while SNAP was trained to obtain a file.hmm specific for our genomes, following the instructions at the following link: *https://wiki.cyverse.org/wiki/display/TUT/Training+ab+initio+Gene+Predictors+for+MAKER+genome+annotation*. The necessary repeat libraries were constructed using the basic procedure (*http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/Repeat_Library_Construction--Basic*). The external data provided to MAKER, and used for the training of SNAP, were all the ESTs, protein sequences and transcript sequences of *F. fujikuroi* available on National Center for Biotechnology Information (NCBI). The option "correct_est_fusion" was activated in the control files to prevent merging of genes.

The annotated genomes of SG4, C2S and I1.3 were deposited in the European Nucleotide Archive with the accession numbers GCA_901677965, GCA_901677955 and GCA_902702945, respectively.

SNP calling.

Reads were mapped on the genome of *F. fujikuroi* IMI 58289 (RefSeq: GCF_900079805.1) using bowtie2 (Langmead and Salzberg, 2012), duplicated reads were identified with MarkDuplicates v. 4.1.2.0 (https://software.broadinstitute.org/gatk/documentation/tooldocs/4.0.3.0/picard_sam_markduplicates_MarkDuplicates.php) and the actual SNP calling was performed with HaplotypeCaller v. 4.1.2.0 (Poplin *et al.*, 2017). Identified polymorphisms were evaluated with SNPEff v. 4.2 (Cingolani *et al.*, 2012), and polymorphisms not covered by reads in all the strains were not considered for further analysis.

Other bioinformatic analyses.

The proteins of *F. fujikuroi* IMI 58289 that had a missense, nonsense or frameshift polyorphisms in SG4 but not in I1.3 strain were blasted against the SG4 proteome, and their homologs were identified.

BLAST2GO v. 2.8 (Conesa *et al.*, 2005) and InterProScan v. 5.24.63 (Quevillon *et al.*, 2005) were used to conduct the functional analysis of the whole I1.3 proteome, and of proteins of SG4 that had a deleterious missense, nonsense or frameshift polymorphisms that were absent in I1.3.

The secretome of I1.3 was predicted following the steps presented by Levin *et al.* (2019), analyzing sequentially the proteome with SignalP 5.0 (Nielsen *et al.*, 2017), TargetP 1.1 (Emanuelsson *et al.*, 2000), TMHMM 2.0 (Krogh *et al.*, 2001) and big-PI (Eisenhaber *et al.*, 2004).

The proteins present in I1.3, but not in SG4, as well as the proteins of SG4 that had a missense, nonsense or frameshift polymorphisms that were absent in I1.3, were

also analyzed using BlastKOALA (Kanehisa *et al*, 2016) and CAT (Park *et al.*, 2010). The same proteins were also blasted against the PHI-base v. 4.6 database (Winnenburh *et al.*, 2006; Urban *et al.*, 2016), with a threshold of e^{-5} in the evalue and a minimum query coverage of 50%. PROVEAN Protein (Choi *et al*., 2012) was used to analyse proteins of SG4 having a missense polymorphism that was absent in I1.3, if they had a hit with 70% identity and 50% query coverage on a PHI-base protein whose knockout could reduce virulence or indicated as an effector.

Comparative genomics.

The proteins putatively related to virulence and pathogenicity, obtained from the previous analyses, were blasted against the proteomes of other provenly virulent strains of *F. fujikuroi*. The considered strains were: Augusto2 (GCA_009663095.1), CSV1 (GCA_009663055.1), B14 (GCA_900096505.1), B20 (GenBank: GCA_900096605.1), C1995 (GenBank: GCA_900096645.1), E282 (GenBank: GCA_900096705.1), FSU48 (GenBank: GCA_900096685.1), IMI58289 (GenBank: GCA_900079805.1), m567 (GenBank: GCA_900096615.1), MRC2276 (GenBank: GCA_900096635.1) and NCIM1100 (Genbank: GCA_900096625.1).

Preparation of samples for gene expression analysis.

The *F. fujikuroi* strains I1.3, C2S and SG4 were grown on 6 PDA Petri dishes for 7 days, with a temperature of 22°C and a photoperiod composed of 12 hours of light and 12 of darkenss. On the 7th day, rice seeds cv. 'Galileo', susceptible to bakanae disease (Amatulli *et al.*, 2010), were deposed on 3 of the Petri dishes, while the others were kept as controls. The seeds were previously thermally treated by dipping in water at 60°C for 5 min, immersed in a solution of 1% NaClO for 2 min and then washed three times with sterile water for 5 min, to ensure proper decontamination. After the introduction of seeds, Petri dishes were left at 22°C, with a photoperiod composed of 12 hours of light and 12 of darkenss. After 48 hours, each seed was

collected with sterilized tweezers and rubbed on the micelium underneath it, in order to collect the fungal mycelium grown in close contact with the seed itself.

RNA extraction and qPCR

RNA was extracted from the mycelium growing directly on or under the seeds in the target plates, while all the mycelium of the plate was collected for the controls. The extraction was performed with the Spectrum Plant Total RNA Kits (Sigma-Aldrich, St. Louis, Missouri, United States), following manufacturers'instructions. The extracted RNA was quantified by Nanodrop (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and purified using the TURBO DNA-free kit (Ambion, Foster City, California, United States). The samples were then checked for DNA contamination by PCR, using primers for the gene CCT71712.1 (**Supplementary Table 1**). After verificating the absence of DNA contamination, the cDNA was obtained using with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, California, United States). The samples were then used in real time qPCR (Applied Biosystems, Foster City, California, United States), with primers for CCT71712.1. The PCR mix were composed of 5 µl of Applied Biosystems SYBR Green Power Mix, 2 µl of cDNA, 0.15 µl of each primer (10 µM) and 2.7 µl of nuclease free water. The thermal cycler protocol was the following: 95°C for 10 min, followed by 40 cycles (95°C for 15 s; 58°C for 60 s) and 95°C for 15 s. The ubiquitin *F. fujikuroi* gene (Wiemann *et al.*, 2013) was chosen as housekeeping gene, and the sequences of the primers used are presented in **Supplementary Table 1**. REST (Pfaffl *et al*., 2002) was used to determine which genes were overexpressed in the mycelium growing around the seeds.

A standard curve built upon five serial dilutions (1:5), using the DNA of strain I1.3, was used to test the efficiency of the newly designed primers.

RESULTS

Pathogenicity trials.

Three strains of *F. fujikuroi* were inoculated in rice seeds of the susceptible cultivar 'Galileo' in order to investigate their ability to induce bakanae when rice was grown at 22°C or at 28°C. SG4 and C2S resulted non pathogenic in both trials, even if some C2S-inoculated plants presented mild bakanae symptoms. I1.3 was highly virulent in every situation, especially at the higher temperature (**Figure 3.1**). The symptoms showed by the plants grown at 22°C and 28°C at 3 weeks after germination are reported in **Figures 3.2** and **3.3**, respectively. Plants inoculated with SG4 had a disease index similar to the controls, while C2S inoculated plants occasionally showed symptoms, but the disease index was not statistically different at either temperatures.

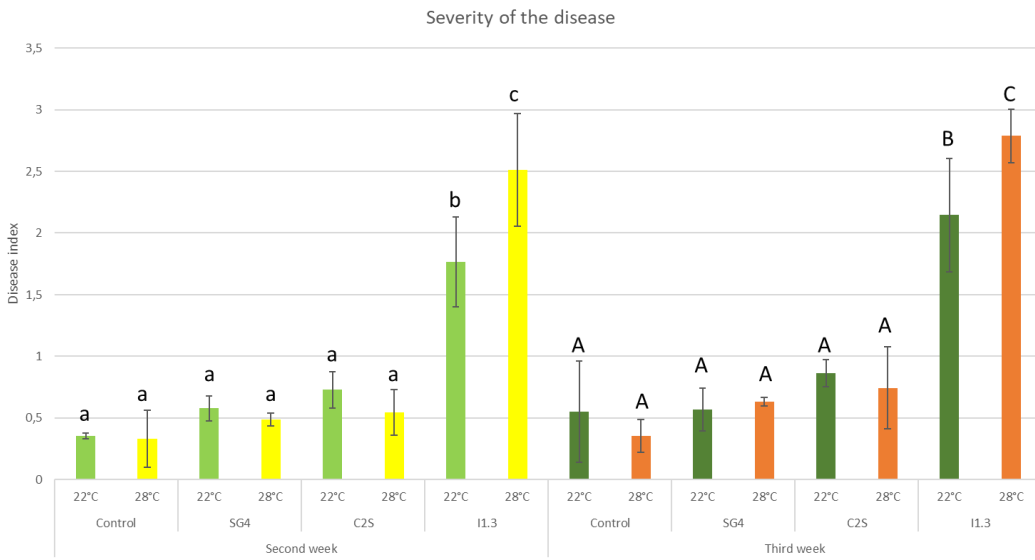

Figure 3.1: disease indexes of rice plants (cv. Galileo) inoculated with the 3 studied strains of *F. fujikuroi*. Values with the same letter are not statistically different according to the Duncan's test (P<0.05). Statistical analysis was performed separately for the two weeks.

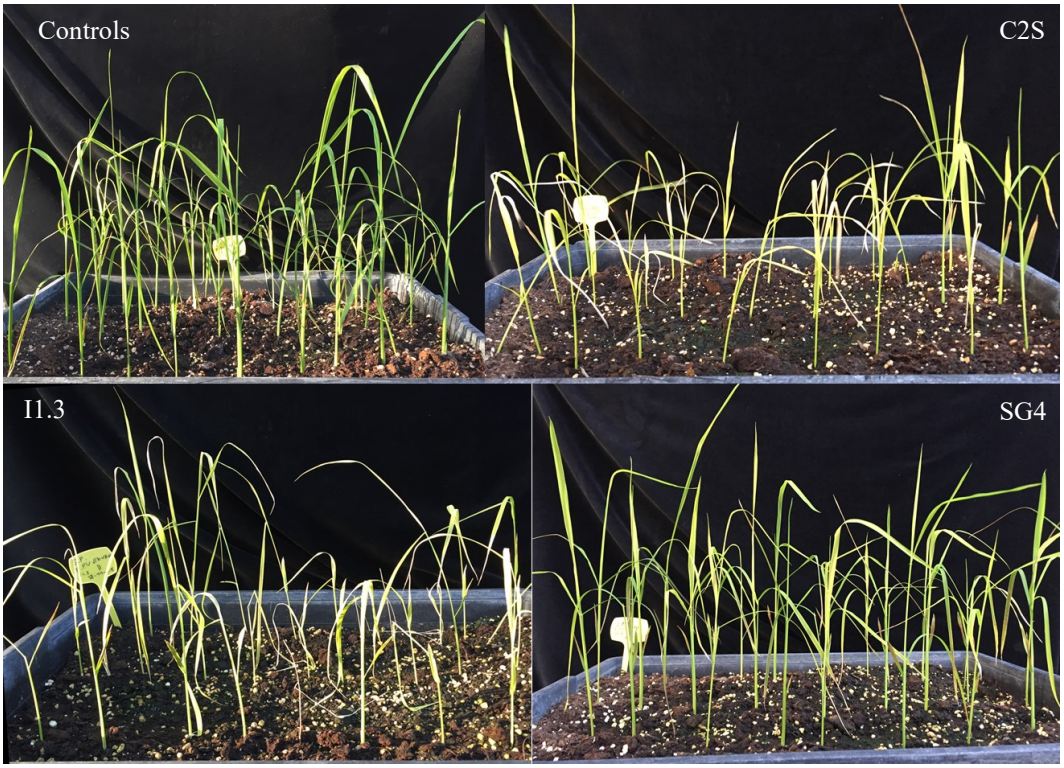

Figure 3.2: rice plants cv. Galileo inoculated with strains I1.3, C2S and SG4 of *F. fujikuroi*. The plants were grown at 22°C, with a photoperiod composed of 12 hours of dark and 12 of light. The pictures were taken at the third week after germination.

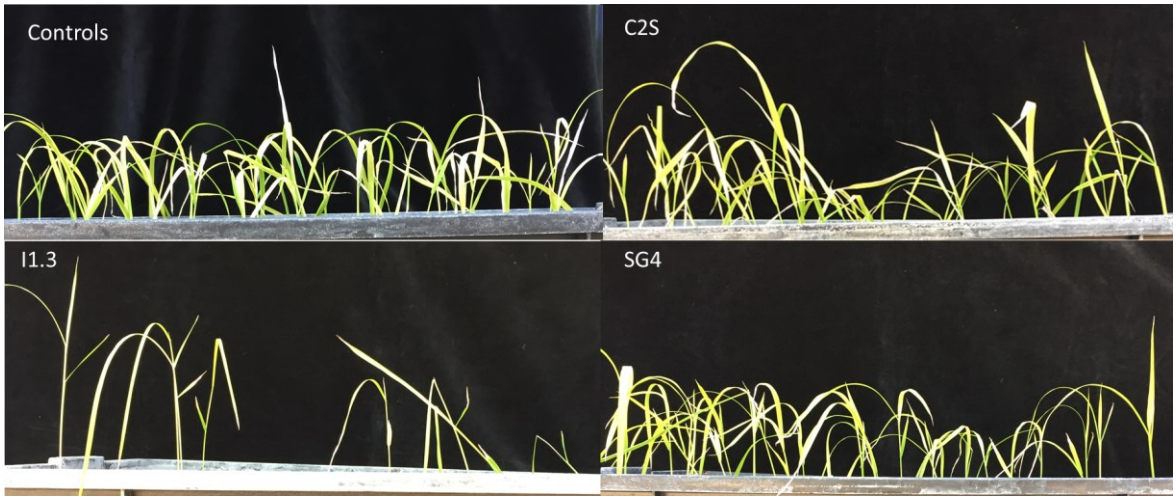

Figure 3.3: rice plants cv. Galileo inoculated with strains I1.3, C2S and SG4 of *F. fujikuroi*. The plants were grown at 28°C, with a photoperiod composed of 8 hours of dark and 16 of light. The pictures were taken at the third week after germination.

***In vitro* growth assays.**

The results of the growth assays performed on YES Agar are presented in **Figure 3.4**. At 22°C, strain I1.3 had a faster growth, reaching an average colony diameter higher than the other strains at almost every time point, even if its growth was not statistically different from SG4 at 3 days after inoculation. At 28°C, however, I1.3 growth was comparable to that of the low virulent C2S, but still higher than that of the avirulent SG4. The higher temperature had a negative effect on the growth speed of all the strains, but it impaired SG4 more than the others, while C2S showed a significant difference between the two temperatures only at 9 days after inoculation.

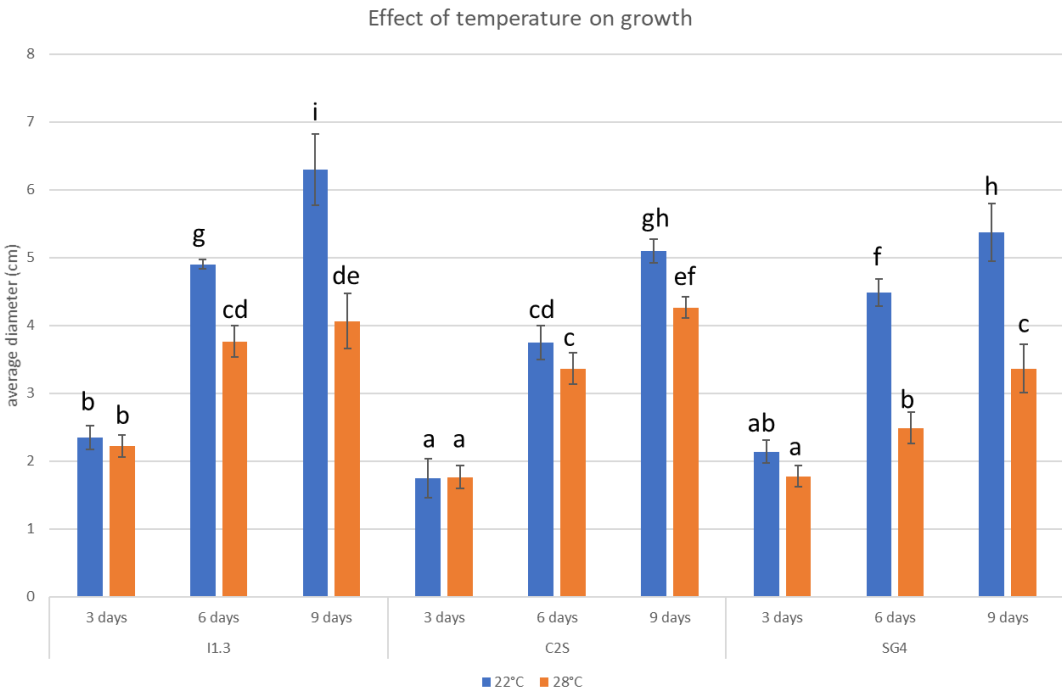

Figure 3.4: average colony diameter measured after 3, 6 and 9 days of growth in YES Agar plates inoculated with the strains I1.3, C2S and SG4 of *F. fujikuroi.* The fungi were kept in the dark at 22°C and 28°C. Values with the same letter are not statistically different according to the Duncan's test (P<0.05).

Sequencing, gene prediction and annotation.

F. fujikuroi strains I1.3, C2S and SG4 were sequenced with a respective coverage of 62X, 272X and 254X, producing 2.12 millions of raw paired ends reads and 16.57 millions mate pair reads for I1.3, 83 millions of paired end reads for C2S, and 79 millions of paired end reads for SG4. The results of assembly and gap-closing are presented in **table 3.1**. The obtained genomes had a quality sufficient to proceed with the other analysis, with a low number of scaffolds and a size similar to the expected one.

Strain	SG4	C2S	I1.3
Number of scaffolds (Mb)	544	490	1024
Biggest scaffold	1.78	1.37	2.82
Total length (Mb)	46.5	45.82	45.85
GC percentage	47.32	47.34	47.37
N50 (bp)	546,219	436,322	790,579
#_N's_each_100_kbp	4.28	0.07	179.49
Coverage	254X	272X	62X

Table 3.1. Statistics of the three assemblies obtained in this study.

MAKER predicted 13,527, 14,795 and 14,534 genes, respectively for I1.3, SG4 and C2S. Among them, 201 putative proteins were present only in the *F. fujikuroi* virulent strain I1.3, while 228 were predicted in I1.3 and C2S but not in SG4.

After SNP calling (**Supplementary File 3.1**) 158,434 polymorphisms were identified, and 56% of them were present in the three strains (**Figure 3.5**). 1,247 proteins had missense mutations only in SG4, while 320 presented missense mutations in SG4 and C2S, but not in I1.3.

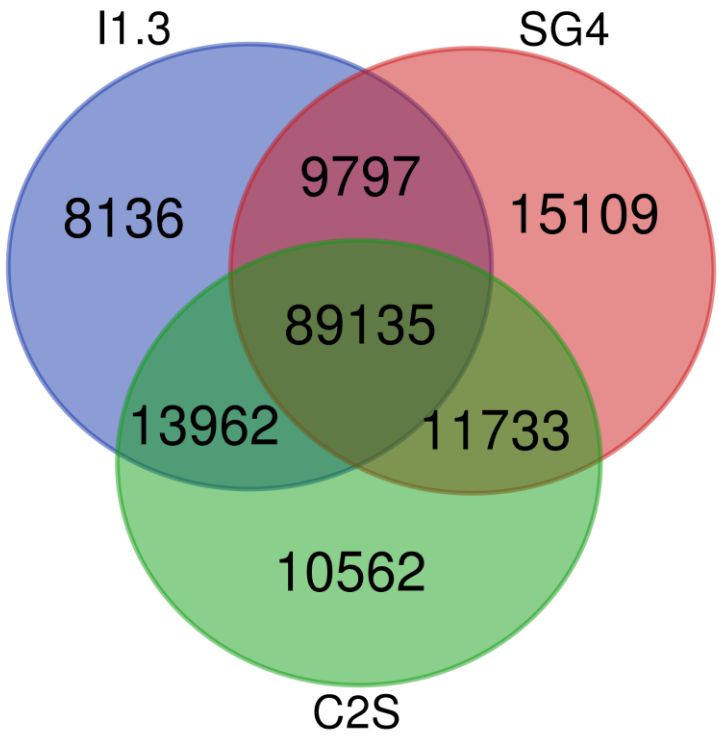

Figure 3.5: Venn graphic showing the distribution of polymorphisms among the *F. fujikuroi* strains C2S, SG4 and I1.3. The genome of strain IMI 58289 was used as reference in the SNP calling. Image obtained with the software at the following link: *http://bioinformatics.psb.ugent.be/webtools/Venn/*

Comparative genomics

All the genes of the gibberellin gene cluster were present in all the genomes, and no missense, frameshift or nonsense polymorphisms were identified in any of the strains.

Furthermore, we looked for secreted I1.3 proteins not present in SG4: 1,028 proteins of I1.3 were predicted to be secreted, 13 of them are unique to I1.3 and 10 shared with C2S but not present in SG4. This group includes three putative hydrolases and one protein with oxydoreductase activity (**Supplementary Table 3.2**). Among the secreted proteins shared by all the strains, 28 presented a missense polymorphism

putatively deleterious to their function in SG4 but not in I1.3. One of these proteins presented the same polymorphisms in C2S.

The analysis with blast2GO assigned 2,598 Gene Onthology (GO) terms to 9,764 I1.3 proteins (**Supplementary File 3.2**). The most common GO terms for the biological process were “metabolic process” and “cellular process”. Regarding the molecular function, the most frequent GO terms were “catalytic activity” and “binding”, while “cell”, “cell part”, “membrane” and “membrane part” were the most common GO terms describing the cellular components. 1,002 proteins of I1.3 were assigned GO terms associated with transcription or regulation of transcription: 9 of these were unique to I1.3 and 16 were present also in C2S but absent in SG4 (**Supplementary Table 3.3**). The SNP calling revealed that 15 of these proteins contain a missense polymorphism putatively deleterious for their function in SG4, but not in I1.3. One of these proteins was mutated in C2S as well.

The blast analysis against Pathogen Host Interactions databese (PHI-base) made it possible to identify genes associated to reduced virulence when knocked out in other pathogens (DeZwaan *et al.*, 1999; Ogiso *et al.*, 2004). 18 of these genes were absent in *F. fujikuroi* avirulent strain SG4, 7 of them were unique to I1.3 and 11 were present also in low virulent C2S strain. Furthermore, 3 genes presenting deleterious missense polymorphisms in SG4, but not in I1.3, were identified. Of these, one had the same polymorphism in C2S, while the remaining 2 were present only in SG4. Together, these genes codify for 24 proteins that may be involved in the pathogenicity of *F. fujikuroi* on rice (**Supplementary Table 3.4**). Among the transcription factors, secreted proteins and factors of virulence identified with the previous analyses, 13 were considered as putatively related to virulence in *F. fujikuroi*, according to their functional annotation and their absence or non-functioning in the genomes of the avirulent strains. These genes, presented in table 2, include 5 secreted proteins, 5 transcription factors and 3 putative factors of virulence identified with PHI-base.

Gene	Related to	Role	Absent/not functioning in strain	Number of virulent strains with the protein (out of 12)
I13_2472	Endo-arabinase	Degradation of the plant cell wall	SG4, C2S	12
CCT71712.1	Dipeptidyl aminopeptidase B	Cellulose-response	SG4	12
CCT74357.1	Alpha/beta-glucosidase agdC	Degradation of cellulose	SG4	12
I13_2983	Cholinesterase precursor		SG4, C2S	10
CCT61573.1	Cholinesterase		SG4	12
I13_12502	Regulator of cutinases	Degradation of the plant cell wall	SG4, C2S	1
I13_10754	Regulator of conidial development *fluffy*	Source of inoculum	SG4, C2S	11
I13_10486	Regulator of sterol metabolism UPC2	Growth in anaerobic environment	SG4, C2S	6
CCT74231.1	Homolog of transcriptional activator acu15	Utilization of "difficult" carbon sources like acetate	SG4	12
CCT70723.1	Homolog of transcriptional activator acu15	Utilization of "difficult" carbon sources like acetate	SG4	12
I13_7187	Ubiquitin homeostasis protein lub1	Resistance to various stresses, temperature included	SG4, C2S	11
I13_7541	PTH11-like membrane protein	Appressorium formation and pathogenicity	SG4, C2S	7
I13_7295	Nitrogen metabolic regulator NMR	Repression of AreA-dependent genes	SG4, C2S	11

Table 3.2. Identified genes putatively involved in the virulence of *F. fujikuroi.*

Gene-expression analysis

The three strains were grown on PDA plates with or without the addition of rice seeds, and the gene codifying the protein CCT71712.1, related to dipeptidyl aminopeptidase B, a class of enzymes involved in cellulose-response in *Aspergillus aculeatus* (Tani *et al.*, 2017), was selected for gene expression analysis (**Figure 3.6**). It resulted overexpressed in the virulent strain I1.3 and in the scarcely virulent C2S when they were grown on Petri dishes with rice seeds, compared to the same strains grown on PDA. The gene was expressed also in the strain SG4, where this protein was predicted to be not functioning, but no overexpression was detected when this isolate was grown on rice seeds.

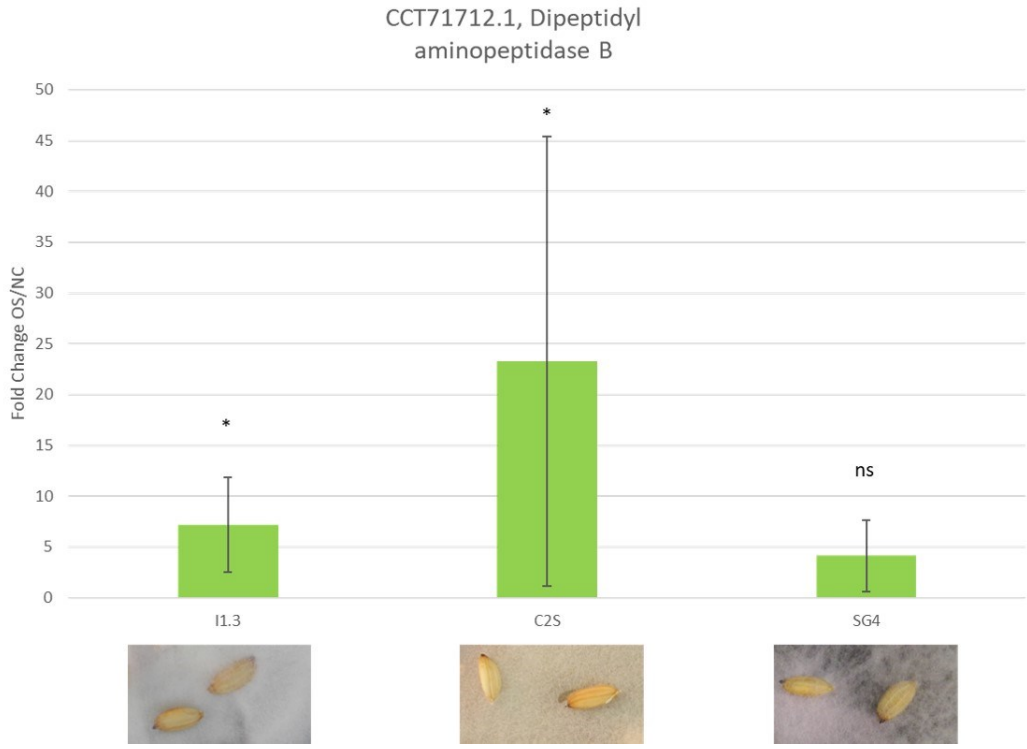

Figure 3.6: CCT71712.1 expression levels in strains I1.3, SG4 and C2S of *F. fujikuroi* grown in contact with rice seeds, divided by the PDA contols. The error bars represent the standard deviation. The asterisk (*) indicate strains in which the gene was significantly overexpressed in the mycelium grown in contact with rice seeds, while "ns" stands for "not significant".

DISCUSSION

The *in vivo* trials demonstrated that the strain I1.3 is highly virulent, SG4 is not pathogenic and C2S shows low to no virulence depending on environmental conditions. As already observed in Piombo *et al.* (2020), I1.3 was more virulent at higher temperatures, even if in the current study the same strain showed higher speed growth at 22°C. This suggests that the higher virulence at higher temperatures is probably due to the production of secondary metabolites related to the disease, and not to the faster growth of the fungus. Frequently, secondary metabolites are produced at temperatures different from the optimal growth conditions of the fungus: this happens for aflatoxins in *Aspergillus spp.* (Alkhayyat and Yu, 2014), trehalose in *Fusarium verticilloides* (Boudreau *et al*., 2013) and fumonisins in both *Fusarium verticilloides* and *Fusarium proliferatum* (Marin *et al*., 2004). In particular, *F. fujikuroi* has been observed to have higher gibberellin production at 30°C (Dendooven *et al*., 2000), and gibberellins are one of the main factors of virulence contributing to bakanae disease. SG4 seemed particularly impaired in the growth at 28°C, and its reduced capacity to grow at the temperatures favorable for the induction of the disease could play a role in its avirulence.

After assembly and annotation, the comparison of the genomes of the strains allowed for the identification of many genes putatively involved in pathogenicity. Three main bioinformatic approaches were followed to identify these genes: secretome analysis, investigation of transcription factors and analysis of virulence genes previously studied in other organisms.

Secretome

Effectors are a group of proteins and metabolites, which are used by pathogens to bypass the plant immune response and cause the necessary changes for the induction of disease (Fouché *et al*., 2018). Effectors are able to function only when in contact

with plant cells, and therefore they are commonly found in the pathogen secretome (Stergiopoulos and de Wit, 2009; Djamei *et al.*, 2011).

The secreted proteins of the three sequenced strains of *F. fujikuroi* characterized by different virulence were compared. Among the 1,028 secreted proteins of I1.3, 23 were not present or putatively not functioning in the avirulent strain SG4 and, considering their biological function, 5 might contribute to the avirulence of this strain. Two of them, I1.3_2983 and CCT61573.1, are respectively a cholinesterase precursor and a cholinesterase. It has been observed that knockout mutants of Ric8 proteins (resistant to inhibitors of cholinesterase 8 proteins) in *F. graminearum* have reduced virulence, as well as reduced vegetative growth, conidiation, pigment production and deoxynivalenol biosynthesis (Wu *et al.*, 2015). I1.3_2983 is absent in two of the virulent strains, while CCT61573.1 was found in all of them.

I1.3_2472, on the other hand, is an endo-arabinase, a class of proteins normally involved in the degradation of the plant cell wall (De Vries *et al.*, 2000). This is of interest because *F. fujikuroi* was observed to express this kind of genes more on susceptible genotypes than on resistant ones (Bashyal *et al.*, 2017), suggesting that the breaching of the host cell barrier during plant-pathogen interactions is an important step in the pathogenicity. All of the considered virulent strains had a gene codifying for this protein in their genomes. Another enzyme relevant for similar reasons is CCT74357.1, with a putatively deleterious polymorphisms present in SG4 but not in I1.3. It is present in all the virulent strains, and it is similar to agdC, an alpha/beta-glucosidase involved in the degradation of cellulosic biomass (Bauer *et al*., 2006), with homologs in both *A. flavus* and *A. nidulans*. Glucosidases are necessary for cell wall integrity and infectious growth in *M. oryzae* (Li *et al*., 2016). Another protein of interest is CCT71712.1, predicted not to function in SG4 and present in the virulent strains as well. The protein is related to dipeptidyl aminopeptidase B, and this class of enzymes is involved in cellulose-response in *Aspergillus aculeatus* (Tani *et al.*, 2017). Since cellulose is the main component of

plant cell wall, this gene could be involved in pathogenicity for the same reason of I1.3_2472 and CCT74357.1. This gene was selected for gene expression analysis, and it resulted overexpressed in I1.3 and C2S mycelium when grown in presence of rice seeds, compared to the same strains on PDA, while its expression did not change significantly in SG4, in which this protein was putatively not functioning (**Figure 3.6**).

Transcription factors

Transcription factors are fundamental for pathogen adaptation to the host environment, induction of disease and tolerance of plants (Van der Does and Rep, 2017).

Among the 24 transcription factors present and functioning in *F. fujikuroi* strain I1.3 and not in SG4, 5 are of particular interest. I1.3_12502 is similar to a regulator of expression of cutinases, a class of enzymes necessary to degrade the plant cell wall, and factors of virulence in many pathogens (Köller *et al*., 1982; Liu *et al.*, 2016; *Lu et al*., 2018). However, this gene was found only in I1.3 and not in any other virulent strain, so, even if it could have a role in virulence, it is certainly not a fundamental part of the pathosystem. I1.3_10754 is a homolog of *fluffy*, responsible for conidial development (Bailey and Ebbole, 1998; Bailey-Shrode and Ebbole, 2004). The conidia adhering to the seed are the primary source of inoculum in bakanae (Gupta *et al.*, 2010; Sun, 1975), and a reduced conidial production could diminish the ability of a strain of causing the disease. However, it must be noted that the starting conidia concentration used for seeds inoculation in this study was the same for all the strains, so a reduced conidial production cannot be solely responsible for the avirulence of SG4. Furthermore, this gene was absent in one of the considered virulent strains. Another gene that could play a role in virulence is I1.3_10486, a homolog of UPC2, which regulates sterol metabolism in *Candida albicans*, where the knockout of this gene causes the inability of growing in anaerobic environment (White and Silver,

2005). The ability of growing in low-oxygen conditions proved to be important for the virulence of another hemibiotrophic rice pathogen, *M. oryzae* (Choi *et al.*, 2015), and it could be so for *F. fujikuroi* too. However, probably I1.3_10486 is not the only gene able to permit anaerobic growth in *F. fujikuroi*, because it was found only in 6 of the considered virulent strains. Both CCT74231.1 and CCT70723.1, homologs to the transcriptional activator acu-15, were predicted to be not functioning in SG4 (but not in I1.3 or C2S). This transcription factor usually mediates the utilization of "difficult" carbon sources, such as acetate and lactate, even if it was observed in *A. niger* that a basal regulation of genes involved in the utilisation of these sources remain active even when the homolog of acu-15 is knocked out (Meijer *et al.*, 2009). Still, the putative inactivity of both genes in SG4 might suggest a role of acetate utilisation in the virulence of *F.* fujikuroi, expecially because both genes are present in all the virulent strains of *F. fujikuroi*. Furthermore, the expression of acu-15 is normally repressed by CREA, a transcription factor whose glucose repression activity also regulates SUC2 invertases (Tanaka *et al*., 1998). This is particularly interesting because protein CCT73860.1, a probable SUC2-invertase, is among the proteins predicted to not function both in C2S and SG4.

Comparison with PHI-base

PHI-base is a database containing pathogen proteins and the phenotype of their knockout or overexpression (Winnenburg *et al.*, 2006). A group of 21 genes putatively absent or not working in SG4, but present and active in I1.3, had a good blast hit against this database. The proteins codified by these genes have several functions, ranging from transporters to methyltransferases and hydrolases (**Supplementary Table 4**). Within this set, 3 proteins were especially interesting. I1.3_7187 is a homolog of the ubiquitin homeostasis protein LUB1, whose absence causes a diminished resistance to various stresses, temperature included, in *Saccharomyces pombe* (Ogiso *et al.*, 2004). *F. fujikuroi* is commonly more virulent

at high temperature (Piombo *et al*., 2020), so a diminished resistance to heat might impair virulence. The gene was present in all but one of the considered virulent *F. fujikuroi* genomes. I1.3_7541 is a PTH11-like membrane protein, and proteins of this class exhibit host-preferential expression in *F. graminearum* (Harris *et al.*, 2016) and are necessary for appressorium formation and pathogenicity in *Magnaporthe oryzae* (DeZwaan *et al.*, 1999), so they might also play a role in bakanae disease: *F. fujikuroi* was recently proven to form globose appressoria during its penetration in rice roots (Sunani *et al*., 2019). However, this gene is absent in 5 of the virulent strains, so the formation of appressorium-like structures can probably be managed by other genes in *F. fujikuroi*. I1.3_7295 is related to nitrogen metabolic regulation protein NMR, a protein involved in the repression of AreA-dependent genes in presence of rich nitrogen sources (Schönig *et al*., 2008). It is present in all but one of the virulent strains, but absent in both SG4 and C2S.

Putting the data together (**Table 2**), four of the 13 descripted proteins (13_2472, CCT71712.1, I1.3_12502 and CCT74357.1) are in some way related to the degradation of plant cell wall. This result suggests that the breaching of the host cell barrier during plant–pathogen interactions is an important step in the pathogenicity of *F. fujikuroi*, which was observed to express this kind of proteins more on susceptible genotypes than on resistant ones (Bashyal *et al.*, 2017).

Potential biocontrol action.

SG4 is the first avirulent *F. fujikuroi* strain whose genome is deposited on public databases. Avirulent *Fusarium spp*. isolates have been proved to be effective in the biocontrol of *Fusarium* wilt diseases (Garibaldi et al., 1994; Janvier et al., 2007; Spadaro and Gullino, 2005). Yeast or bacterial antagonists have already been evaluated for the biological control of bakanae of rice (Matic et al., 2014), but there is a lack of knowledge about the potential use of avirulent strains of *F. fujikuroi*. The publication of an avirulent genome of this species constitutes the first step

towards the comprehension and development of bioproducts able to protect rice seeds from this disease.

CONCLUSIONS

This work presents the first publicly available genome of avirulent *F. fujikuroi* isolated from rice plants. This genome is compared with two other *F. fujikuroi* strains from the same area, one very virulent and the other scarcely virulent. The comparison made it possible to identify many putatively proteins correlated to *F. fujikuroi* pathogenicity on rice, whose presence was checked in all the available genomes of virulent *F. fujikuroi* strains. Of these, 13 constitute expecially good candidates for knockout experiments aimed to prove their role in *F. fujikuroi* pathogenicity. One of them was found to be overexpressed during *F. fujikuroi* interaction with rice seeds. The mechanisms by which the bakanae disease is induced on rice are not clear yet, and the characterization of new genes related to the virulence is fundamental to add new tiles to the complex puzzle of rice-*F. fujikuroi* interaction. Beside the results presented here, the availability in databases of an avirulent *F. fujikuroi* genome will make it easier to identify more genes related to virulence in this rice pathogen, as more *F. fujikuroi* strains are sequenced.
Furthermore, this work constitutes the basis for the utilization of *F. fujikuroi* strain SG4 in the biocontrol of the bakanae disease on rice.

SUPPLEMENTARIES

Supplementary Table 3.1. List of primers used in the study.

Supplementary Table 3.2. Functional annotation of the proteins putatively secreted by I1.3 not present or predicted as not functioning in SG4. The functional annotation was done with BLAST2GO, InterProScan, CAT and BlastKOALA.

Supplementary Table 3.3. Functional annotation of the putative I1.3 transcription factors not present or predicted as not functioning in SG4. The functional annotation was done with BLAST2GO, InterProScan, CAT and BlastKOALA.

Supplementary Table 3.4. Functional annotation of proteins predicted to be either not present or not functioning in the avirulent *F. fujkuroi* strain SG4. All the proteins have a hit on PHI-base with a maximum evalue of e^{-5} and a minimum query coverage of 50%. The functional annotation was done with BLAST2GO, InterProScan, CAT and BlastKOALA.

Supplementary file 3.1. SNP calling, performed with HaplotypeCaller. The reference genome is *F. fujikuroi* strain IMI 58289.

Supplementary file 3.2. Blast2GO analysis of the proteome of *F. fujikuroi* strain I1.3.

All the supplementary material is available online at: https://drive.google.com/drive/folders/1MAF2Vag3B5sAgS9BPtPI4AOFZWItRAkB?usp=sharing

4: IMAGING THE INVASION OF RICE ROOTS BY THE BAKANAE AGENT *FUSARIUM FUJIKUROI* USING A GFP TAGGED ISOLATE

Maria Aragona[1], Lidia Campos Soriano[2], Edoardo Piombo[3,4], Elena Romano[5], Blanca San Segundo[2], Davide Spadaro[3,4], Alessandro Infantino[1]

[1]Council for Agricultural Research and Economics (CREA), Research Centre for Plant Protection and Certification, Via C.G. Bertero 22, 00156 Rome, Italy.

[2]Centre for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Edifici CRAG, Campus UAB Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.

[3]Dept. Agricultural, Forestry and Food Sciences (DISAFA), AGROINNOVA - Centre of Competence, University of Torino, Largo Braccini 2, 10095 Grugliasco (TO), Italy

[4]AGROINNOVA – Centre of Competence for the Innovation in the Agro-environmental Sector, University of Turin, Largo Braccini 2, 10095 Grugliasco, TO, Italy.

[5]Centre of Advanced Microscopy "P. Albertano", Department of Biology, University of Rome "Tor Vergata", 00133 Rome, Italy

Corresponding author: Maria Aragona, maria.aragona@crea.gov.it

INTRODUCTION

Bakanae is a rice disease caused by the hemibiotrophic fungal pathogen *Fusarium fujikuroi* [teleomorph *Gibberella fujikuroi* (Sawada) Ito in Ito & K. Kimura]. It was originally observed in Japan in 1828 (Ito and Kimora, 1931), but it is now present in several countries in America, Europe, Asia, Oceania and Africa (Amatulli *et* al., 2010; Carter *et* al., 2008; Chen *et* al., 2016; Desjardins *et* al., 2000; Jeon *et al*., 2013; Karov *et* al., 2005; Khan *et al*, 2000; Kim *et al*., 2015; Zainudin *et al*., 2008). The meaning of the name is "foolish seedling", and it is due to the main symptom of the fungus: the elongation and thinning of internodes, inducing frail stems and abnormal height, besides thin leaves and grains entirely or partially empty. The losses caused by bakanae depend on climate, rice cultivars and disease strain. In Thailand losses of 3-15% (Kanjanasoon *et al*., 1965) were registered, while in Macedonia they are between 2% and 20% (Karov *et* al., 2005) and in Japan they go from 20% to 50% (Ito and Kimura, 1931). In Iran losses of even 75% were reported (Saremi *et al*., 2008).

An excellent overview of the rice-*F. fujikuroi* pathosystem has been given by Gupta and colleagues (2015). The disease is both seedborne and soilborne. However, the potential for infection in soil rapidly decreases. 93% of seeds planted immediately after soil inoculation develop the disease, while only 0.7 % of seeds does it if they are planted after 90 days, with no disease at all occurring after 180 days (Kanjanasoon *et al*., 1965). However, the fungus can survive as hyphae on infested seeds for much longer, lasting 4-10 months at room temperature and more than 3 years at 7°C (Kanjanasoon *et al*., 1965). Seeds can become infested when they come in contact with ascospores, which are produced on diseased plants and moved to other plants during flowering by wind. Seeds can also be contaminated by the fungus during the harvest, when they can be reached by conidia produced on diseased and dead plants. A third source of infection is represented by conidia and mycelium contaminating the water used to stimulate germination in soaked seeds (Karov *et al*.,

2009). The fungus infects seedlings through the roots and crown (Sun, 1975), allowing the fungus to invade the plant without producing visible symptoms, so that *F. fujikuroi* can be found in apparently healthy seeds.

In a recent work Sunani and colleagues (2019) studied infectious structures, penetration and colonization of *F. fujikuroi* in seeds and seedlings of rice. They were able to observe that the fungus usually enters roots by penetration with infection hyphae, infection cushions and globose appressoria, structures that make the pathogen able to colonise both the intracellular and intercellular spaces of rice roots. Recently, a *F. fujikuroi* isolate has been transformed by introducing the green fluorescent protein (GFP) into protoplasts by a polyethylene glycol (PEG)-mediated method (Lee at al., 2018). The basal stem of susceptible and resistant rice varieties, inoculated with a GFP-tagged isolate, were analysed by confocal microscopy to investigate the cultivar-dependent differences in early rice-*F. fujikuroi* interactions. The PEG-mediated method has also been used for introducing the *gfp* and *red* gene into *F. fujikuroi*, for visualizing interaction with biocontrol agents (Watanabe et al., 2007; Kato et al., 2012). The advantage of GFP protein as a reporter is that it allows in vivo imaging of fungal hyphae during its interaction with the host plant. Hyphae of gfp-expressing fungal strains can be visualized both in real time and in living tissue using fluorescence microscopy without extensive manipulation. Organisms that undergo transformation and at the same time carry fluorescent reporter proteins are very important tools to study plant-pathogens interaction and were frequently used to observe pathogen behaviours in plant tissues under various physiological conditions (Lagopodi et al., 2002; Oren et al., 2003; O'Connell et al., 2004).

In the present study, a *A. tumefaciens*-based method has been developed for introducing the *gfp* gene into a virulent *F. fujikuroi* strain. A GFP-tagged isolate was then used to investigate the infection and colonization in susceptible and resistant interactions with host plant by confocal microscopy analysis of rootlets.

Quantification of gene expression of genes related to bakanae disease in the rootlets was also performed.

MATERIALS AND METHODS

Fungal strains, growth conditions and seed inoculation

Fusarium fujikuroi strains, listed in **supplementary table 4.1** were grown on potato dextrose broth (PDB) at 23˚C. Conidia were collected from liquid cultures and adjusted to the concentration of 10^6 ml^{-1} for inoculation of seeds of the susceptible rice cultivar Galileo, as described in Valente *et al*. (2016). Plants were grown in the greenhouse at 25-28 °C for 30 days, to test pathogenicity of GFP transformants, compared to the wild type isolates. Bakanae symptoms were evaluated as described by Kim et al. (2014), by calculating the incidence of the disease in the inoculated vs untreated plants. Plants showing thinness, elongated seedlings, yellowish colouring or dead, after infection, were classified as unhealthy plants.

Development of a gfp-expressing Fusarium fujikuroi strain

The *F. fujikuroi* isolates named Ff 192, Ff 297, Ff 364 and Ff 1550, were transformed with the plasmid pCAM*gfp* (kindly provided by A. Sesma, John Innes Center, UK) (Sesma and Osbourn, 2004). The *pCAMgfp* plasmid contains the *sgfp* gene (Chiu *et al.*, 1996) under the control of the *ToxA* promoter from *Pyrenophora tritici-repens* (Lorang *et al.*, 2001) and the hygromycin phosphotransferase (*hph*) gene as the selectable marker gene . The pCAM*gfp* plasmid was introduced into the *Agrobacterium tumefaciens* AGL-1 strain, the virulent strain required for fungal transformation.

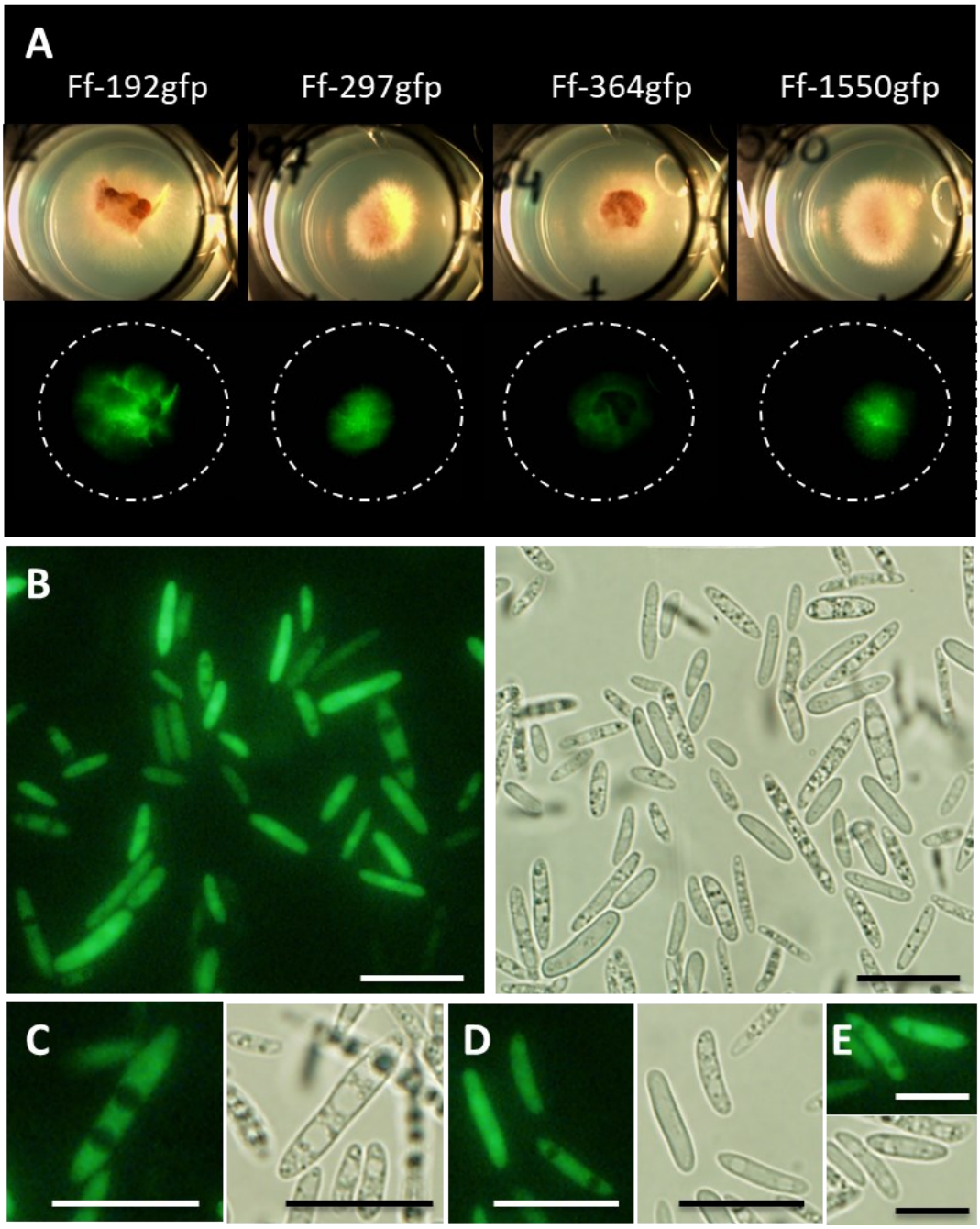

Figure 4.1: Morphological characteristics of transformed isolates of *F. fujikuroi* (A) Typical growth of gfp-expressing *F. fujikuroi* isolates. (B) Transmission and fluorescent micrographs of the gfp-expressing *F. fujikuroi* spores (strain 192). Similar results were obtained for the other F. fujikuroi strains. Bars, 20 µm (B, C and D) and 10 µm (E).

F. fujikuroi transformation was carried out using the *A. tumefaciens* AGL-1-transformed strain following the protocol described by Campos-Soriano & San Segundo (2009). Co-cultivation was performed at 25˚C and selection was done at 28˚C. Potato Dextrose Agar (PDA) medium with hygromycin B (250 µg ml^{-1} final concentration) was used as selective medium to grow the *F. fujikuroi* transformed isolates (**Supplementary figure 4.1**). Fungal colonies were transferred to 24-well plates containing selection medium (PDA medium plus hygromycin B) (**Figure 4.1 A**). A stereomicroscope (Olympus SZX16) with 480-nm excitation and 500 to 550-nm emission filter block was used to identify GFP-transformed fungal colonies. The number of pCAM*gfp* copies integrated into the genome of transformants was assessed by qPCR, using the primers Hyg588U and Hyg588L (**Supplementary table 4.2**).

Root infection assay

Seeds of one susceptible (Galileo) and one resistant (Selenio) rice cultivar were surface sterilized in 2% NaOCl for 2 min and rinsed in sterile H_2O before plating on sterile wet paper for germination. After 5 days at 30°C in the dark, young emerged roots were inoculated by applying 100 ul spore suspension at 10^6 spores ml^{-1} in the middle of the tissue. Both wild type isolates and GFP-derived transformants were inoculated as described, then seedlings were allowed to grow from 21 hours to 9 days, before confocal laser scanning microscopic (CLSM) analysis.

Epi-fluorescence microscopic analysis

GFP-labelled *F. fujikuroi* mycelium and spores, grown on PDA plates or inoculated roots, were photographed using an epifluorescence microscope equipped with a GFP filter and a camera to capture images of GFP fluorescence (excitation at 455 to 490 nm and emission at 515 to 560 nm).

Confocal microscopy of infected roots

Infected rice roots were stained with propidium iodide (0.2 μg ml^{-1}) for 3 min before microscopical observation. At different times after inoculation, unaltered and hand-sectioned roots were analysed. Images of GFP-labelled *F. fujikuroi* strains in host roots were captured using a confocal laser scanning microscope FV1000 Olympus (Tokyo, Japan) equipped with inverted microscope IX 81. Images were acquired in z stack with objective 10x (N.A. 0,40), using 488nm (argon Ion, emission 520nm) for GFP fluorescence, and 543nm (HeNe, emission 570 nm) laser for Propidium Iodide staining of root bark. Subsequently, they were processed using Imaris 6.2.1 software (Bitplane, Switzerland).

Gene expression analyses

The QIAGEN "RNeasy" kit was used to extract RNA from 0.1 g of root at 72 hours since inoculation. The extracted RNA was quantified by Nanodrop (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and purified using the TURBO DNA-free kit (Ambion, Foster City, California, United States). The samples were then checked for DNA contamination by PCR. The gene used was the rice elongation factor 1-alpha (Manosalva *et al.*, 2009). After verifying the samples purity, the RNA was used to obtain cDNA, using the iScript cDNA synthesis kit (Biorad, Hercules, California, United States). The samples were then used in real time qPCR (Applied Biosystem StepOnePlus, Foster City, California, United States), with primers for the rice genes *peroxidase P7*, *chitinase 1* and *gibberellin 20 oxidase 1* (*Gib20ox1*). The PCR mix were composed of 5 μl of Applied Biosystem SYBR Green Power Mix, 2 μl of cDNA, 0.15 μl of each primer (10 μM) and 2.4 μl of nuclease free water. The thermal cycler protocol was the following: 95°C for 10 min, followed by 40 cycles (95°C for 15 s; 60°C for 60 s) and 95°C for 15 s. The rice elongation factor 1-alpha (Manosalva *et al.*, 2009) was used as

housekeeping. The sequences of the used primers are reported in supplementary table 2. The efficiency of the primers was tested with a standard curve built upon five serial dilutions (1:10).

RESULTS AND DISCUSSION

Expression of the *gfp* gene in *Fusarium fujikuroi*

Four different *F. fujikuroi* isolates were transformed with the plasmid pCAM*gfp* (Sesma and Osbourn, 2004) containing the *sgfp* gene (Chiu *et al.*, 1996). The transformed isolates retained the colony morphology typical of the wild-type isolate indicating that *gfp* expression did not affect the parental phenotype (Fig 1A). Approximately 80-85% of the transformants showed strong fluorescent signal. Furthermore, strong fluorescence could be visualized in fungal spores and mycelium (Figure 1).

The stability of transformants was assessed. For this, some transformants were grew onto PDA medium without antibiotic and sub-cultured for 5 generations, before transferring again on PDA medium plus hygromycin B. The fluorescence of GFP in transformed *F. fujikuroi* strains remained stable through successive transfers. The number of pCAM*gfp* copies integrated into the transformant genomes varied from 1 to 2, Ff 297-GFP had only one copy.

All GFP isolates tested by infection of rice seeds were pathogenic, but only for Ff 297-GFP tagged isolate virulence was comparable to that of the recipient strain (**Figure 2**), while in the other transformants virulence was severely reduced, suggesting that, in these isolates, *sgfp* gene insertion affected a locus involved in the virulence trait. For all these features, only Ff 297-GFP isolate has been considered for microscopic analyses.

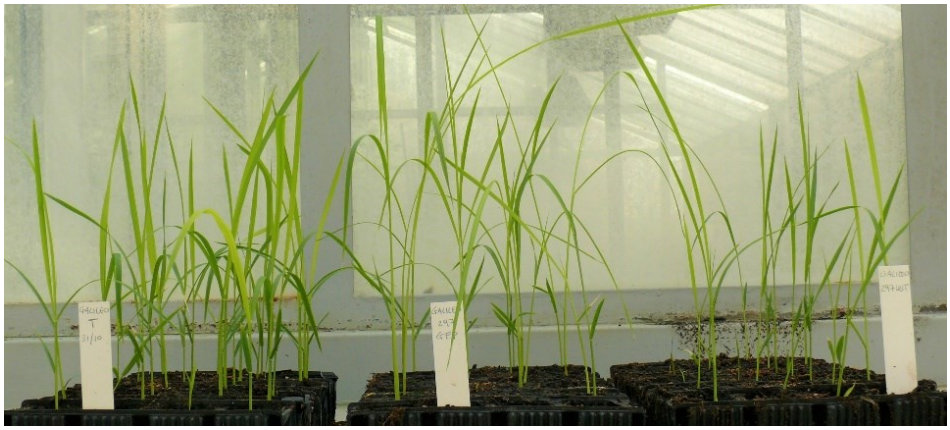

Figure 4.2: phenotype of the rice susceptible variety Galileo infected with the transformed (in the middle) and untransformed (on the right) *F. fujikuroi* isolate Ff 297. T, in the left, represents the not infected control.

Fluorescent behaviour of GFP-marked *F. fujikuroi* in susceptible and resistant cultivars

The GFP-tagged *F. fujikuroi* Ff 297 was used to inoculate two rice varieties and visualize the growth of this pathogen in the host infected rootlets. The infection of roots of the highly susceptible cv Galileo and the resistant cv Selenio was observed by confocal microscopy analysis starting from 48 hours (hpi) up to 8 days (dpi) post infection. Hyphae were predominantly observed along the root surface and among the root hairs and, after 48 h, the penetration site appeared green, because of penetration of the epidermal root cells by the pathogen (**Figure 4.3**).

Confocal imaging of radial sections of the roots showed that the fungus penetrated the stele in both varieties, and was more abundant in the susceptible variety than it was in the resistant one. Microscopical observations of infected roots began 48 h after inoculation with gfp-*F. fujikuroi* mycelium and continued up to 8 days after inoculation. In non-hairy regions, hyphae preferentially grew along the root surface.

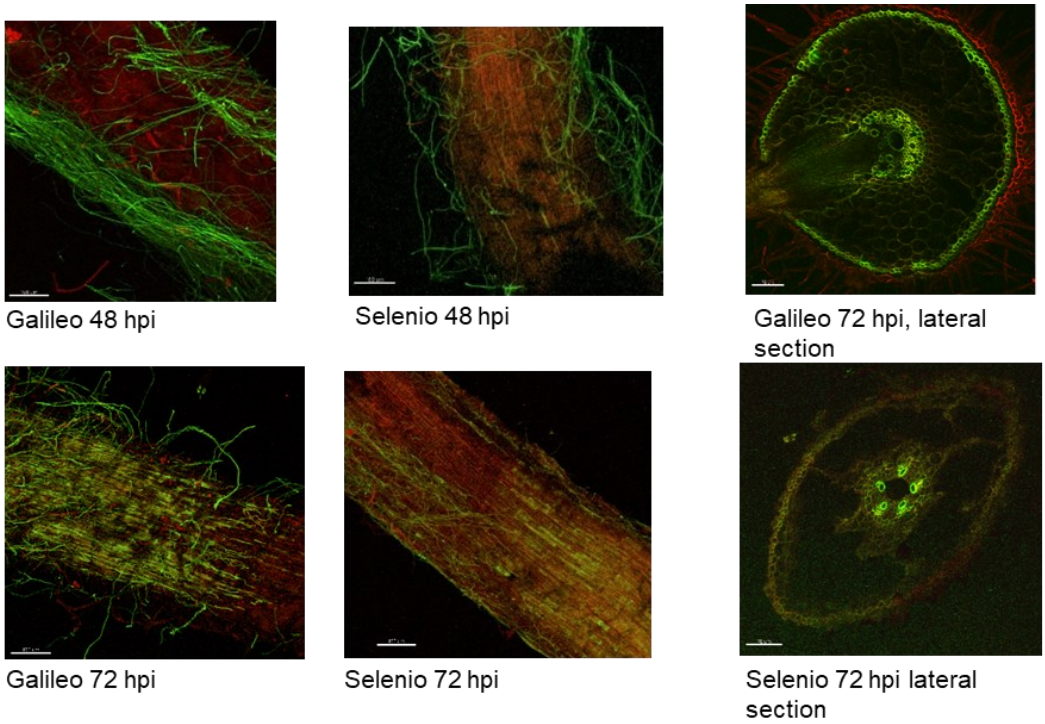

Figure 4.3: Rootlets of rice cv. Galileo (susceptible) and Selenio (resistant), inoculated with the transformed *F. fujikuroi* isolate Ff 297. Lateral and longitudinal sections are shown, at various time points after inoculation.

At 72 hours after inoculation, cell invasion was detected in a large number of epidermal cells scattered along the root. Mycelia extensively colonized the root tissues. The hyphal network showing high levels of gfp expression tightly embraced most of the root surface 7 days after inoculation. Contiguous cortical cells invaded by hyphae, forming green fluorescent lines along the root, were frequently observed, even in regions in which external hyphae were not present. Once inside the cortex, the fungus was highly invasive. Thus, the fungus progressed into the cortical cells and colonized the central cylinder (transverse sections). A close inspection of CLSM serial sections of the infected rice roots revealed colonization of the xylem vessels.

Eight days after inoculation, the cortex was almost completely colonized by the fungus.

Fungal growth happened also in the resistant cultivar Selenio, even if the fungal presence was less abundant than in the susceptible variety Galileo (**Figure 4.3**). This suggests that *Fusarium fujikuroi* can colonize the tissues of both susceptible and resistant plants, even if it does not cause symptoms in the former, as proposed by Matić and colleagues (2016). It is possible that rice resistance to *F. fujikuroi* is based on stopping the pathogen from entering specific areas of the root, while letting it colonise other parts of the tissue. A similar mechanism seems to be the base for lettuce resistant to *Verticillium dahlia*e (Vallad and Subbarao, 2008). This is in contrast with what can be observed in rice cultivars resistant to *Magnaporthe oryzae*, whose resistance is based on stopping the pathogen from penetrating the host tissues (Campos-Soriano *et al*., 2013).

Gene expression analyses

Chitinase1, *Gib20ox1* and *peroxidase P7* are rice genes previously observed to be differentially expressed in rice plants of cv. Selenio (resistant) and Dorella (susceptible) inoculated with *F. fujikuroi* (Matić *et al.* 2016). Their expression was quantified in this study at 72 hours after inoculation in order to compare the response to *F. fujikuroi* of young rootlets to that of adult plants. At this time point the differences between the susceptible and the resistant cultivar were already evident with confocal microscopy (**Figure 4.3**).

Chitinases are proteins involved in the plant defence against pathogens because of their ability to hydrolise chitin in the cell wall of fungi (Sharma *et al*., 2011). In this work, we observed that Selenio and Galileo expressed *chitinase1* at similar levels in the not inoculated samples, but the presence of this transcript increased greatly after inoculation in the resistant cultivar. Conversely, in the susceptible cultivar *chitinase1* expression did not change significantly (**Figure 4.4A**). It is possible that

the high expression of this chitinase is necessary to keep the pathogen outside of key root tissues, even if it does not stop the fungus from colonising other areas (**Figure 4.3**). On the contrary, Matić and colleagues (2016) observed that this gene was more expressed in the inoculated plants of the susceptible cultivar, at both 1 and 3 weeks since germination. Maybe a greater expression of this gene in the first phases of the infection, as we observed in the resistant cultivar, makes it unnecessary for the plant to maintain a high production of the protein in the following weeks. Another possible explanation is that this gene acts differently in the roots in respect to the green parts of the plant, which were the tissues analysed by Matić and colleagues (2016).

Gibberellins are involved in the regulation of bolting, germination and elongation of roots and internodes (Hedden and Sponsel, 2015) by inducing relaxation of cell wall (Cosgrove and Sovonick-Dunford, 1989). Many members of the gibberellin biosynthetic cluster are overexpressed in susceptible rice plants inoculated with *F. fujikuroi* (Matić *et al.*, 2016), and their production is the main cause for the bakanae-like symptoms caused by the fungus (Niehaus *et al*., 2017). However, due to an extremely high variability among replicates, there were no significant differences in *Gib20ox1* expression among the samples. This is probably caused by the fact that, at the time of the RNA extraction, some plants were starting to develop shoots, while others were not. It is known that gibberellins are produced during seedling growth (Hedden and Thomas, 2012), so the high variability can be explained by the fact that plants with developing shoots and internodes naturally overexpress the gibberellin biosynthetic gene cluster. It is however worth noticing how the average *Gib20ox1* expression was much lower in inoculated rootlets than in healthy ones (**Figure 4.4B**), suggesting an inhibitory effect of *F. fujikuroi* on gibberellin production at this early stage of the disease.

Peroxidases are enzymes involved in stopping pathogens from spreading through the tissues of the host, and they perform their function by producing high quantities

of reactive oxygen and nitrogen species, or by developing and reinforcing structural barriers (Passardi *et al*., 2005). However, at this early stage of the infection there was no significant difference in the expression of *peroxidase P7* in the analysed samples (**Figure 4.4C**). It is possible that this gene is not involved in rice defence against pathogens when the plant is so young, or that its role is played out in the green parts of the plant but not in the roots.

None of the considered genes showed an expression profile similar to that observed in the adult plants by Matić and collaborators (2016), suggesting that the rice genes involved in the resistance to *F. fujikuroi* at the early stages of host-pathogen interaction are different from the ones active in the adult plants.

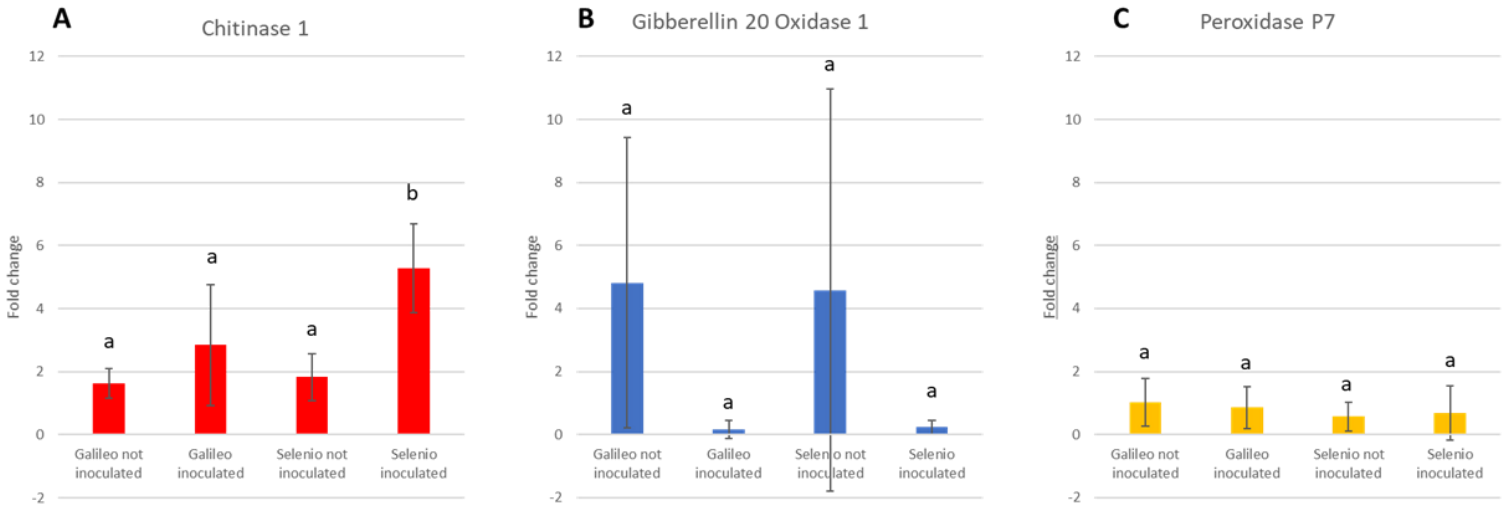

Figure 4.4: gene expression of genes *Chitinase 1* (**A**), *Gibberellin 20 oxidase 1* (**B**) and *Peroxidase P7* (**C**). Data obtained by reverse transcriptase real time PCR. The error bar is the standard deviation, and the letters indicate groups not considered to be statistically different using the Duncan test.

CONCLUSIONS

In this study a virulent GFP-transformed *F. fujikuroi* strain was obtained, enabling us to study the early stages of *F. fujikuroi* infection of rice roots in resistant and susceptible cultivars. We verified that the fungus spreads even in the roots of resistant plants, although its growth there is less abundant than in the susceptible variety. This suggests that *F. fujikuroi* is able to survive and grow inside its hosts even when not causing symptoms. At the same time, gene expression analyses revealed that the transcriptomic response of the plant to *F. fujikuroi* is likely to be different in the early and later stages of the disease. Until now, transcriptomics and proteomics studies on the rice-*F. fujikuroi* interaction have been carried out at several weeks after inoculation (Ji *et al*., 2016; Ji *et al*., 2019; Matić *et al*., 2016), but similar works conducted at earlier time periods may provide the scientific community with interesting new data concerning the genes involved in the resistance to *F. fujikuroi* showed by cultivars such as Selenio.

SUPPLEMENTARIES

Supplementary figure 4.1. *F. fujikuroi* strains used in this work.

Supplementary table 4.1. *F. fujikuroi* strains used in this study.

Supplementary table 4.2. Primers used in this study.

All the supplementary material is available online at: https://drive.google.com/drive/folders/1MAF2Vag3B5sAgS9BPtPI4AOFZWItRAkB?usp=sharing

5: DISCUSSION

F. fujikuroi is a complex pathogen, inducing different symptoms depending on both genotype (Niehaus *et al*., 2017) and temperature (Piombo *et al*., 2019), with small genomic differences greatly affecting the phenotype of the strains. Its effect on plants strongly depend on the cultivar and some susceptible and resistant rice varieties have been identified. Nevertheless, pathogen strains can survive inside resistant cultivars, even when they do not cause evident symptoms (Matić *et al*., 2016). Many studies have been conducted to identify genomic regions related to resistance, but the molecular basis of resistance to bakanae in rice is still unknown (Chen *et* al., 2019; Ji *et* al., 2018; Kang *et* al., 2018; Lee *et al.*, 2018; Lee *et al.*, 2019; Volante *et al*., 2017). Ji and colleagues (2016) and Matić and colleagues (2016) have compared the transcriptomic response to *F. fujikuroi* of susceptible and resistant rice cultivars, identifying MAP kinases and WRKY transcriptional factors upregulated in the 93-11 and Selenio resistant cultivars, but due to complex post-translational mechanisms only around 2% of differentially expressed genes maintain their up-regulation at the protein level in rice (Ji *et al*., 2019). Additionally, the main bakanae symptoms are deeply related to the action of gibberellins, which the pathogen can produce both directly and by inducing the overexpression of their cluster in the plant, making it difficult to properly study the disease with omics technologies unless looking at both sides of the pathosystem.

However, comparative genomics is an adequate tool for investigating the pathosystem and finding genes related to the typology of induced symptoms or the ability to cause the disease in the first place. In the second chapter of this thesis, we identified through comparative genomics a pisatin demethylase, an integral membrane protein and a fructosyl amino acid oxidase that could be involved in the shift from elongation-inducing to stunting-inducing phenotype. Furthermore, in chapter 3 the comparison with an avirulent strain and a scarcely virulent one, never sequenced before, permitted the identification of 13 genes putatively affecting the

pathogen ability to cause the disease (**Table 3.2**). These included genes involved in degradation of the plant cell wall, cellulose-response, degradation of cellulose, conidia development, growth in anaerobic environment, utilization of "difficult" carbon sources, resistance to stresses, appressorium formation and repression of AreA-dependent genes. Five of these genes are involved in either degradation or response to some of the components of plant cell wall or in appressorium formation, suggesting that the early stages of the pathogen-host interaction are fundamental for the development of the disease. One of these genes resulted overexpressed in a virulent strain after 48 hours of growth in the presence of rice seeds, while this did not happen in the considered avirulent strain. This piece of data correlates well with the results presented in chapter 4, where differences between fungal growth in susceptible and resistant cultivars are observed at a few hours after inoculation with *F. fujikuroi* strain 297. Until now, transcriptomics and proteomics studies have focused on the response in the aerial part of the plant, at several weeks after inoculation (Ji *et al*., 2016; Ji *et al*., 2019; Matić *et al*., 2016). This time point is commonly used in pathogenicity trials to establish the virulence of *F. fujikuroi* strains, but it might not be the best one for studying gene expression and plant defence. Therefore, omics studies at earlier stages of the plant-pathogen interaction are necessary to identify the genes involved in the mechanisms of resistance (invisible at later stages) that allow some cultivars, like Selenio, to cope with the pathogen in a better way than others. Of course, the early interactions happen in seeds and rootlets, making RNA and DNA extractions problematic. There are, however, a few available protocols that can provide good nucleic acid from seed tissues (Botelho et al., 2016; Sangha et al., 2010; Suzuki et al., 2008; Wang et al., 2012). The literature about *F. fujikuroi* is currently also lacking in epigenetic studies, which could be fundamental to further clarify the pathogenesis mechanism of this fungus. In *Magnaporthe oryzae*, another rice pathogen, histone methylation activates the gene MoCel7C, a cellulase upregulated upon contact with the host (Van

Vu *et al*., 2013), and the expression of effector genes can be regulated by methylation as well, as proven in *Leptosphaeria maculans* (Soyer *et al*., 2014). Even secondary metabolism is influenced by epigenetics, for example the production of fumonisins in *F. verticilloides* is controlled by histone acetylation-mediated epigenetic modification (Visentin *et al*., 2012), and it is known that many fungal pathogens depend on epigenetic regulation for some aspects of their interaction with the host (Dubey and Jeon, 2017). This type of regulation could be important in the development of bakanae disease, and it should be investigated in the future.

Avirulent *F. fujikuroi* strains are rare, with Amatulli and colleagues (2010) finding only 9 in a study that evaluated the virulence of 75 ones. However, they have a potential in the application for biocontrol of bakanae disease by seed dressing. Treating rice seeds with thermotherapy and antagonistic yeasts, such as *Pichia guillermondii* or *Metschnikowia pulcherrima*, can control bakanae disease and increase the germination rate (Matić *et al*., 2014), but there are currently no scientific works studying the efficacy of avirulent *F. fujikuroi* strains in biocontrol. The completely avirulent strain SG4, sequenced in chapter 3, is a good candidate for such a study.

Lastly, another important result of the thesis has been to show, in chapter 2, that many *F. fujikuroi* strains are more virulent at higher temperatures. While the effect on rice of the abiotic stresses caused by rising temperatures are being investigated (Mohammed and Tarpley, 2009), and genetically modified climate-resilient rice cultivars are being developed (Sreenivasulu *et al.*, 2015), there are only a few studies regarding the effect of global warming on the severity of fungal diseases on rice (Onaga *et al*., 2017), despite the consistent effect that pathogens could have in the production of staple food for the future years.

CONCLUSIONS

The studies presented in this thesis have contributed to advance the current knowledge on *F. fujikuroi*, identifying genes probably involved in the host-pathogen interaction with rice. Furthermore, the sequencing of the strain SG4 poses the basis for the utilization of avirulent *F. fujikuroi* in the biocontrol of bakanae disease.

The work with GFP-transformed rice provides the scientific community with useful information for future studies as well. In particular, it highlights how transcriptomic and proteomic studies concerning the resistance to the disease should be conducted during the first stages of the pathogen-host interaction.

Finally, the thesis draws attention to the danger that global warming poses to rice cultivation around the world, demonstrating that the virulence of this pathogen tends to increase at higher temperatures.

REFERENCES

- Ádám, A. L., García-Martínez, J., Szűcs, E. P., Avalos, J., & Hornok, L. (2011). The MAT1-2-1 mating-type gene upregulates photo-inducible carotenoid biosynthesis in Fusarium verticillioides. *FEMS microbiology letters*, *318*(1), 76-83. https://doi.org/10.1111/j.1574-6968.2011.02241.x
- Adorisio, S., Fierabracci, A., Muscari, I., Liberati, A. M., Cannarile, L., Thuy, T. T., ... & Riccardi, C. (2019). Fusarubin and Anhydrofusarubin Isolated from A Cladosporium Species Inhibit Cell Growth in Human Cancer Cell Lines. *Toxins*, *11*(9), 503. https://doi.org/10.3390/toxins11090503
- Alexander, N. J., Proctor, R. H., & McCormick, S. P. (2009). Genes, gene clusters, and biosynthesis of trichothecenes and fumonisins in Fusarium. *Toxin Reviews*, *28*(2-3), 198-215. https://doi.org/10.1080/15569540903092142
- Alkhayyat, F., & Yu, J. H. (2014). Upstream regulation of mycotoxin biosynthesis. In *Advances in Applied Microbiology* (Vol. 86, pp. 251-278). Academic Press. https://doi.org/10.1016/B978-0-12-800262-9.00005-6
- Amatulli, M. T., Spadaro, D., Gullino, M. L., & Garibaldi, A. (2010). Molecular identification of Fusarium spp. associated with bakanae disease of rice in Italy and assessment of their pathogenicity. *Plant Pathology*, *59*(5), 839-844. https://doi.org/10.1111/j.1365-3059.2010.02319.x
- Amoah, B. K., Rezanoor, H. N., Nicholson, P., & MacDonald, M. V. (1995). Variation in the Fusarium section Liseola: pathogenicity and genetic studies of isolates of *Fusarium moniliforme* Sheldon from different hosts in Ghana. *Plant Pathology*, *44*(3), 563-572. https://doi.org/10.1111/j.1365-3059.1995.tb01678.x
- Avalos, J., Prado-Cabrero, A., & Estrada, A. F. (2012). Neurosporaxanthin production by Neurospora and Fusarium. *Microbial Carotenoids From Fungi*, 263-274. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-61779-918-1_18
- Avalos, J., Pardo-Medina, J., Parra-Rivero, O., Ruger-Herreros, M., Rodríguez-Ortiz, R., Hornero-Méndez, D., & Limón, M. C. (2017). Carotenoid biosynthesis in Fusarium. *Journal of Fungi*, *3*(3), 39. https://doi.org/10.3390/jof3030039
- Bacon, C. W., Porter, J. K., Norred, W. P., & Leslie, J. F. (1996). Production of fusaric acid by Fusarium species. *Applied and Environmental Microbiology*, *62*(11), 4039-4043.

- BAGGA, P. S., & KUMAR, V. (2000). Resistance to bakanae or foot-rot disease in basmati rice. *Indian Phytopathology*, *53*(3), 321-322.
- Bailey, L. A., & Ebbole, D. J. (1998). The fluffy gene of Neurospora crassa encodes a Gal4p-type C6 zinc cluster protein required for conidial development. *Genetics*, *148*(4), 1813-1820.
- Balan, J., Fuska, J., Kuhr, I., & Kuhrova, V. (1970). Bikaverin, an antibiotic from *Gibberella fujikuroi*, effective against *Leishmania brasiliensis*. *Folia Microbiologica*, *15*(6), 479-484.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pyshkin, A. V. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455-477. https://doi.org/10.1089/cmb.2012.0021
- Barrero, A. F., Sánchez, J. F., Oltra, J. E., Tamayo, N., Cerdá-Olmedo, E., Candau, R., & Avalos, J. (1991). Fusarin C and 8Z-fusarin C from *Gibberella fujikuroi*. *Phytochemistry*, *30*(7), 2259-2263. https://doi.org/10.1016/0031-9422(91)83625-U
- Bashyal, B. M., Aggarwal, R., Sharma, S., Gupta, S., Rawat, K., Singh, D., ... & Krishnan, S. G. (2016a). Occurrence, identification and pathogenicity of Fusarium species associated with bakanae disease of basmati rice in India. *European Journal of Plant Pathology*, *144*(2), 457-466. https://doi.org/10.1007/s10658-015-0783-8
- Bashyal, B. M., Aggarwal, R., Sharma, S., Gupta, S., & Singh, U. B. (2016b). Single and combined effects of three Fusarium species associated with rice seeds on the severity of bakanae disease of rice. *Journal of Plant Pathology*, 405-412.
- Bashyal, B. M., Rawat, K., Sharma, S., Kulshreshtha, D., Gopala Krishnan, S., Singh, A. K., ... & Aggarwal, R. (2017). Whole genome sequencing of *Fusarium fujikuroi* provides insight into the role of secretory proteins and cell wall degrading enzymes in causing bakanae disease of rice. *Frontiers in Plant Science*, *8*, 2013. https://doi.org/10.3389/fpls.2017.02013
- Bauer, S., Vasu, P., Persson, S., Mort, A. J., & Somerville, C. R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analysing plant cell walls. *Proceedings of the National Academy of Sciences*, *103*(30), 11417-11422. https://doi.org/10.1073/pnas.0604632103

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170
- Bolger, M., Coker, R. D., Dinovi, M., Gaylor, D., Gelderblom, M. O., Paster, N., ... & Speijers, J. A. (2001). Fumonisins. In Safety Evaluation of Certain Mycotoxins in Food, Food and Agriculture Organization of the United Nations. *World Health Organization. Geneva: Food Additives, Series*, *47*, 103-297.
- Bömke, C., & Tudzynski, B. (2009). Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. *Phytochemistry*, *70*(15-16), 1876-1893. https://doi.org/10.1016/j.phytochem.2009.05.020
- Boonman, N., Prachya, S., Boonmee, A., Kittakoop, P., Wiyakrutta, S., Sriubolmas, N., ... & Chusattayanond, A. D. A. (2012). *In vitro* acanthamoebicidal activity of fusaric acid and dehydrofusaric acid from an endophytic fungus *Fusarium sp.* Tlau3. *Planta medica*, *78*(14), 1562-1567. https://doi.org/10.1055/s-0032-1315146
- Botelho, S. R., Martins, T. P., Duarte, M. F., Barbosa, A. V., Lau, D., Fernandes, F. R., & Sanches, M. M. (2016). Development of methodologies for virus detection in soybean and wheat seeds. *MethodsX*, *3*, 62-68. https://doi.org/10.1016/j.mex.2016.01.005
- Boudreau, B. A., Larson, T. M., Brown, D. W., Busman, M., Roberts, E. S., Kendra, D. F., & McQuade, K. L. (2013). Impact of temperature stress and validamycin A on compatible solutes and fumonisin production in *F. verticillioides*: role of trehalose-6-phosphate synthase. *Fungal Genetics and Biology*, *57*, 1-10. https://doi.org/10.1016/j.fgb.2013.06.001
- Brown, D. W., Lee, S. H., Kim, L. H., Ryu, J. G., Lee, S., Seo, Y., ... & Lee, T. (2015). Identification of a 12-gene fusaric acid biosynthetic gene cluster in Fusarium species through comparative and functional genomics. *Molecular Plant-Microbe Interactions*, *28*(3), 319-332. https://doi.org/10.1094/MPMI-09-14-0264-R
- Butchko, R. A., Brown, D. W., Busman, M., Tudzynski, B., & Wiemann, P. (2012). Lae1 regulates expression of multiple secondary metabolite gene clusters in *Fusarium verticillioides*. *Fungal Genetics and Biology*, *49*(8), 602-612. https://doi.org/10.1016/j.fgb.2012.06.003

- Cambaza, E. (2018). Comprehensive description of Fusarium graminearum pigments and related compounds. *Foods*, *7*(10), 165. https://doi.org/10.3390/foods7100165
- Campos-Soriano, L. & San Segundo, B. (2009) Assessment of blast disease resistance in transgenic PRms rice using a gfp-expressing *Magnaporthe oryzae* strain. *Plant Pathology,* 58, 677–689. https://doi.org/10.1111/j.1365-3059.2009.02061.x
- Campos-Soriano, L., Valè, G., Lupotto, E., & San Segundo, B. (2013). Investigation of rice blast development in susceptible and resistant rice cultivars using a gfp-expressing Magnaporthe oryzae isolate. *Plant Pathology*, *62*(5), 1030-1037. https://doi.org/10.1111/ppa.12023
- Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., ... & Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188-196. https://doi.org/10.1101/gr.6743907
- Carter, L. L. A., Leslie, J. F., & Webster, R. K. (2008). Population structure of *Fusarium fujikuroi* from California rice and water grass. *Phytopathology*, *98*(9), 992-998. https://doi.org/10.1094/PHYTO-98-9-0992
- Chiara, M., Fanelli, F., Mule, G., Logrieco, A. F., Pesole, G., Leslie, J. F., ... & Toomajian, C. (2015). Genome sequencing of multiple isolates highlights subtelomeric genomic diversity within *Fusarium fujikuroi*. *Genome Biology and Evolution*, *7*(11), 3062-3069. https://doi.org/10.1093/gbe/evv198
- Chen, S. Y., Lai, M. H., Tung, C. W., Wu, D. H., Chang, F. Y., Lin, T. C., & Chung, C. L. (2019). Genome-wide association mapping of gene loci affecting disease resistance in the rice-*Fusarium fujikuroi* pathosystem. *Rice*, *12*(1), 85. https://doi.org/10.1186/s12284-019-0337-3
- Chen, Y. C., Lai, M. H., Wu, C. Y., Lin, T. C., Cheng, A. H., Yang, C. C., ... & Lin, G. C. (2016). The genetic structure, virulence, and fungicide sensitivity of *Fusarium fujikuroi* in Taiwan. *Phytopathology*, *106*(6), 624-635. https://doi.org/10.1094/PHYTO-11-15-0285-R
- Chiara, M., Fanelli, F., Mule, G., Logrieco, A. F., Pesole, G., Leslie, J. F., ... & Toomajian, C. (2015). Genome sequencing of multiple isolates highlights subtelomeric genomic diversity within *Fusarium fujikuroi*. *Genome biology and evolution*, *7*(11), 3062-3069. https://doi.org/10.1093/gbe/evv198

- Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. & Sheen, J. (1996) Engineered GFP as a vital reporter in plants. *Current Biology, 6*(3), 325–330. https://doi.org/10.1016/S0960-9822(02)00483-9
- Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745-2747. https://doi.org/10.1093/bioinformatics/btv195
- Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., and Chan, A. P. 2012. Predicting the functional effect of amino acid substitutions and indels. PloS one. 7:e46688. https://doi.org/10.1371/journal.pone.0046688
- Chu, F. S., & Li, G. Y. (1994). Simultaneous occurrence of fumonisin B1 and other mycotoxins in moldy corn collected from the People's Republic of China in regions with high incidences of esophageal cancer. *Applied and Environmental Microbioogy*, *60*(3), 847-852.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80-92. https://doi.org/10.4161/fly.19695
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610
- Cornforth, J. W., Ryback, G., Robinson, P. M., & Park, D. (1971). Isolation and characterization of a fungal vacuolation factor (bikaverin). *Journal of the Chemical Society C: Organic*, 2786-2788. https://doi.org/10.1039/J39710002786
- Cosgrove, D. J., & Sovonick-Dunford, S. A. (1989). Mechanism of gibberellin-dependent stem elongation in peas. *Plant Physiology*, *89*(1), 184-191. https://doi.org/10.1104/pp.89.1.184
- Creppy, E. E. (2002). Update of survey, regulation and toxic effects of mycotoxins in Europe. *Toxicology Letters*, *127*(1-3), 19-28. https://doi.org/10.1016/S0378-4274(01)00479-9
- Cruz, A., Marín, P., González-Jaén, M. T., Aguilar, K. G. I., & Cumagun, C. J. R. (2013). Phylogenetic analysis, fumonisin production and pathogenicity of *Fusarium*

fujikuroi strains isolated from rice in the Philippines. *Journal of the Science of Food and Agriculture*, *93*(12), 3032-3039. https://doi.org/10.1002/jsfa.6136

- Dendooven, L., Magaña, I. P., & De la Torre, M. (2000). Optimization of gibberellic acid production by immobilized *Gibberella fujikuroi* mycelium in fluidized bioreactors. *Journal of Biotechnology*, *76*(2-3), 147-155. https://doi.org/10.1016/S0168-1656(99)00182-0
- Desjardins, A. E., Manandhar, H. K., Plattner, R. D., Manandhar, G. G., Poling, S. M., & Maragos, C. M. (2000). Fusarium species from Nepalese rice and production of mycotoxins and gibberellic acid by selected species. *Applied and Environmental Microbioogy*, *66*(3), 1020-1025. https://doi.org/10.1128/AEM.66.3.1020-1025.2000
- Desjardins, A. E., Plattner, R. D., & Nelson, P. E. (1997). Production of Fumonisin B (inf1) and Moniliformin by *Gibberella fujikuroi* from Rice from Various Geographic Areas. *Applied and Environmental Microbioogy*, *63*(5), 1838-1842.
- De Vries, R. P., Kester, H. C., Poulsen, C. H., Benen, J. A., & Visser, J. (2000). Synergy between enzymes from Aspergillus involved in the degradation of plant cell wall polysaccharides. *Carbohydrate Research*, *327*(4), 401-410. https://doi.org/10.1016/S0008-6215(00)00066-5
- DeZwaan, T. M., Carroll, A. M., Valent, B., & Sweigard, J. A. (1999). *Magnaporthe grisea* pth11p is a novel plasma membrane protein that mediates appressorium differentiation in response to inductive substrate cues. *The Plant Cell*, *11*(10), 2013-2030. https://doi.org/10.1105/tpc.11.10.2013
- Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., ... & Feussner, K. (2011). Metabolic priming by a secreted fungal effector. *Nature*, *478*(7369), 395-398. https://doi.org/10.1038/nature10454
- Drori, E., Levy, D., Smirin-Yosef, P., Rahimi, O., & Salmon-Divon, M. (2017). CircosVCF: circos visualization of whole-genome sequence variations stored in VCF files. *Bioinformatics*, *33*(9), 1392-1393. https://doi.org/10.1093/bioinformatics/btw834
- Dubey, A., & Jeon, J. (2017). Epigenetic regulation of development and pathogenesis in fungal plant pathogens. *Molecular plant pathology*, *18*(6), 887-898. https://doi.org/10.1111/mpp.12499
- Eisenhaber, B., Schneider, G., Wildpaner, M., & Eisenhaber, F. (2004). A sensitive predictor for potential GPI lipid modification sites in fungal protein sequences and its application to genome-wide studies for *Aspergillus nidulans, Candida albicans,*

Neurospora crassa, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Journal of Molecular Biology*, *337*(2), 243-253. https://doi.org/10.1016/j.jmb.2004.01.025

- Eisenreich, W., Rohdich, F., & Bacher, A. (2001). Deoxyxylulose phosphate pathway to terpenoids. *Trends in Plant Science*, *6*(2), 78-84. https://doi.org/10.1016/S1360-1385(00)01812-4
- Emanuelsson, O., Nielsen, H., Brunak, S., & Von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *Journal of Molecular Biology*, *300*(4), 1005-1016. https://doi.org/10.1006/jmbi.2000.3903
- Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology*, *16*(1), 157. https://doi.org/10.1186/s13059-015-0721-2
- Emms, D., & Kelly, S. (2018). STAG: species tree inference from all genes. *BioRxiv*, 267914. https://doi.org/10.1101/267914
- Fan, X., Abbott, T. E., Larson, D., & Chen, K. (2014). BreakDancer: Identification of genomic structural variation from paired-end read mapping. *Current Protocols in Bioinformatics*, *45*(1), 15-6. https://doi.org/10.1002/0471250953.bi1506s45
- Fiyaz, R. A., Yadav, A. K., Krishnan, S. G., Ellur, R. K., Bashyal, B. M., Grover, N., ... & Prabhu, K. V. (2016). Mapping quantitative trait loci responsible for resistance to Bakanae disease in rice. *Rice*, *9*(1), 45. https://doi.org/10.1186/s12284-016-0117-2.
- Flaherty, J. E., Pirttilä, A. M., Bluhm, B. H., & Woloshuk, C. P. (2003). PAC1, a pH-regulatory gene from *Fusarium verticillioides*. *Applied and Environmental Microbioogy*, *69*(9), 5222-5227.https://doi.org/10.1128/AEM.69.9.5222-5227.2003
- Flaherty, J. E., & Woloshuk, C. P. (2004). Regulation of fumonisin biosynthesis in *Fusarium verticillioides* by a zinc binuclear cluster-type gene, ZFR1. *Applied and Environmental Microbioogy*, *70*(5), 2653-2659. https://doi.org/10.1128/AEM.70.5.2653-2659.2004
- Fouché, S., Plissonneau, C., & Croll, D. (2018). The birth and death of effectors in rapidly evolving filamentous pathogen genomes. *Current Opinion in Microbiology*, *46*, 34-42. https://doi.org/10.1016/j.mib.2018.01.020

- Frandsen, R. J., Schütt, C., Lund, B. W., Staerk, D., Nielsen, J., Olsson, S., & Giese, H. (2011). Two novel classes of enzymes are required for the biosynthesis of aurofusarin in *Fusarium graminearum*. *Journal of Biological Chemistry*, *286*(12), 10419-10428. https://doi.org/10.1074/jbc.M110.179853
- Fuska, J., Proksa, B., & Fuskova, A. (1975). New potential cytotoxic and antitumor substances I. *In vitro* effect of bikaverin and its derivatives on cells of certain tumors. *Neoplasma*, *22*(3), 335-338.
- Garibaldi, A., Minuto, G., & Gullino, M. L. (1994). Biological control of Fusarium wilt of basil (*Ocimum basilicum* L.). *Brighton Crop Protection Conference - Pest and Disease 2,* 811-815.
- Gopalakrishnan, S., Beale, M. H., Ward, J. L., & Strange, R. N. (2005). Chickpea wilt: identification and toxicity of 8-O-methyl-fusarubin from *Fusarium acutatum*. *Phytochemistry*, *66*(13), 1536-1539. https://doi.org/10.1016/j.phytochem.2005.04.037
- Gupta, A. K., Singh, Y., Jain, A. K., & Singh, D. (2014). Prevalence and incidence of bakanae disease of rice in northern India. *Journal of AgriSearch*, *1*(4), 233-7.
- Gupta, A. K., Solanki, I. S., Bashyal, B. M., Singh, Y., & Srivastava, K. (2015). Bakanae of rice-an emerging disease in Asia. *Journal of Animal and Plant Sciences, 25*, 1499-1514.
- Gupta, A. K., Solanki, I. S., Bashyal, B. M., Singh, Y., & Srivastava, K. (2015). Bakanae of rice-an emerging disease in Asia. *Journal of Animal and Plant Sciences*, *25*, 1499-1514.
- Haidar, S., Aichele, D., Birus, R., Hielscher, J., Laitinen, T., Poso, A., & Jose, J. (2019). *In Vitro* and *In Silico* Evaluation of Bikaverin as a Potent Inhibitor of Human Protein Kinase CK2. *Molecules*, *24*(7), 1380. https://doi.org/10.3390/molecules24071380
- Harris, L. J., Balcerzak, M., Johnston, A., Schneiderman, D., & Ouellet, T. (2016). Host-preferential *Fusarium graminearum* gene expression during infection of wheat, barley, and maize. *Fungal Biology*, *120*(1), 111-123. https://doi.org/10.1016/j.funbio.2015.10.010
- Hedden, P., & Sponsel, V. (2015). A century of gibberellin research. *Journal of Plant Growth Regulation*, *34*(4), 740-760. https://doi.org/10.1007/s00344-015-9546-1

- Ito, S., Kimura, J. (1931). Studies on the 'bakanae' disease of the rice plant. *Rep Hokkaido Natl Agric Exp Stat, 27*, 1-95
- Janevska, S., & Tudzynski, B. (2018). Secondary metabolism in *Fusarium fujikuroi*: strategies to unravel the function of biosynthetic pathways. *Applied Microbiology and Biotechnology*, *102*(2), 615-630. https://doi.org/10.1007/s00253-017-8679-5
- Janvier, C., Villeneuve, F., Alabouvette, A., Edel-Hermann, V., Mateille, T., & Steinberg, C. (2007) Soil health through soil disease suppression: which strategy from descriptors to indicators? *Soil Biology and Biochemistry, 39*(1), 1–23. https://doi.org/10.1016/j.soilbio.2006.07.001
- Jeon, Y. A., Yu, S. H., Lee, Y. Y., Park, H. J., Lee, S., Sung, J. S., ... & Lee, H. S. (2013). Incidence, molecular characteristics and pathogenicity of *Gibberella fujikuroi* species complex associated with rice seeds from Asian countries. *Mycobiology*, *41*(4), 225-233. https://doi.org/10.5941/MYCO.2013.41.4.225
- Jeong, H., Lee, S., Choi, G. J., Lee, T., & Yun, S. H. (2013). Draft genome sequence of *Fusarium fujikuroi* B14, the causal agent of the bakanae disease of rice. *Genome Announcements.*, *1*(1), e00035-13. https://doi.org/10.1128/genomeA.00035-13
- Jeong, H. Y., Song, M., Back, J., Han, D. M., Wu, X., Monnier, V., ... & Chae, K. S. (2002). The veA gene is necessary for the inducible expression by fructosyl amines of the *Aspergillus nidulans* faoA gene encoding fructosyl amino acid oxidase (amadoriase, EC 1.5. 3). *Archives of Microbiology*, *178*(5), 344-350. https://doi.org/10.1007/s00203-002-0463-x
- Ji, H., Kim, T. H., Lee, G. S., Kang, H. J., Lee, S. B., Suh, S. C., ... & Kim, K. H. (2018). Mapping of a major quantitative trait locus for bakanae disease resistance in rice by genome resequencing. *Molecular Genetics and Genomics*, *293*(3), 579-586. https://doi.org/10.1007/s00438-017-1407-0
- Ji, Z., Zeng, Y., Liang, Y., Qian, Q., & Yang, C. (2016). Transcriptomic dissection of the rice–*Fusarium fujikuroi* interaction by RNA-Seq. *Euphytica*, *211*(1), 123-137. https://doi.org/10.1007/s10681-016-1748-5
- Ji, Z., Zeng, Y., Liang, Y., Qian, Q., & Yang, C. (2019). Proteomic dissection of the rice-*Fusarium fujikuroi* interaction and the correlation between the proteome and transcriptome under disease stress. *BMC Genomics*, *20*(1), 91. https://doi.org/10.1186/s12864-019-5435-5

- Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*, *428*(4), 726-731. https://doi.org/10.1016/j.jmb.2015.11.006
- Kang, D. Y., Cheon, K. S., Oh, J., Oh, H., Kim, S. L., Kim, N., ... & Chung, N. J. (2019). Rice Genome Resequencing Reveals a Major Quantitative Trait Locus for Resistance to Bakanae Disease Caused by *Fusarium fujikuroi*. *International Journal of Molecular Sciences*, *20*(10), 2598. https://doi.org/10.3390/ijms20102598
- Kanjanasoon, P. (1965). Studies on the bakanae disease of rice in Thailand. *Doc Agr Thesis, Tokyo University, Japan*.
- Karov, I., Mitrev, S., & Arsov, E. (2009). *Giberella fujikuroi* (Wollenweber) the new parasitical fungus on rice in the Republic of Macedonia. *Proceedings in Natural Science of the Matica Srpska Novi Sad*, (116), 175-182.
- Karov, I., Mitrev, S., Mihajlov, L., Ristova, D., Arsov, E., & Kovacevik, B. (2005). *Gibberella fujikuroi* (Sawada) Wollenweber, the new parasitical fungus on rice in region of Kocani. *Yearbook, Institute of Southern Crops Strumica*, 157-162.
- Kasahara, H., Hanada, A., Kuzuyama, T., Takagi, M., Kamiya, Y., & Yamaguchi, S. (2002). Contribution of the mevalonate and methylerythritol phosphate pathways to the biosynthesis of gibberellins in Arabidopsis. *Journal of Biological Chemistry*, *277*(47), 45188-45194. https://doi.org/10.1074/jbc.M208659200
- Kato, A., Miyake, T., Nishigata, K., Tateishi, H., Teraoka, T., & Arie, T. (2012). Use of fluorescent proteins to visualize interactions between the Bakanae disease pathogen *Gibberella fujikuroi* and the biocontrol agent Talaromyces sp. KNB-422. *Journal of general plant pathology*, *78*(1), 54-61. https://doi.org/10.1007/s10327-011-0343-9
- Khan, J. A., Jamil, F. F., & Gill, M. A. (2000). Screening of rice varieties/lines against bakanae and bacterial leaf blight (BLB). *Pakistan Journal of Phytopathology*, *12*(1), 6-11.
- Kim, B. R., Han, K. S., Hahm, S. S., Kwon, M. K., & Nam, Y. G. (2015). Occurrence of the rice bakanae disease in Chungnam province. *Research in Plant Disease*, *21*, 154.
- Kim, M. H., Lee, S. B., Kwon, T., Hwang, U. H., Park, S. K., Youn, Y. N., ... & Yeo, U. S. (2013). A large-scale screening analysis for the evaluation of Bakanae

disease in rice. *Journal of General Plant Pathology*, 41-41. https://doi.org/10.1007/s10327-014-0528-0

- Kim, S. Y., Park, D. S., Lee, I. J., Park, Y. G., Seo, J. H., Bae, H. K., ... & Ko, J. M. (2018). Disease incidence, growth, and physiological characterization of rice cultivars with different susceptibility to bakanae disease. *Philippine Journal of Crop Science (PJCS)*, *43*(2), 38-46.
- Khan, J. A., Jamil, F. F., & Gill, M. A. (2000). Screening of rice varieties/lines against bakanae and bacterial leaf blight (BLB). *Pakistan Journal of Phytopathology*, *12*(1), 6-11.
- Köller, W., Allan, C. R., & Kolattukudy, P. E. (1982). Role of cutinase and cell wall degrading enzymes in infection of *Pisum sativum* by *Fusarium solani f. sp. pisi*. *Physiological Plant Pathology*, *20*(1), 47-60. https://doi.org/10.1016/0048-4059(82)90023-6
- Krogh, A., Larsson, B., Von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology*, *305*(3), 567-580. https://doi.org/10.1006/jmbi.2000.4315
- Kumar, S., Tamura, K., & Nei, M. (1994). MEGA: molecular evolutionary genetics analysis software for microcomputers. *Bioinformatics*, *10*(2), 189-191. https://doi.org/10.1093/bioinformatics/10.2.189
- Kwon, H. R., Son, S. W., Han, H. R., Choi, G. J., Jang, K. S., Choi, Y. H., ... & Kim, J. C. (2007). Nematicidal activity of bikaverin and fusaric acid isolated from *Fusarium oxysporum* against pine wood nematode, *Bursaphelenchus xylophilus*. *The Plant Pathology Journal*, *23*(4), 318-321. https://doi.org/10.5423/PPJ.2007.23.4.318
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357. https://doi.org/10.1038/nmeth.1923
- Levin, E., Kishore, A., Ballester, A. R., Raphael, G., Feigenberg, O., Liu, Y., ... & Droby, S. (2019). Identification of pathogenicity-related genes and the role of a subtilisin-related peptidase S8 (PePRT) in authophagy and virulence of *Penicillium expansum* on apples. *Postharvest Biology and Technology*, *149*, 209-220. https://doi.org/10.1016/j.postharvbio.2018.10.011

- Lee, S. B., Hur, Y. J., Cho, J. H., Lee, J. H., Kim, T. H., Cho, S. M., ... & Park, Y. J. (2018). Molecular mapping of qBK1 WD, a major QTL for bakanae disease resistance in rice. *Rice*, *11*(1), 3. https://doi.org/10.1186/s12284-017-0197-7
- Lee, S. B., Kim, N., Hur, Y. J., Cho, S. M., Kim, T. H., Lee, J. Y., ... & Ko, J. M. (2019). Fine mapping of qBK1, a major QTL for bakanae disease resistance in rice. *Rice*, *12*(1), 36. https://doi.org/10.1186/s12284-019-0295-9
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: detection of (co-) orthologs in large-scale analysis. *BMC Bioinformatics*, *12*(1), 124. https://doi.org/10.1186/1471-2105-12-124
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754-1760. https://doi.org/10.1093/bioinformatics/btp324
- Li, M., Liu, X., Liu, Z., Sun, Y., Liu, M., Wang, X., ... & Zhang, Z. (2016). Glycoside hydrolase MoGls2 controls asexual/sexual development, cell wall integrity and infectious growth in the rice blast fungus. *PloS One*, *11*(9), e0162243. https://doi.org/10.1371/journal.pone.0162243
- Liu, T., Hou, J., Wang, Y., Jin, Y., Borth, W., Zhao, F., ... & Zuo, Y. (2016). Genome-wide identification, classification and expression analysis in fungal–plant interactions of cutinase gene family and functional analysis of a putative ClCUT7 in Curvularia lunata. *Molecular Genetics and Genomics*, *291*(3), 1105-1115. https://doi.org/10.1007/s00438-016-1168-1
- Liuzzi, V. C., Mirabelli, V., Cimmarusti, M. T., Haidukowski, M., Leslie, J. F., Logrieco, A. F., ... & Mulè, G. (2017). Enniatin and beauvericin biosynthesis in Fusarium species: production profiles and structural determinant prediction. *Toxins*, *9*(2), 45. https://doi.org/10.3390/toxins9020045
- Lorang, J. M., Tuori, R. P., Martinez, J. P., Sawyer, T. L., Redman, R. S., Rollins, J. A., ... & Ciuffetti, L. M. (2001). Green fluorescent protein is lighting up fungal biology. *Applied Environmental Microbiology*, *67*(5), 1987-1994. https://doi.org/10.1128/AEM.67.5.1987-1994.2001
- Lu, L., Rong, W., Massart, S., & Zhang, Z. (2018). Genome-wide identification and expression analysis of cutinase gene family in *Rhizoctonia cerealis* and functional study of an active cutinase RcCUT1 in the fungal-wheat interaction. *Frontiers in Microbiology*, *9*, 1813. https://doi.org/10.3389/fmicb.2018.01813

- Malapi-Wight, M., Smith, J., Campbell, J., Bluhm, B. H., & Shim, W. B. (2013). Sda1, a Cys2-His2 Zinc finger transcription factor, is involved in polyol metabolism and fumonisin B1 production in *Fusarium verticillioides*. *PloS One*, *8*(7), e67656. https://doi.org/10.1371/journal.pone.0067656
- Malonek, S., Bömke, C., Bornberg-Bauer, E., Rojas, M. C., Hedden, P., Hopkins, P., & Tudzynski, B. (2005a). Distribution of gibberellin biosynthetic genes and gibberellin production in the Gibberella fujikuroi species complex. *Phytochemistry*, *66*(11), 1296-1311. https://doi.org/10.1016/j.phytochem.2005.04.012
- Malonek, S., Rojas, M. C., Hedden, P., Hopkins, P., & Tudzynski, B. (2005b). Restoration of gibberellin production in Fusarium proliferatum by functional complementation of enzymatic blocks. *Applied Environmental Microbiology*, *71*(10), 6014-6025. https://doi.org/10.1128/AEM.71.10.6014-6025.2005
- Manosalva, P. M., Davidson, R. M., Liu, B., Zhu, X., Hulbert, S. H., Leung, H., & Leach, J. E. (2009). A germin-like protein gene family functions as a complex quantitative trait locus conferring broad-spectrum disease resistance in rice. *Plant Physiology*, *149*(1), 286-296. https://doi.org/10.1104/pp.108.128348
- Marin, S., Magan, N., Ramos, A. J., & Sanchis, V. (2004). Fumonisin-producing strains of Fusarium: a review of their ecophysiology. *Journal of Food Protection*, *67*(8), 1792-1805. https://doi.org/10.4315/0362-028X-67.8.1792
- Matić, S., Bagnaresi, P., Biselli, C., Carneiro, G. A., Siciliano, I., Valé, G., ... & Spadaro, D. (2016). Comparative transcriptome profiling of resistant and susceptible rice genotypes in response to the seedborne pathogen *Fusarium fujikuroi*. *BMC Genomics*, *17*(1), 608. https://doi.org/10.1186/s12864-016-2925-6
- Matic, S., Gullino, M. L., & Spadaro, D. (2017). The puzzle of bakanae disease through interactions between *Fusarium fujikuroi* and rice. *Frontiers in Bioscience (Elite Edition)*, *9*, 333-344. https://doi.org/10.2741/e806
- Matić, S., Spadaro, D., Garibaldi, A., & Gullino, M. L. (2014). Antagonistic yeasts and thermotherapy as seed treatments to control *Fusarium fujikuroi* on rice. *Biological Control*, *73*, 59-67. https://doi.org/10.1016/j.biocontrol.2014.03.008
- Matić, S., Spadaro, D., Prelle, A., Gullino, M. L., & Garibaldi, A. (2013). Light affects fumonisin production in strains of *Fusarium fujikuroi*, *Fusarium proliferatum*, and

Fusarium verticillioides isolated from rice. *International journal of food microbiology*, *166*(3), 515-523. https://doi.org/10.1016/j.ijfoodmicro.2013.07.026

- Meijer, S., De Jongh, W. A., Olsson, L., & Nielsen, J. (2009). Physiological characterisation of acuB deletion in *Aspergillus niger*. *Applied Microbiology and Biotechnology*, *84*(1), 157. https://doi.org/10.1007/s00253-009-2027-3
- Michielse, C. B., Studt, L., Janevska, S., Sieber, C. M., Arndt, B., Espino, J. J., ... & Tudzynski, B. (2015). The global regulator FfSge 1 is required for expression of secondary metabolite gene clusters but not for pathogenicity in *Fusarium fujikuroi*. *Environmental Microbiology*, *17*(8), 2690-2708. https://doi.org/10.1111/1462-2920.12592
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., ... & Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proceedings of the National Academy of Sciences*, *98*(23), 12902-12907. https://doi.org/10.1073/pnas.231473398
- Milne, I., Stephen, G., Bayer, M., Cock, P. J., Pritchard, L., Cardle, L., ... & Marshall, D. (2012). Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics*, *14*(2), 193-202. https://doi.org/10.1093/bib/bbs012
- Mohammed, A. R., & Tarpley, L. (2009). Impact of high nighttime temperature on respiration, membrane stability, antioxidant capacity, and yield of rice plants. *Crop Science*, *49*(1), 313-322. https://doi.org/10.2135/cropsci2008.03.0161
- Moretti, A. A. (1996). Diversity in beauvericin and fusaproliferin production by different populations of *Gibberella fujikuroi* (Fusarium section Liseola). *Sydowia*, *48*, 44-56.
- Munkvold, G. P. (2009). Seed pathology progress in academia and industry. *Annual Review of Phytopathology*, *47*. https://doi.org/10.1146/annurev-phyto-080508-081916
- Myung, K., Zitomer, N. C., Duvall, M., Glenn, A. E., Riley, R. T., & Calvo, A. M. (2012). The conserved global regulator VeA is necessary for symptom production and mycotoxin synthesis in maize seedlings by *Fusarium verticillioides*. *Plant Pathology*, *61*(1), 152-160. https://doi.org/10.1111/j.1365-3059.2011.02504.x
- Niehaus, E. M., Janevska, S., von Bargen, K. W., Sieber, C. M., Harrer, H., Humpf, H. U., & Tudzynski, B. (2014). Apicidin F: characterization and genetic manipulation

of a new secondary metabolite gene cluster in the rice pathogen *Fusarium fujikuroi*. *PLoS One*, *9*(7), e103336. https://doi.org/10.1371/journal.pone.0103336

- Niehaus, E. M., Kim, H. K., Münsterkötter, M., Janevska, S., Arndt, B., Kalinina, S. A., ... & Kim, D. W. (2017). Comparative genomics of geographically distant *Fusarium fujikuroi* isolates revealed two distinct pathotypes correlating with secondary metabolite profiles. *PLoS Pathogens*, *13*(10), e1006670. https://doi.org/10.1371/journal.ppat.1006670
- Niehaus, E. M., Münsterkötter, M., Proctor, R. H., Brown, D. W., Sharon, A., Idan, Y., ... & Tarkowská, D. (2016). Comparative "omics" of the Fusarium fujikuroi species complex highlights differences in genetic potential and metabolite synthesis. *Genome biology and evolution*, *8*(11), 3574-3599. https://doi.org/10.1093/gbe/evw259
- Nielsen, H. (2017). Predicting secretory proteins with SignalP. In *Protein Function Prediction* (59-73). Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7015-5_6
- Ogiso, Y., Sugiura, R., Kamo, T., Yanagiya, S., Lu, Y., Okazaki, K., ... & Kuno, T. (2004). Lub1 participates in ubiquitin homeostasis and stress response via maintenance of cellular ubiquitin contents in fission yeast. *Molecular and Cellular Biology*, *24*(6), 2324-2331. https://doi.org/10.1128/MCB.24.6.2324-2331.2004
- Onaga, G., Wydra, K. D., Koopmann, B., Séré, Y., & von Tiedemann, A. (2017). Elevated temperature increases in planta expression levels of virulence related genes in *Magnaporthe oryzae* and compromises resistance in *Oryza sativa* cv. Nipponbare. *Functional plant biology*, *44*(3), 358-371. https://doi.org/10.1071/FP16151
- Ortega, S. F., Tomlinson, J., Hodgetts, J., Spadaro, D., Gullino, M. L., & Boonham, N. (2018). Development of Loop-Mediated Isothermal Amplification Assays for the Detection of Seedborne Fungal Pathogens *Fusarium fujikuroi* and *Magnaporthe oryzae* in Rice Seed. *Plant Disease*, *102*(8), 1549-1558. https://doi.org/10.1094/PDIS-08-17-1307-RE
- Paccanaro, M. C., Sella, L., Castiglioni, C., Giacomello, F., Martínez-Rocha, A. L., D'Ovidio, R., ... & Favaron, F. (2017). Synergistic Effect of Different Plant Cell Wall–Degrading Enzymes Is Important for Virulence of *Fusarium graminearum*. *Molecular Plant-Microbe Interactions*, *30*(11), 886-895. https://doi.org/10.1094/MPMI-07-17-0179-R

- Pan, J. H., Chen, Y., Huang, Y. H., Tao, Y. W., Wang, J., Li, Y., ... & Lin, Y. C. (2011). Antimycobacterial activity of fusaric acid from a mangrove endophyte and its metal complexes. *Archives of Pharmacal Research*, *34*(7), 1177. https://doi.org/10.1007/s12272-011-0716-9
- Park, B. H., Karpinets, T. V., Syed, M. H., Leuze, M. R., & Uberbacher, E. C. (2010). CAZymes Analysis Toolkit (CAT): web service for searching and analysing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology*, *20*(12), 1574-1584. https://doi.org/10.1093/glycob/cwq106
- Passardi, F., Cosio, C., Penel, C., & Dunand, C. (2005). Peroxidases have more functions than a Swiss army knife. *Plant cell reports*, *24*(5), 255-265. https://doi.org/10.1007/s00299-005-0972-6
- Paulino, D., Warren, R. L., Vandervalk, B. P., Raymond, A., Jackman, S. D., & Birol, I. (2015). Sealer: a scalable gap-closing application for finishing draft genomes. *BMC bioinformatics*, *16*(1), 230. https://doi.org/10.1186/s12859-015-0663-4
- Pfaffl, M. W., Horgan, G. W., & Dempfle, L. (2002). Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic acids research*, *30*(9), e36-e36. https://doi.org/10.1093/nar/30.9.e36
- Pfannmüller, A., Leufken, J., Studt, L., Michielse, C. B., Sieber, C. M., Güldener, U., ... & Tudzynski, B. (2017). Comparative transcriptome and proteome analysis reveals a global impact of the nitrogen regulators AreA and AreB on secondary metabolism in *Fusarium fujikuroi*. *PloS One*, *12*(4), e0176194. https://doi.org/10.1371/journal.pone.0176194
- Pharis, R. P., & King, R. W. (1985). Gibberellins and reproductive development in seed plants. *Annual Review of Plant Physiology*, *36*(1), 517-568. https://doi.org/10.1146/annurev.pp.36.060185.002505
- Piombo, E., Bosio, P., Acquadro, A., Abbruscato, P., & Spadaro, D. (2019). Different phenotypes, similar genomes: three newly sequenced *Fusarium fujikuroi* strains induce different symptoms in rice depending on temperature. *Phytopathology*, https://doi.org/10.1094/PHYTO-09-19-0359-R

- Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., ... & Shakir, K. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, 201178. https://doi.org/10.1101/201178
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research*, *33*(2), 116-120.https://doi.org/10.1093/nar/gki442
- Radwan, O., Gunasekera, T. S., & Ruiz, O. N. (2018). Draft genome sequence of *Fusarium fujikuroi*, a fungus adapted to the fuel environment. *Genome Announcements.*, *6*(3), e01499-17. https://doi.org/10.1128/genomeA.01499-17
- Rocha, L. O., Laurence, M. H., Ludowici, V. A., Puno, V. I., Lim, C. C., Tesoriero, L. A., ... & Liew, E. C. Y. (2016). Putative effector genes detected in *Fusarium oxysporum* from natural ecosystems of Australia. *Plant Pathology*, *65*(6), 914-929. https://doi.org/10.1111/ppa.12472
- Sangha, J. S., Gu, K., Kaur, J., & Yin, Z. (2010). An improved method for RNA isolation and cDNA library construction from immature seeds of Jatropha curcas L. *BMC Research Notes*, *3*(1), 126. https://doi.org/10.1186/1756-0500-3-126
- Saremi, H., Ammarellou, A., Marefat, A., & Okhovvat, S. M. (2008). Binam a rice cultivar, resistant for root rot disease on rice caused by *Fusarium moniliforme* in Northwest, Iran. *International Journal of Botany*, *4*, 383-389. https://doi.org/10.3923/ijb.2008.383.389
- Saremi, H., and Farrokhi, F. 2004. Study on bakanae disease of rice and evaluation of cultivars in Gilan and Zanjan provinces, Iran. *Proceedings of the Fourth International Iran and Russia Conference*: 358-364.
- Scala, V., Giorni, P., Cirlini, M., Ludovici, M., Visentin, I., Cardinale, F., ... & Galaverna, G. (2014). LDS1-produced oxylipins are negative regulators of growth, conidiation and fumonisin synthesis in the fungal maize pathogen *Fusarium verticillioides*. *Frontiers in Microbiology*, *5*, 669. https://doi.org/10.3389/fmicb.2014.00669
- Schönig, B., Brown, D. W., Oeser, B., & Tudzynski, B. (2008). Cross-species hybridization with *Fusarium verticillioides* microarrays reveals new insights into *Fusarium fujikuroi* nitrogen regulation and the role of AreA and NMR. *Eukaryotic Cell*, *7*(10), 1831-1846. https://doi.org/10.1128/EC.00130-08

- Scott, P. M. (2012). Recent research on fumonisins: a review. *Food Additives & Contaminants: part A*, *29*(2), 242-248. https://doi.org/10.1080/19440049.2010.546000
- Sesma, A., & Osbourn, A. E. (2004). The rice leaf blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature*, *431*(7008), 582-586. https://doi.org/10.1038/nature02880
- Sharma, N., Sharma, K. P., Gaur, R. K., & Gupta, V. K. (2011). Role of chitinase in plant defense. *Asian Journal of Biochemistry*, *6*(1), 29-37. https://doi.org/10.3923/ajb.2011.29.37
- Siciliano, I., Amaral Carneiro, G., Spadaro, D., Garibaldi, A., & Gullino, M. L. (2015). Jasmonic acid, abscisic acid, and salicylic acid are involved in the phytoalexin responses of rice to *Fusarium fujikuroi,* a high gibberellin producer pathogen. *Journal of Agricultural and Food Chemistry*, *63*(37), 8134-8142. https://doi.org/10.1021/acs.jafc.5b03018
- Son, S. W., Kim, H. Y., Choi, G. J., Lim, H. K., Jang, K. S., Lee, S. O., ... & Kim, J. C. (2008). Bikaverin and fusaric acid from Fusarium oxysporum show antioomycete activity against *Phytophthora infestans*. *Journal of Applied Microbiology*, *104*(3), 692-698. https://doi.org/10.1111/j.1365-2672.2007.03581.x
- Sondergaard, T. E., Hansen, F. T., Purup, S., Nielsen, A. K., Bonefeld-Jørgensen, E. C., Giese, H., & Sørensen, J. L. (2011). Fusarin C acts like an estrogenic agonist and stimulates breast cancer cells *in vitro*. *Toxicology Letters*, *205*(2), 116-121. https://doi.org/10.1016/j.toxlet.2011.05.1029
- Soyer, J. L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., ... & Fudal, I. (2014). Epigenetic control of effector gene expression in the plant pathogenic fungus *Leptosphaeria maculans*. *PLoS genetics*, *10*(3). https://doi.org/https://dx.doi.org/10.1371%2Fjournal.pgen.1004227
- Spadaro, D., & Gullino, M. L. (2005). Improving the efficacy of biocontrol agents against soilborne pathogens. *Crop protection*, *24*(7), 601-613. https://doi.org/10.1016/j.cropro.2004.11.003
- Sperschneider, J., Gardiner, D. M., Dodds, P. N., Tini, F., Covarelli, L., Singh, K. B., ... & Taylor, J. M. (2016). EffectorP: predicting fungal effector proteins from secretomes using machine learning. *New Phytologist*, *210*(2), 743-761. https://doi.org/10.1111/nph.13794

- Sponsel, V. M. (2001). The deoxyxylulose phosphate pathway for the biosynthesis of plastidic isoprenoids: early days in our understanding of the early stages of gibberellin biosynthesis. *Journal of Plant Growth Regulation*, *20*(4), 332-345. https://doi.org/10.1007/s003440010032
- Sreenivasulu, N., Butardo Jr, V. M., Misra, G., Cuevas, R. P., Anacleto, R., & Kavi Kishor, P. B. (2015). Designing climate-resilient rice with ideal grain quality suited for high-temperature stress. *Journal of Experimental Botany*, *66*(7), 1737-1748. https://doi.org/10.1093/jxb/eru544
- Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, *33*(2), 465-467. https://doi.org/10.1093/nar/gki458
- Stanke, M., & Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, *19*(2), 215-225. https://doi.org/10.1093/bioinformatics/btg1080
- Stępień, Ł., Koczyk, G., & Waśkiewicz, A. (2011). FUM cluster divergence in fumonisins-producing Fusarium species. *Fungal Biology*, *115*(2), 112-123. https://doi.org/10.1016/j.funbio.2010.10.011
- Stergiopoulos, I., & de Wit, P. J. (2009). Fungal effector proteins. *Annual Review of Phytopathology*, *47*, 233-263. https://doi.org/10.1146/annurev.phyto.112408.132637
- Studt, L., Janevska, S., Arndt, B., Boedi, S., Sulyok, M., Humpf, H. U., ... & Strauss, J. (2017). Lack of the COMPASS component Ccl1 reduces H3K4 trimethylation levels and affects transcription of secondary metabolite genes in two plant–pathogenic Fusarium species. *Frontiers in Microbiology*, *7*, 2144. https://doi.org/10.3389/fmicb.2016.02144
- Studt, L., Janevska, S., Niehaus, E. M., Burkhardt, I., Arndt, B., Sieber, C. M., ... & Tudzynski, B. (2016). Two separate key enzymes and two pathway-specific transcription factors are involved in fusaric acid biosynthesis in *Fusarium fujikuroi*. *Environmental Microbiology*, *18*(3), 936-956. https://doi.org/10.1111/1462-2920.13150
- Studt, L., Wiemann, P., Kleigrewe, K., Humpf, H. U., & Tudzynski, B. (2012). Biosynthesis of fusarubins accounts for pigmentation of *Fusarium fujikuroi* perithecia. *Applied Environmental Microbiology*, *78*(12), 4468-4480. https://doi.org/10.1128/AEM.00823-12

- Sun, S. K. (1975). The disease cycle of rice bakanae disease in Taiwan. *Proceedings of the National Science Council*, *8*(2), 245-256.
- Sunani, S. K., Bashyal, B. M., Kharayat, B. S., Prakash, G., Krishnan, S. G., & Aggarwal, R. (2019). Identification of rice seed infection routes of *Fusarium fujikuroi* inciting bakanae disease of rice. *Journal of Plant Pathology*, 1-9. https://doi.org/10.1007/s42161-019-00390-8
- Sunani, S. K., Bashyal, B. M., Rawat, K., Manjunatha, C., Sharma, S., Prakash, G., ... & Aggarwal, R. (2019). Development of PCR and loop mediated isothermal amplification assay for the detection of bakanae pathogen *Fusarium fujikuroi*. *European Journal of Plant Pathology*, *154*(3), 715-725. https://doi.org/10.1007/s10658-019-01694-2
- Suzuki, Y., Mae, T., & Makino, A. (2008). RNA extraction from various recalcitrant plant tissues with a cethyltrimethylammonium bromide-containing buffer followed by a acid guanidium thiocyanate-phenol-chloroform treatment. *Bioscience, Biotechnology, and Biochemistry*, 0806060902-0806060902. https://doi.org/10.1271/bbb.80084
- Sydenham, E. W., Thiel, P. G., Marasas, W. F., Shephard, G. S., Van Schalkwyk, D. J., & Koch, K. R. (1990). Natural occurrence of some Fusarium mycotoxins in corn from low and high esophageal cancer prevalence areas of the Transkei, Southern Africa. *Journal of Agricultural and Food Chemistry*, *38*(10), 1900-1903. https://doi.org/10.1021/jf00100a004
- Taevernier, L., Wynendaele, E., De Vreese, L., Burvenich, C., & De Spiegeleer, B. (2016). The mycotoxin definition reconsidered towards fungal cyclic depsipeptides. *Journal of Environmental Science and Health, Part C*, *34*(2), 114-135. https://doi.org/10.1080/10590501.2016.1164561
- Tanaka, N., Ohuchi, N., Mukai, Y., Osaka, Y., Ohtani, Y., Tabuchi, M., ... & Takegawa, K. (1998). Isolation and Characterization of an Invertase and Its Repressor Genes from *Schizosaccharomyces pombe*. *Biochemical and Biophysical Research Communications*, *245*(1), 246-253. https://doi.org/10.1006/bbrc.1998.8406
- Tani, S., Yuki, S., Kunitake, E., Sumitani, J. I., & Kawaguchi, T. (2017). Dipeptidyl peptidase IV is involved in the cellulose-responsive induction of cellulose biomass-degrading enzyme genes in *Aspergillus aculeatus*. *Bioscience, Biotechnology, and Biochemistry*, *81*(6), 1227-1234.https://doi.org/10.1080/09168451.2017.1295800

- Teichert, S., Rutherford, J. C., Wottawa, M., Heitman, J., & Tudzynski, B. (2008). Impact of ammonium permeases mepA, mepB, and mepC on nitrogen-regulated secondary metabolism in *Fusarium fujikuroi*. *Eukaryotic Cell*, *7*(2), 187-201. https://doi.org/10.1128/EC.00351-07
- Teichert, S., Wottawa, M., Schönig, B., & Tudzynski, B. (2006). Role of the *Fusarium fujikuroi* TOR kinase in nitrogen regulation and secondary metabolism. *Eukaryotic Cell*, *5*(10), 1807-1819. https://doi.org/10.1128/EC.00039-06
- Tsavkelova, E. A., Bömke, C., Netrusov, A. I., Weiner, J., & Tudzynski, B. (2008). Production of gibberellic acids by an orchid-associated Fusarium proliferatum strain. *Fungal genetics and Biology*, *45*(10), 1393-1403. https://doi.org/10.1016/j.fgb.2008.07.011
- Tudzynski, B., & Hölter, K. (1998). Gibberellin Biosynthetic Pathway in *Gibberella fujikuroi*: Evidence for a Gene Cluster. *Fungal Genetics and Biology*, *25*(3), 157-170. 25:157-170. https://doi.org/10.1006/fgbi.1998.1095
- Tudzynski, B., Homann, V., Feng, B., & Marzluf, G. A. (1999). Isolation, characterization and disruption of the areA nitrogen regulatory gene of *Gibberella fujikuroi*. *Molecular and General Genetics MGG*, *261*(1), 106-114. https://doi.org/10.1007/s004380050947
- Tudzynski, B., Mihlan, M., Rojas, M. C., Linnemannstöns, P., Gaskin, P., & Hedden, P. (2003). Characterization of the final two genes of the gibberellin biosynthesis gene cluster of *Gibberella fujikuroi* des and P450-3 encode GA4 desaturase and the 13-hydroxylase, respectively. *Journal of Biological Chemistry*, *278*(31), 28635-28643. https://doi.org/10.1074/jbc.M301927200
- Urban, M., Cuzick, A., Rutherford, K., Irvine, A., Pedro, H., Pant, R., ... & Hammond-Kosack, K. E. (2016). PHI-base: a new interface and further additions for the multi-species pathogen–host interactions database. *Nucleic Acids Research*, *45*(1), 604-610. https://doi.org/10.1093/nar/gkw1089
- Urbaniak, C., Dadwal, S., Bagramyan, K., & Venkateswaran, K. (2018). Draft Genome Sequence of a Clinical Isolate of *Fusarium fujikuro*i Isolated from a Male Patient with Acute Myeloid Leukemia. *Genome Announcements*, *6*(25). https://doi.org/10.1128/genomeA.00476-18
- Valente, M. T., Desiderio, F., Infantino, A., Valè, G., Abbruscato, P., & Aragona, M. (2017). Genetic Variability of *Fusarium fujikuroi* Populations associated with

Bakanae of rice in Italy. *Plant Pathology*, 66(3), 469-479. https://doi.org/10.1111/ppa.12575

- Vallad, G. E., & Subbarao, K. V. (2008). Colonization of resistant and susceptible lettuce cultivars by a green fluorescent protein-tagged isolate of Verticillium dahliae. *Phytopathology*, *98*(8), 871-885. https://doi.org/10.1094/PHYTO-98-8-0871
- Van der Does, H. C., & Rep, M. (2017). Adaptation to the host environment by plant-pathogenic fungi. *Annual Review of Phytopathology*, *55*, 427-450. https://doi.org/10.1146/annurev-phyto-080516-035551
- Van Vu, B., Pham, K. T. M., & Nakayashiki, H. (2013). Substrate-induced transcriptional activation of the MoCel7C cellulase gene is associated with methylation of histone H3 at lysine 4 in the rice blast fungus *Magnaporthe oryzae*. *Applied Environmental Microbiology*, *79*(21), 6823-6832. https://doi.org/10.1128/AEM.02082-13
- Visentin, I., Montis, V., Döll, K., Alabouvette, C., Tamietti, G., Karlovsky, P., & Cardinale, F. (2012). Transcription of genes in the biosynthetic pathway for fumonisin mycotoxins is epigenetically and differentially regulated in the fungal maize pathogen *Fusarium verticillioides*. *Eukaryotic cell*, *11*(3), 252-259. https://doi.org/10.1128/EC.05159-11
- Volante, A., Tondelli, A., Aragona, M., Valente, M. T., Biselli, C., Desiderio, F., ... & Spadaro, D. (2017). Identification of bakanae disease resistance loci in japonica rice through genome wide association study. *Rice*, *10*(1), 29. https://doi.org/10.1186/s12284-017-0168-z
- Wang, G., Wang, G., Zhang, X., Wang, F., & Song, R. (2012). Isolation of high quality RNA from cereal seeds containing high levels of starch. *Phytochemical Analysis*, *23*(2), 159-163. https://doi.org/10.1002/pca.1337
- Wasmann, C. C., & VanEtten, H. D. (1996). Transformation-mediated chromosome loss and disruption of a gene for pisatin demethylase decrease the virulence of *Nectria haematococca* on pea. *Molecular Plant-Microbe Interactions,* 9(9), 793-803. https://doi.org/10.1094/MPMI-9-0793
- Watanabe, S., Kumakura, K., Izawa, N., Nagayama, K., Mitachi, T., Kanamori, M., ... & Arie, T. (2007). Mode of action of Trichoderma asperellum SKT-1, a biocontrol agent against *Gibberella fujikuroi*. *Journal of Pesticide Science*, 0706230005-0706230005. https://doi.org/10.1584/jpestics.G06-35

- White, T. C., & Silver, P. M. (2005). Regulation of sterol metabolism in *Candida albicans* by the UPC2 gene. *Seventh Yeast Lipid Conference. Biochemical Society, 33*, 1215-1218. https://doi.org/10.1042/BST0331215
- Wiebe, L. A., & Bjeldanes, L. F. (1981). Fusarin C, a mutagen from *Fusarium moniliforme* grown on corn. *Journal of Food Science*, *46*(5), 1424-1426. https://doi.org/10.1111/j.1365-2621.1981.tb04189.x
- Wiemann, P., Brown, D. W., Kleigrewe, K., Bok, J. W., Keller, N. P., Humpf, H. U., & Tudzynski, B. (2010). FfVel1 and FfLae1, components of a velvet-like complex in *Fusarium fujikuroi*, affect differentiation, secondary metabolism and virulence. *Molecular Microbiology*, *77*(4), 972-994. https://doi.org/10.1111/j.1365-2958.2010.07263.x
- Wiemann, P., Sieber, C. M., Von Bargen, K. W., Studt, L., Niehaus, E. M., Espino, J. J., ... & Bergner, S. V. (2013). Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathogens*, *9*(6), e1003475. https://doi.org/10.1371/journal.ppat.1003475
- Winnenburg, R., Baldwin, T. K., Urban, M., Rawlings, C., Köhler, J., & Hammond-Kosack, K. E. (2006). PHI-base: a new database for pathogen host interactions. *Nucleic Acids Research*, *34*(1), 459-464. https://doi.org/10.1093/nar/gkj047
- Wu, J., Liu, Y., Lv, W., Yue, X., Que, Y., Yang, N., ... & Wang, Z. (2015). FgRIC8 is involved in regulating vegetative growth, conidiation, deoxynivalenol production and virulence in *Fusarium graminearum*. *Fungal Genetics and Biology*, *83*, 92-102. https://doi.org/10.1016/j.fgb.2015.08.012
- Wulff, E. G., Sørensen, J. L., Lübeck, M., Nielsen, K. F., Thrane, U., & Torp, J. (2010). *Fusarium spp.* associated with rice Bakanae: ecology, genetic diversity, pathogenicity and toxigenicity. *Environmental Microbiology*, *12*(3), 649-657. https://doi.org/10.1111/j.1462-2920.2009.02105.x
- Xu, J. R., & Leslie, J. F. (1996). A genetic map of *Gibberella fujikuroi* mating population A (*Fusarium moniliforme*). *Genetics*, *143*(1), 175-189.
- Xu, Y., Orozco, R., Wijeratne, E. K., Gunatilaka, A. L., Stock, S. P., & Molnár, I. (2008). Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus *Beauveria bassiana*. *Chemistry & Biology*, *15*(9), 898-907. https://doi.org/10.1016/j.chembiol.2008.07.011

- Yang, Y., Bouras, N., Yang, J., Howard, R. J., & Strelkov, S. E. (2011). Mycotoxin production by isolates of Fusarium lactis from greenhouse sweet pepper (Capsicum annuum). *International journal of food microbiology*, *151*(2), 150-156. https://doi.org/10.1016/j.ijfoodmicro.2011.08.013
- Zainudin, N. I. M., Razak, A., & Salleh, B. (2008). Bakanae disease of rice in Malaysia and Indonesia: etiology of the causal agent based on morphological, physiological and pathogenicity characteristics. *Journal of Plant Protection Research*, *48*(4), 475-485. https://doi.org/10.2478/v10045-008-0056-z
- Zeevaart, J. A. (1976). Physiology of flower formation. *Annual Review of Plant Physiology*, *27*(1), 321-348. https://doi.org/10.1146/annurev.pp.27.060176.001541
- Zhan, J., Burns, A. M., Liu, M. X., Faeth, S. H., & Gunatilaka, A. L. (2007). Search for cell motility and angiogenesis inhibitors with potential anticancer activity: beauvericin and other constituents of two endophytic strains of *Fusarium oxysporum*. *Journal of Natural Products*, *70*(2), 227-232. https://doi.org/10.1021/np060394t
- Zhang, S. Y., Dai, D. J., Wang, H. D., & Zhang, C. Q. (2019). One-step loop-mediated isothermal amplification (LAMP) for the rapid and sensitive detection of *Fusarium fujikuroi* in bakanae disease through NRPS31, an important gene in the gibberellic acid bio-synthesis. *Scientific Reports*, *9*(1), 3726. https://doi.org/10.1038/s41598-019-39874-z

APPENDIX A: Genome sequence, assembly and characterization of two *Metschnikowia fructicola* strains used as biocontrol agents of postharvest diseases

Edoardo Piombo[1]*, Noa Sela[2]*, Michael Wisniewski[3], Maria Hoffmann[4], Maria Lodovica Gullino[1,5], Marc W Allard[4], Elena Levin[6], Davide Spadaro[1] and Samir Droby[6]**

[1]Department of Agricultural, Forestry and Food Sciences (DISAFA), University of Torino, Grugliasco (TO), 10095, Italy.
[2]Dept. Plant Pathology and weed research, ARO, The Volcani Center, Rishon Le-Zion, 7505101, Israel
[3]United States Department of Agriculture – Agricultural Research Service (USDA-ARS), Kernersville, WV 254430, USA
[4]United States Food and Drug Administration (FDA), Division of Microbiology, Office of Regulatory Science, College Park, MD 20740, USA
[5]Centre of Competence for the Innovation in the Agroenvironmental Sector (AGROINNOVA), University of Torino, Grugliasco (TO), 10095, Italy.
[6]Dept. Postharvest Science, ARO, the Volcani Center, P. O. Box 15159, Rishon Le-Zion, 7505101, Israel

*The two authors contributed at the same level for the realization of the presented work.

**corresponding author samird@volcani.agri.gov.il

ABSTRACT

The yeast *Metschnikowia fructicola* was reported as an efficient biological control agent of postharvest diseases of fruits and vegetables, and it is the bases of the

commercial formulated product "Shemer". Several mechanisms of action by which *M. fructicola* inhibits postharvest pathogens were suggested including iron-binding compounds, induction of defence signaling genes, production of fungal cell wall degrading enzymes and relatively high amounts of superoxide anions. We assembled the whole genome sequence of two strains of *M. fructicola* using PacBio and Illumina shotgun sequencing technologies. Using the PacBio, a high-quality draft genome consisting of 93 contigs, with an estimated genome size of approximately 26 Mb, was obtained. Comparative analysis of *M. fructicola* proteins with the other three available closely related genomes revealed a shared core of homologous proteins coded by 5,776 genes. Comparing the genomes of the two *M. fructicola* strains using a SNP calling approach resulted in the identification of 564,302 homologous SNPs with 2,004 predicted high impact mutations. The size of the genome is exceptionally high when compared with those of available closely related organisms, and the high rate of homology among *M. fructicola* genes points towards a recent whole-genome duplication event as the cause of this large genome. Based on the assembled genome, sequences were annotated with a gene description and gene ontology (GO term) and clustered in functional groups. Analysis of CAZymes family genes revealed 1,145 putative genes, and transcriptomic analysis of CAZyme expression levels in *M. fructicola* during its interaction with either grapefruit peel tissue or *Penicillium digitatum* revealed a high level of CAZyme gene expression when the yeast was placed in wounded fruit tissue.

INTRODUCTION

The yeast *Metschnikowia fructicola* (type strain NRRL Y-27328, CBS 8853) was first isolated from grapes and identified as a new species by Kurtzman and Droby in 2001 (Kurtzman and Droby, 2001). The identification was achieved by comparing its nucleotide sequence in the species-specific ca. 500–600-nucleotide D1/D2 domain of 26S ribosomal DNA (rDNA) with a database of D1/D2 sequences from all the

recognized ascomycetous yeasts available at that time (Kurtzman and Robnett, 1998), and subsequent entries in GenBank.

Yeasts have been identified by many workers as potential biological control agents suitable for the prevention of postharvest diseases, especially since they are naturally occurring on fruits and vegetables, and exhibit a number of traits that favor their use as fungal antagonists. These traits include high tolerance to environmental stresses (low and high temperatures, desiccation, wide fluctuations in relative humidity, low oxygen levels, pH fluctuations, UV radiation) encountered during fruit and vegetable production before and after harvest, and their ability to adapt to the micro-environment present in wounded fruit tissues, characterized by high sugar concentration, high osmotic pressure, low pH and conditions that conducive to oxidative stress. These traits are especially beneficial for their use as biocontrol agents, since the majority of postharvest decay pathogens are necrotrophic and infect fruit through wounded tissues (Droby et al., 2016; Wisniewski et al., 2016). Additionally, many yeast species can grow rapidly on inexpensive substrates in fermenters, a trait that is conducive to their large-scale commercial production and use (Spadaro and Droby, 2016). Moreover, in contrast to filamentous fungi, the vast majority of naturally occurring yeasts do not produce allergenic spores or mycotoxins, and have simple nutritional requirements that enable them to colonize dry surfaces for long periods of time (Spadaro et al., 2008).

Significant progress has been made in the development, registration and commercialization of postharvest biocontrol products (Droby et al., 2009; Droby et al., 2016) and a variety of different biocontrol agents have reached advanced stages of development and commercialization. "Shemer", based on the yeast *M. fructicola* (Droby et al., 2009), is one of the commercial products that has reached the market.

Several studies have documented the biocontrol efficacy of *M. fructicola* and its ability to prevent or limit the infection of harvested products by postharvest pathogens (Karabulut et al., 2003; Karabulut et al., 2004; Spadaro et al., 2013). Similar to other

postharvest biocontrol agents, *M. fructicola* exhibits several modes of action to achieve its ability to act as an antagonist. Like its sister species *M. pulcherrima*, *M. fructicola* produces the red pigment, pulcherrimin, which is formed non-enzymatically from pulcherriminic acid and ferric ions. Pulcherrimin has been reported to play a role in the control of *Botrytis cinerea*, *Alternaria alternata,* and *Penicillium expansum* on apple (Saravanakumar et al., 2008). Enhanced expression of several genes involved in defence signaling, including PRP genes and MAPK cascade genes was demonstrated in grapefruit when surface wounds were treated with *M. fructicola* cells (Hershkovitz et al., 2012). The enhanced gene expression was consistent with an induced resistance response suggesting that induced host resistance plays a role in the biocontrol of *M. fructicola* against postharvest pathogens such as *P. digitatum* (Hershkovitz et al., 2012). *M. fructicola* also exhibits chitinase activity and the chitinase gene, *MfChi,* was demonstrated to be highly induced in yeast cells when cell walls of *Monilinia fructicola*, the causal agent of brown rot in stone fruit, was added to the growth medium. These data suggest that *MfChi* may also play a role in the biocontrol activity exhibited by Metschnikowia species (Banani et al., 2015). Macarisin et al (2010) demonstrated that yeast antagonists, including *M. fructicola*, used to control postharvest diseases have the ability to produce relatively high amounts of superoxide anions. They also demonstrated that yeast cells applied to surface wounds of fruits produce greater levels of superoxide anions than yeast grown *in vitro* in artificial media.

Several studies have examined differential gene expression during the interaction of the yeast *M. fructicola* with host fruit tissue or with the mycelium of the postharvest pathogen *P. digitatum* (Hershkovitz et al., 2012; Hershkovitz et al., 2013). Due to the lack of an assembled genome sequence, de-novo assembly of the transcriptome of *M. fructicola* was performed, which resulted in the identification of 9,674 unigenes, half of which could be annotated based on homology to genes in the NCBI database (Hershkovitz et al., 2013). Approximately, 69% of the unigene sequences identified

in *M. fructicola* showed high homology to genes of the yeast *Clavispora lusitaniae.* Thus, the RNA-Seq-based transcriptome analysis generated a large number of newly identified *M. fructicola* yeast genes and significantly increased the number of sequences available for *Metschnikowia* species in the NCBI database. Shotgun sequencing data enabled to construct a draft genome of *M. fructicola* based on Illumina paired-end assembly with ~7000 contigs that was submitted to Genbank (Hershkovitz et al., 2013).

Details about the structure and annotation of the genomes of yeast biocontrol agents are lacking. Such information would be a valuable tool for analysing the sequences of putative "biocontrol-related" genes among different species of yeast biocontrol agents, characterizing gene clusters with known and unknown functions, as well as studying global changes in gene transcription rather than just specific, targeted genes. Obtaining full genome sequences would also allow comparative genomic analyses to be conducted among closely related yeast species that do not exhibit antagonist properties (Massart et al., 2015).

In the present study, a whole genome sequence of the 277 type-strain of *M. fructicola* (NRRL Y-27328) was assembled using PacBio technology. Results indicate that the genome of *M. fructicola* (Mf genome) is approximately 26 Mbp and contains 8,629 gene coding sequences. The new assembly resulted in a high quality assembly consisting of 93 contigs – the longest one is 2,548,689 bp – with 439X average genome coverage.

In parallel, the genome of another biocontrol strain of *M. fructicola* (strain AP47) isolated in northern Italy from apple fruit surfaces and used to control brown rot of peaches (Zhang et al., 2010), was assembled by aligning Illumina shotgun sequences (with a genome coverage of 161.8 X), using the genome assembly of the strain 277 as a reference. The mutation rate between the two biocontrol strains of *M. fructicola* was also determined.

RESULTS AND DISCUSSION

Assembly, gene prediction and functional annotation of the genome of *Metschnikowia fructicola* strain 277

A new assembly of the *M. fructicola* (type strain NRRL Y-27328, CBS 8853) genome (Genbank accession ANFW02000000) was constructed using sequence data obtained from the Pacific Biosciences (PacBio) RS II Sequencer. The PacBio genomic sequences were assembled with the HGAP3.0 program (Chin et al., 2013) and yielded a high-quality draft genome consisting of 93 contigs with an N50 of 957,836 bp. The estimated genome size is approximately 26 Mb. Total of 8,629 genes were predicted with

	New sequence	**Old sequence** (Hershkovitz et al., 2013
Sequencing technology	PacBio	Illumina
Genome Size	~26Mb	~23Mb
Sequencing coverage	20X	700X
Number of contigs	93	8430
Number of Large contigs (>100 Kb)	84	2
N50 (base pairs)	957,836 bp	3,784 bp
GC content (%)	45.8%	45.5%
N50 of transcript length (nucleotides)	5033bp	589bp
Number of genes	8,629	15,803
Annotated genes	6,277	-

Table A.1: Summary of the main assembly and annotation features of the genome of the sequenced *Metschnikowia fructicola* strain 277.

MAKER, and 6,262 were successfully annotated with Blast2GO (Conesa et al., 2005)

and InterProScan (Finn et al., 2016). The results of assembly, gene prediction and annotation are presented in **Table A.1**. In contrast to the previous assembly (Hershkovitz et al., 2013), where 9,674 transcriptswere identified, the current high-quality assembly provided a more accurate estimate of the transcript number (8,629) and size of the *M. fructicola* genome.

We believe that the current number is more accurate because it was estimated by using the MAKER gene predictor (Cantarel et al., 2008), trained with the transcript sequences obtained by mapping the RNA reads obtained by Hershkovitz et al. (2013) on a high-quality genomic sequence. On the other hand, the 9,674 predicted by Hershkovitz et al. (2013) were obtained by *de novo* assembly with the Trinity software (Grabherr et al., 2011), which can be prone to the overestimation of the number of transcripts (Cerveau and Jackson, 2016). The annotated transcripts are listed in supplementary **Table A.1**, and their sequences, CDSs and protein sequences are presented in **supplementary Files A.1**, **A.2** and **A.3**. **Supplementary File A.4** contains the gene coordinates. The main characteristics of the current *M. fructicola* genome assembly and a comparison to the previous assembly (Hershkovitz et al., 2013) are summarized in **Table A.1**. Comparative analysis of *M. fructicola* proteins with the other three available closely related genomes of *Clavispora lusitaniae*, *Candida auris*, and *M. bicupsidata* revealed a shared core of homologous proteins coded by 5,776 genes (**supplementary File A.5**). A recently published work describing the phylogeny of strains belonging to Metschnikowia species isolated from the guts of flower-visiting insects (Lachance et al., 2016) allowed us to construct a phylogenetic tree of *Metschnikowia* spp that is based on the fastq raw-data deposited in Genbank (**Figure A.1**). The tree was constructed using an assembly and alignment-free method of phylogeny reconstruction (Fan et al., 2015). Interestingly, the phylogenetic analysis showed that the two *M. fructicola* strains described in our study were grouped together and were separate from other *Metschnikowia* species described

by Lachance et al. (2016). This difference in phylogeny may be related to evolutionary history and niche colonization of fruit surfaces versus insect guts.

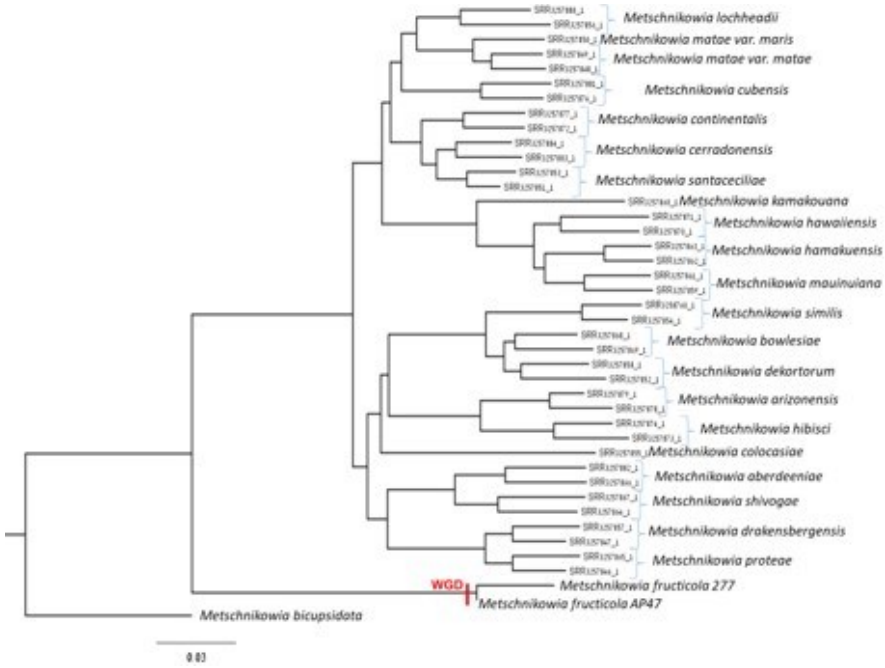

Figure A.1: Phylogenetic tree comprised of *Metschnikowia fructicola* 277, *Metschnikowia fructicola* AP47, and other *Metchnikowia* species. The tree was constructed using an assembly and alignment-free method of phylogeny reconstruction (Fan et al., 2015). The whole genome duplication event was indicated on the tree with "WGD".

The GO analysis revealed that 6,262 of the 8,629 identified *M. fructicola* genes were characterized with 4,493 GO terms (**supplementary File A.6**). The most common descriptors concerning the cellular component were "Cell" and "Cell Part", followed by "Organelle", while "Cellular process" and "Metabolic Process", followed by "Localization", "Establishment of Localization", "Biological Regulation", "Pigmentation" and "Response to stimulus" were the most common in the biological processes. Regarding the molecular function, the most common descriptors were "Binding" and "Catalytic", followed by "Transporter". The same descriptors in the

three categories were the most common in the genes characterized in the paper of Hershkovitz et al (2013).

Sequencing data	**Library PE1**	**Library PE2**	**Library MP1**
Number of raw reads	3717646	2599548	10188012
Number of clean reads	2545140	2546666	9126542
Total length (Mb)	301.257	927.79	2977.528
GC percentage	43 % GC	45% GC	43% GC

Table A.2: Sequencing data of the two pair end libraries used to sequence the genome of *Metschnikowia fructicola*, strain AP47.

Utilisation of *M. fructicola* 277 genome for reference-based assembly of strain AP47

The assembly of the genome of strain 277 presented here is the most comprehensive and complete assembly for *M. fructicola* to date. This assembly was used as a reference to assemble the genome of the AP47 strain of *M. fructicola*, obtained by Illumina MySeq (161.8 X) shotgun sequencing data (**Table A.2**). The reference guided assembly resulted in an N50 of 957,045, which was much higher than the one obtained by *de novo* assembly (**Table A.3**). The length of the AP47 genome was similar to the reference strain 277 (~ 26 Mb), but had a slightly higher GC content (46.3% compared to 45.8%).

*Obtained with SPAdes (*Bankevich et al., 2012*)
**Obtained with IMR-DENOM (http://mus.well.ox.ac.uk/19genomes/IMR-*DENOM/*)

	De novo assembly*	**Reference guided assembly****
Sequence length	~23.3 Mb	~26.2 Mb
Number of scaffolds	10,173	93
Number of scaffolds > 100 Kl	35	53
Number of scaffolds > 1 Kb	3156	93
N50 (base pairs)	63,477 bp	957,045 bp
G+C content (%)	46.3%	46.3%

Table A.3: *De novo* and reference guided assemblies of the genome of the sequenced *Metschnikowia fructicola*, strain AP47.

The assembly presented here was also compared to the AP47 strain assembly using a SNP calling approach. Results of this analysis are presented in **Table A.4**, and the complete vcf is found in **supplementary File A.7**. Considering only homozygous polymorphisms, a total of 546,356 SNPs, 11,987 insertions and 5,959 deletions were identified. Among these mutations, 185,649 were in coding regions, and the vast majority of the variations (135,616) were silent. However, 50,822 were missense mutations, and 212 were nonsense mutations. The differences with strain AP47 were mapped on strain 277 and presented in **Figure A.2**.

The average mutation rate was one every 46 bases, which is exceptionally high in respect to the average reported for other yeast species. For example, the average mutation rate is approximately one SNP every 235 and 269 nucleotides, in *C. albicans* (Hirakawa et al., 2015) and *Saccharomyces cerevisiae,* respectively (Drozdova et al., 2016). The high number of observed mutations may be related to the different geographical origin and host species of the strains. The 277 type-strain of *M. fructicola*

(NRRL Y-27328) was isolated in Israel from the surface of grapes, while the AP47 strain was isolated in Italy from the surface of apples.

Number of mutations	**564,302**
SNPs	546,356
Insertions	11,987
Deletions	5,959
Variant rate	1 variant every 46 bases
Predicted mutation effect	
Silent	135,884
Missense	49,794
Nonsense	212
Mutation impact	
High	2,023 (0.08%)
Moderate	50,032 (1.97%)
Low	136,810 (5.39%)
Negligible	2,348,195 (92.56%)

Table A.4: Number of mutations in the genome sequence of *M. fructicola* strain AP47, compared to the reference genome of *M. fructicola* strain 277, and their predicted effect and impact on coding sequences.

The strain AP47 Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MTJM00000000. The version described in this paper is version MTJM01000000.

The D1/D2 region ribosomal region was identified in strain 277 genome by blasting *M. pulcherrima* D1/D2 region on it. Since we observed that none of the identified SNPs were localized in that region, we can confirm with high confidence that both strains 277 and strain AP47 belong to the same species, which is different from *M. pulcherrima* (Kurtzman and Robnett, 1998).

Stress-induced genomic instability has been studied in various yeast and bacteria, under a variety of stress conditions. Stresses were suggested to induce several genetic changes including small changes (1 to few nucleotides), deletions and insertions, gross

chromosomal rearrangements, copy-number variations and movement of mobile elements (Galhardo et al., 2007).

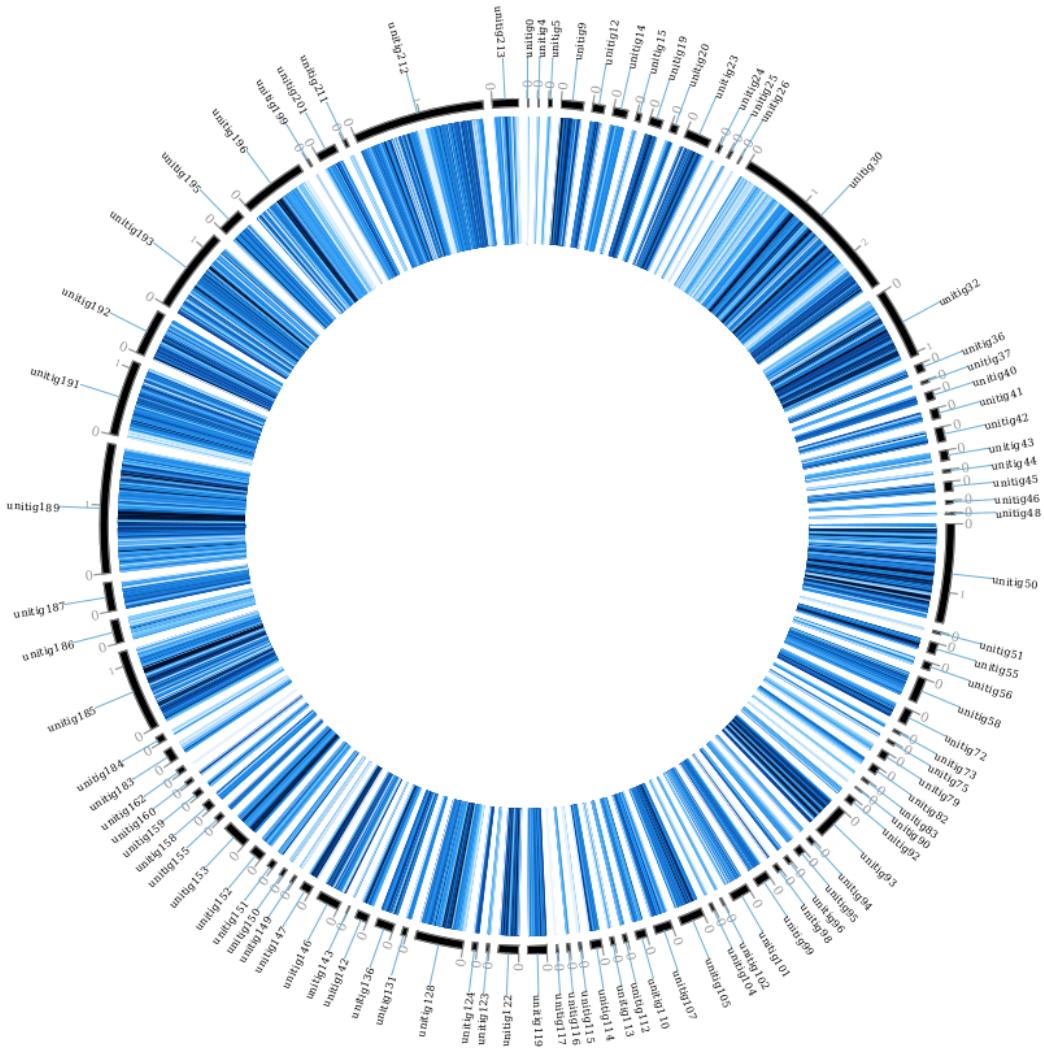

Figure A.2: homologous SNPs of *M. fructicola* strain AP47, mapped on the strain 277 genome. The figure was obtained using circoVCF tool (Drori *et al.*, 2017).

We suggest that *M. fructicola* as a species could undergo genomic changes in order to survive environmental stresses, in particular on the fruit surface. These changes may have led to evolve mechanisms not only to tolerate stresses, but also to generate large-scale genetic variation as a means of adaptation, giving both *M. fructicola* strains the

genetic traits to be successful plant surface colonizers (intact and wounded surfaces) and, possibly, antagonists of fruit pathogens.. A second reason of the high polymorphism-rate between *M. fructicola* strains may be the number of high impact mutations in genes putatively involved in the repair or mutation of the genomic sequences. A list of GO terms related to these processes (**supplementary File A.8**) was used to identify 272 annotated genes, and the percentage of these genes showing a putative high impact polymorphism was calculated. 21% of them (57 out of 272) had a mutation of this kind, while the percentage of total genes showing a similar polymorphism was slightly lower (16%, 1,379 out of 8,629).

Uncommonly large genome

The genome of *M. fructicola* was surprisingly large in size, being 26 Mb long. In fact, the most closely-related available genomes (*M. bicuspidata*, *C. auris* and *C. lusitaniae*), are 16 Mb (BioProject PRJNA207846, Riley et al., 2016), 12.5 Mb (BioProjects PRJNA342691 and PRJNA267757, Chatterjee et al., 2015) and 11.9 Mb (BioProject PRJNA12753, Butler et al., 2009), respectively. The most probable explanation for such a genome size seemed to be a whole genome duplication event. To have evidence of this, we searched the genome for homologs, finding 5,132 genes out of 8,629, all in pairs but for 228, which come in groups of three or more copies. This is a high degree of homology, since in the genomes of *M. bicuspidata*, *C. auris*, and *C. lusitaniae* we found only 71, 69 and 56 homologous genes, respectively.

Ordinarily, after a whole-genome duplication event in yeasts, most of the duplicates of genes situated in low mutation regions are lost, while the ones situated in rapidly evolving regions accumulate mutations and differentiate themselves from their homologs (Fares et al., 2017). We compared the average number of polymorphisms identified between strains 277 and AP47 on homologs and single-copy genes, finding that the first group of genes has a variant rate of 1/65 bases, while for the second group this value is of 1/68. Since divergence between gene copies can also happen at the

expression level, so that each copy can be expressed in a different situation and accumulate mutations useful for a specific environmental condition (Fares et al., 2017), the variant rate in the promoters was also checked. Among the promoters of the homologous genes, the average variant rate is of 1/37 bases, while in the single-copy gene promoters it is of 1/45.

Despite the low difference in the mutation rate of single-copy and homologous genes, particularly in the proper gene sequence and not in the promoters, we believe that the available data strengthen the hypothesis of a whole-genome duplication event being responsible for the large genome of *M. fructicola*. This is due principally to the fact that nearly all the homologous genes come in pairs, with only 228 having more than one homolog. The sequencing of other *M. fructicola* strains will undoubtedly be critical to gain further insight on the reasons of this yeast's large genome.

It should be noted that the strain AP47 has SNPs spread along all the contigs of strain 277 (**Figure A.2**). This seems to indicate that the whole genome duplication event occurred in AP47 as well, and that the strains share a common ancestor. This was observed despite the high mutation rate between the strains.

The genomes of the Metschnikowia spp. present in **Table A.5** were downloaded from ncbi, to look for others whole-genome duplication events. Since *M. bicuspidata* is the only one of these species to have been fully annotated, it was impossible to look for the whole genome duplication event as has been done with *M. fructicola*. Therefore, we blasted both the transcriptomes of *M. fructicola* and *M. bicuspidata* on all the considered genomes, counting how many of these had matches on different contigs: even if not every transcript had a match, the result of the analysis gave us an idea of the level of homology inside the genomes of interest. In *M. fructicola*, 75% of the transcripts had matches on more than one contig. Furthermore, of the *M. bicuspidata* transcripts with a match on the *M. fructicola* genome, 58% had a match on more than one contig. On the contrary, none of the other analysed genomes reached a percentage

	Matched transcripts		Homology level	
	M. fructicola transcriptome	*M. bicuspidata* transcriptome	*M. fructicola* transcriptome	*M. bicuspidata* transcriptome
M. aberdeeniae (GCA_002370615.1)	39.16%	38.89%	3.64%	4.93%
M. arizonensis (GCA_002370875.1)	33.97%	33.3%	4.74%	7.15%
M. bicuspidata (PRJNA207846)	**67.96%**	**100%**	**3.27%**	**9.23%**
M. bowlesiae (GCA_002370295.1)	36.77%	38.02%	5.55%	7.26%
M. cerradonensis (GCA_002370635.1)	37.66%	38.51%	6.98%	8.1%
M. colocasiae (GCA_002370175.1)	39.89%	41.32%	4.71%	6.55%
M. continentalis (GCA_002370835.1)	37.46%	38.05%	8.42%	9.37%
M. cubensis (GCA_002374405.1)	38.3%	38.98%	6.51%	8.53%
M. dekortorum (GCA_002374455.1)	36.46%	38%	5.02%	6.99%
M. drakensbergensis (GCA_002370475.1)	39.02%	40.16%	4.1%	5.25%
M. fructicola	**100%**	**66.52%**	**74.13%**	**58.23%**
M. hawaiiensis (GCA_002370325.1)	40.06%	40.74%	7.52%	9.71%
M. hibisci (GCA_002374725.1)	31.71%	29.57%	3.4%	5.91%
M. kamakouana (GCA_002374535.1)	38.86%	39.3%	3.67%	5.58%
M. lochheadii (GCA_002370915.1)	36.49%	36.3%	7.21%	9.44%
M. matae (GCA_002370695.1)	35.07%	35.12%	7.93%	9.56%
M. mauinuiana (GCA_002374555.1)	38.63%	39.59%	7.47%	9.04%
M. proteae (GCA_002370515.1)	39.65%	40.57%	3.98%	5.83%
M. santaceciliae (GCA_002374485.1)	38.08%	38.74%	6.57%	8.4%
M. shivogae (GCA_002374645.1)	39.85%	40.19%	3.63%	5.33%
M. similis (GCA_002370765.1)	36.93%	38.15%	5.3%	7.5%

Table A.5: Homology level in different *Metschnikowia spp.* genomes. The table is divided in two sections. The left section (Matched transcripts) shows the percentage of *M. fructicola* or *M. bicuspidata* transcripts having a match when blasted on the genome of various *Metschnikowia spp.* The homology level on the right section shows the percentage of matched transcripts which also have a second match on another contig.

of transcripts mapping on different contigs of 10%. Based on this data, it seems that the whole-genome duplication event is unique to *M. fructicola*,This data correlates well with the high homology level found in the genome, because a high number of homologous genes is commonly associated with relatively recent whole genome duplication events (Lenassi et al., 2013).

Carbohydrate active enzymes

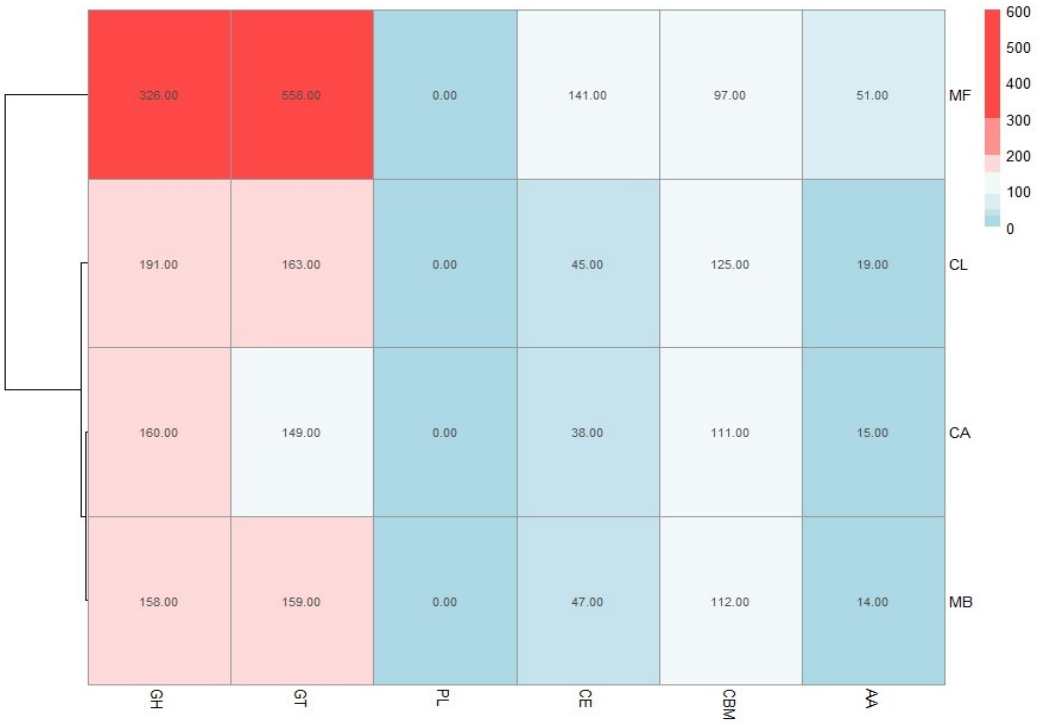

Figure A.3: The number of CAZYenzyme genes in each of the 4 sequenced genomes belonging to the Metschnikowiaceae family: Mf – *Metschnikowia fructicola*, Mb – *Metschnikowia bicuspidate*, CL - *Clavispora lusitaniae*, and CA - *Candida auris*. Different classes of CAZYenzyme genes are designated as GH –Glycoside Hydrolases; GT – Glycosyl Transferases; PL - Polysaccharide Lyases; CE - Carbohydrate Esterases; CBM - Carbohydrate-Binding Modules and AA - Auxiliary Activities. The color reflects the relative number of genes in each of the four species as indicated by the scale in the upper right portion of the figure. 28 genes were included in 2 categories, and therefore the sum of the total of Figure 3 for *M. fructicola* is slightly more than 1,145, which is the reported number of CAZymes.

Plant cell walls consist of a complex network of carbohydrate components, including cellulose, hemicellulose and pectin, as well as a variety of proteins and glycoproteins. These polysaccharides, and other analogous microbial related structural compounds, are targets of carbohydrate-active enzymes (CAZymes) that cleave them into oligomers and simple monomers, which can then be used as nutrients by microorganisms (Cantarel et al., 2009). Bacteria and fungi that are associated with and interact with plants have evolved carbohydrate enzymes strongly linked to the plant environment that these microbes inhabit (Kolton et al., 2013). *M. fructicola* strain 277 MAKER predicted proteins were analysed with CAT (Byung et al., 2010) showing 1,145 putative CAZymes in *M. fructicola* (**Figure A.3**). This represents one of the largest number of potential CAZyme genes that have been reported in Ascomycetes (Amselem et al., 2011). In comparison, the genomes of *Botrytis cinerea* and *Sclerotinia sclerotiorum*, two versatile necrotrophic plant pathogens, contain 367 and 346 putative CAZyme genes, respectively, including 106 and 118 clearly related to cell wall degradation (Apweiler et al., 2001). The impressive repertoire of CAZymes in *M. fructicola* thus may play an important role in its nutritional status and ability to colonize plant surfaces as well as being an effective biocontrol control agent. This role becomes particularly important giving that injured fruit surfaces contain a wide variety of simple and complex carbohydrates that can be consumed by pathogens. Despite different studies characterizing the action of some of these genes ((Jijakli and Lepoivre, 1998 ; Friel et al., 2007), the prospective role of CAZymes in the mechanism of action of microbial antagonists is yet to be fully explored. Among the identified CAZymes in *M. fructicola*, 463 have clear assignments to either glycoside hydrolases (GH) or carbohydrate esterases (CE), all involved in fungal cell wall degradation. Two of the aforementioned genes, unitig185_25 and unitig50_23, have a strong resemblance to MfChi (Genbank accession number: HQ113461.1), a *M. fructicola* chitinase which was shown to inhibit *Monilinia fructicola* and *M. laxa in vitro* and on fruit (*Banani et al., 2015*). A comparison of the number of CAZymes in each of the

four annotated genomes belonging to the Metschnikowiaceae family (Mf – *Metschnikowia fructicola*, Mb – *Metschnikowia bicuspidat*a, CL - *Clavispora lusitaniae*, and CA - *Candida auris*) was conducted (**Figure A.3**). Mb is a fresh-water fish pathogen, while CL and CA are both human pathogens. Results indicated that the *M. fructicola* genome contained a significantly greater variation and number of CAZyme genes, including glycoside hydrolase (GH), glycosyl transferases (GT) and carbohydrate-binding modules (CBM) family genes (**Figure A.3** and **supplementary Table SA.2).** The Mf genome contained several unique CAZymes involved in the metabolism of glucans, arabinose, and rhamnogalacturonan that are exclusively associated with terrestrial plant hemicellulose.

M. fructicola response to *P. digitatum* and to grapefruit peel tissue

The current assembly and genome annotation of Mf enabled us to examine the identification of genes associated with the interaction of Mf with either *P. digitatum* or grapefruit peel tissue and determine the genes that are specific to each interaction. The transcriptomic RNAseq libraries of Mf, available from BioProject PRJNA168317 (Hershkovitz et al., 2013), were then analysed. These libraries were constructed from Mf under four different conditions: 1) Mf growing in NYPD broth (control), 2) Mf in contact with *P. digitatum* (Pd) mycelium for 24h, 3) Mf in contact with *P. digitatum* (Pd) mycelium for 48h, and 4) Mf in contact with grapefruit peel for 24h.

The analysis of DEGs indicated that gene expression in Mf cells that were in contact with fruit peel tissue or had no contact with fruit tissue (control), was more similar to each other than to gene expression in Mf cells that were in contact with *P. digitatum* mycelia. In total, 2,588 DEGs were identified among Mf cells in contact or not in contact with citrus fruit, peel tissue, and Mf cells that were in contact with *P. digitatum* mycelium (**supplementary Table SA.3**). The DEGs could be grouped into three different co-expressed clusters (**Figure A.4A** and **A.4B**).

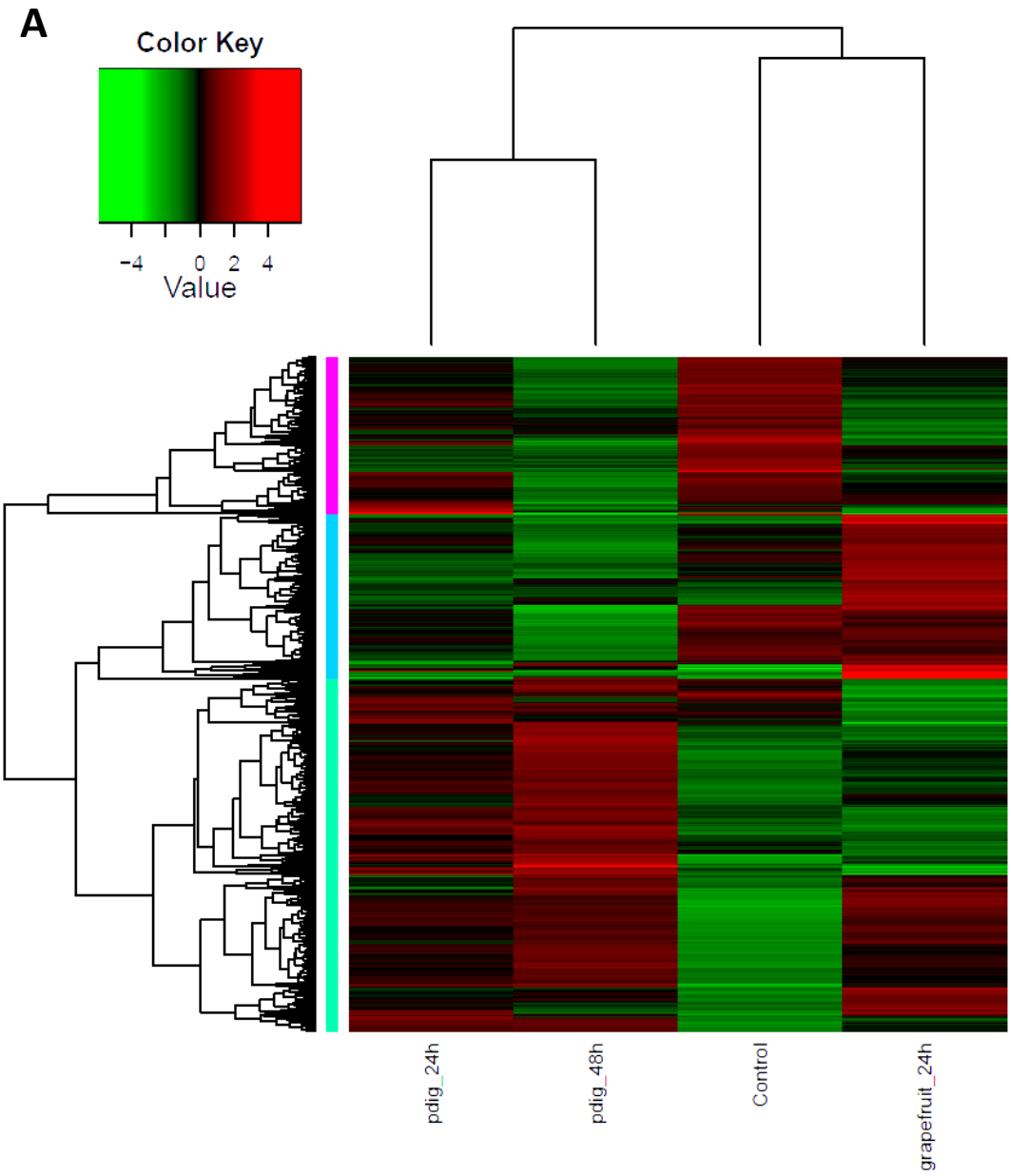

Figure A.4: (A) heatmap and expression profile of differentially expressed genes in *Metschnikowia fructicola* (Mf) grown on different substrates. Three clusters were identified. Cluster 1 – genes with higher expression level when Mf was grown in contact with *Penicillium digitatum* (Pd). Cluster2 – genes with higher expression level when Mf was grown in NYPD broth (control). Cluster 3 - genes with higher expression level when in Mf was grown in contact with grapefruit peel.

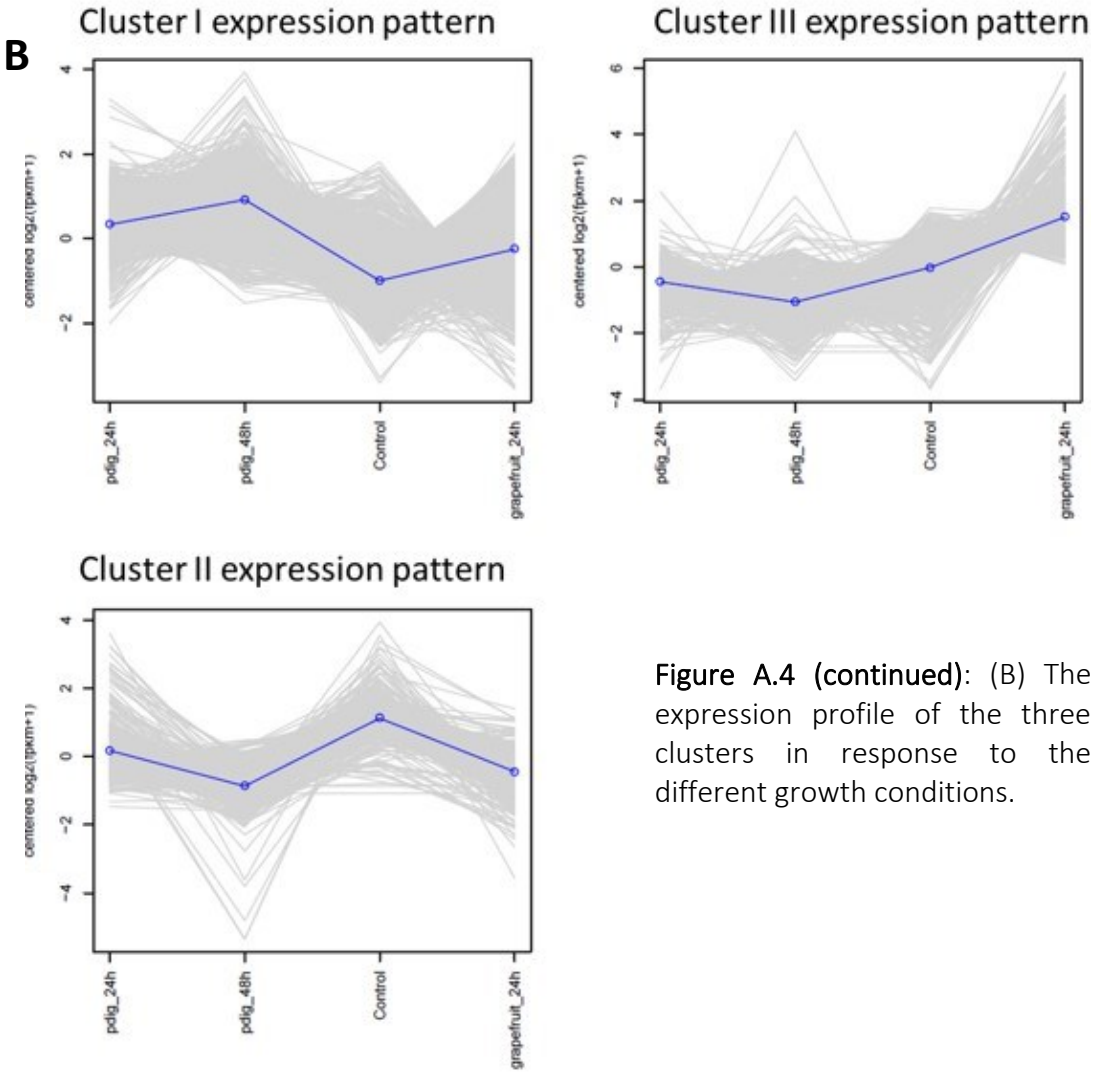

Figure A.4 (continued): (B) The expression profile of the three clusters in response to the different growth conditions.

Cluster1 genes were more highly expressed during contact with *P. digitatum* (Pd) mycelia, relative to cells grown in NYPD broth (control) or on grapefruit peel tissue. We have found 1353 such genes (while only 153 unigenes were found in the previous analysis when using de-novo transcriptome assembly). Cluster 2 genes were more highly expressed in Mf grown in NYPD broth (control) than they were when Mf was in contact with either grapefruit peel tissue or *P. digitatum* mycelium (total of 635 genes). Cluster 3 genes exhibited higher levels of expression when Mf cells were in

contact with grapefruit peel tissue, rather than when grown in NYPD broth (control) or in contact with *P. digitatum* mycelium (600 genes).

Transcriptomic analysis of CAZyme expression levels in *M. fructicola* during its interaction with grapefruit peel tissue or *P. digitatum* mycelium when cultured in a PDB medium revealed a high level of CAZyme gene expression when the yeast was placed in wounded fruit tissue (**Figure A.5**). These results suggest that CAZyme genes may play an important role in the adaptation of *M. fructicola* to a fruit environment.

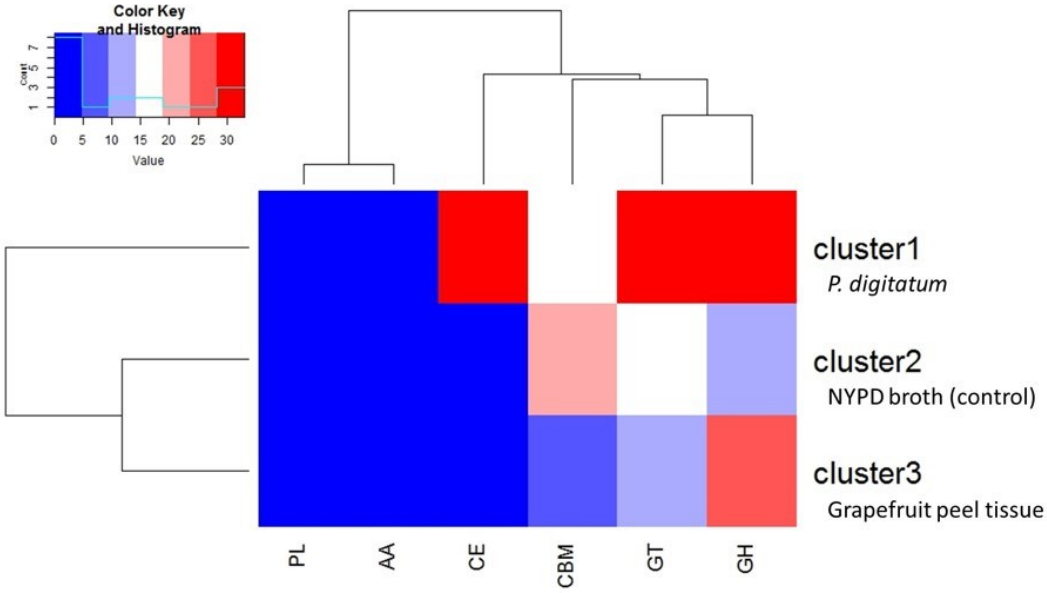

Figure A.5: Diagram showing the relative number of genes within each class of CAZymes in each cluster. Cluster1 genes were more highly expressed, relative to cells grown in NYPD broth (control) or on grapefruit peel tissue, when the yeast cells were in contact with *Penicillium digitatum* (Pd) mycelium. Cluster 2 genes were more highly expressed in Mf grown in NYPD broth (control). Cluster 3 genes exhibited higher levels of expression when Mf cells were grown in contact with grapefruit peel tissue. Cluster 3 show also the highest quantity of CAZY enzymes. The color reflects the relative number of genes in each of the clusters as indicated by the scale in the upper left portion of the figure.
Different classes of CAZYenzyme genes are designated as GH –Glycoside Hydrolases; GT – Glycosyl Transferases; PL - Polysaccharide Lyases; CE - Carbohydrate Esterases; CBM - Carbohydrate-Binding Modules and AA - Auxiliary Activities.

Secondary metabolite clusters present in *M. fructicola*

Secondary metabolite cluster type	Transcripts of Mf found in cluster	location
Terpene cluster	unitig147_4 unitig147_5 unitig147_6 unitig147_7 unitig147_8 unitig147_9 unitig147_10	unitig147 15287 - 36642
Terpene cluster	unitig50_207 unitig50_208 unitig50_209 unitig50_210 unitig50_211 unitig50_212 unitig50_213 unitig50_214	unitig50 578895 - 600250

Table A. 6: Secondary metabolites clusters identified with antiSMASH (Weber et al., 2015) software.

The sequence of the *M. fructicola* genome revealed that this yeast possesses several secondary metabolite (SM) genes. SMs are known to play an important role in the virulence of many plant pathogens (Namdeo, 2007), but limited knowledge is available about the SM repertoire present in *M. fructicola*. Using antiSMASH (Weber et al., 2015) software, the *M. fructicola* genome was analysed for the presence of secondary metabolite clusters or homologs of these genes present in related fungi.

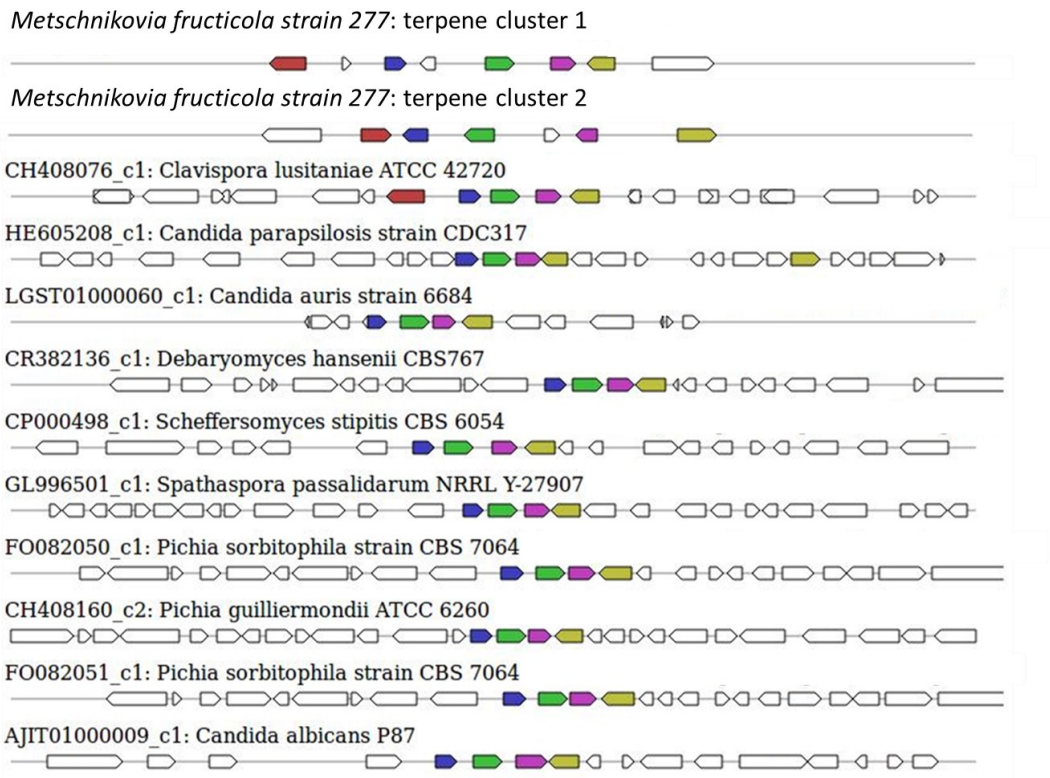

Figure A. 6: Secondary metabolite clusters producing terpene and their homology with terpene-synthesis clusters in closely related fungi. The terpene synthesis clusters as indicated by antiSMASH3.0 software. The uppermost clusters represent the terpene-synthesis clusters in *M. fructicola* while the terpene-synthesis cluster from other yeasts are shown below.

Twenty-six SM gene clusters were identified in *M. fructicola*, four of which are highly conserved in yeast and other fungi. The remaining 22 clusters could only be designated as putative clusters as similar clusters could not be identified in other fungal genomes using the ClusterFinder algorithm (Cimermancic et al., 2014). These 22 potential clusters included putative saccharide and fatty acid biosynthetic clusters. The analysis of secondary metabolite genes indicated that *M. fructicola* is capable of producing small, potentially bioactive molecules. Two of the identified clusters (**Table A.6** and **Figure A.6**) code for the production of a terpene that is conserved within *Candida* species. Terpenoid compounds are known to play a significant role in yeast

antimicrobial defence mechanism (Hyldgaard et al., 2012). The isoprenoid backbones of these compounds are synthesized by terpene synthases (TSs). The classification of various terpene synthases and their catalytic mechanisms have been recently reviewed (Gao et al., 2012). Although terpenoid SMs have not been previously reported in *M. fructicola*, the genome sequence clearly possesses two gene sequences that encode squalene/phytoene synthases: the transcripts unitig50_211 and unitig147_7.

YAP gene expression in *M. fructicola*

Systematic Name in *Saccharomyces cerevisiae*	**Homologue in Mf Genome**	**Gene Name**	**Alias(es)**	**Description**
YDR259C	Not found	YAP6	HAL7	Basic leucine zipper (bZIP) transcription factor
YDR423C	Not found	CAD1	YAP2	AP-1-like basic leucine zipper (bZIP) transcriptional activator
YGR241C	unitig192_208	YAP1802		Protein of the AP180 family, involved in clathrin cage assembly
YHL009C	unitig142_42 unitig187_66	YAP3		Basic leucine zipper (bZIP) transcription factor
YHR161C	Not found	YAP1801		Protein of the AP180 family, involved in clathrin cage assembly
YIR018W	Not found	YAP5		Basic leucine zipper (bZIP) iron-sensing transcription factor
YJR005W	unitig146_71 unitig192_37	APL1	YAP80	Beta-adaptin

YJR058C	unitig122_58 unitig50_345	APS2	YAP17	Small subunit of the clathrin-associated adaptor complex AP-2
YLR120C	unitig104_2 unitig150_6 unitig193_34 9unitig32_12	YPS1	aspartyl protease	Aspartic protease
YLR170C	unitig196_23 4	APS1	YAP19	Small subunit of the clathrin-associated adaptor complex AP-1
YML007W	Not found	YAP1	PDR4,DNA-binding transcription factor YAP1	Basic leucine zipper (bZIP) transcription factor
YOL028C	Not found	YAP7		Putative basic leucine zipper (bZIP) transcription factor
YOR028C	Not found	CIN5	YAP4, HAL6	Basic leucine zipper (bZIP) transcription factor of the yAP-1 family
YPL259C	unitig105_13 unitig193_25 1	APM1	YAP54	Mu1-like medium subunit of the AP-1 complex
YPR199C	Not found	ARR1	ACR1, YAP8	Transcriptional activator of the basic leucine zipper (bZIP) family

Table A.7: Yap family genes and homologs identified in the genome of *M. fructicola*.

The Yap protein family plays a role in cellular response to oxidative stress (Rodrigues-Pousada et al., 2010) and *M. fructicola* has been demonstrated to have a high tolerance to oxidative stress (Macarisin et al., 2010). An analysis of YAP genes

in the *M. fructicola* genome revealed the presence of 14 YAP genes (**Table A.7**). In comparison, 7 YAP genes were found in *C. albicans* (BioProjects PRJNA14005 and PRJNA10701), *C. auris* (BioProjects PRJNA342691 and PRJNA267757) and *M. bicuspidata* (BioProject PRJNA207846), while *C. lusitaniae* (BioProject PRJNA12753) had 6. YAP genes are important for resistant to oxidative stress (Macarisin et al., 2010). a feature that could possibly play a role in the ecological fitness and antagonistic activity of *M. fructicola*.

Pulcherrimin cluster analysis

Pulcherrimin is a *M. fructicola* metabolite of major interest, since it is involved in the biocontrol action of this yeast (Saravanakumar et al., 2008). The genes responsible for the biosynthesis of this siderophore were successfully identified only in *B. subtilis* (Randazzo et al., 2016), and an analysis of orthology with proteinortho and blast showed no homology between the *B. subtilis* pulcherrimin gene cluster and the proteins predicted in *M. fructicola*. It is probable that the *B. subtilis* and *M. fructicola* genes involved in pulcherrimin biosynthesis are the product of different evolutionary processes.

CONCLUSIONS

The genomes of two strains of *M. fructicola* (277 and AP47) were sequenced, assembled and compared. The comparison of the two genomes sequences indicated a very high rate of mutation, even though it will be necessary to sequence additional strains to establish if the average mutation rate in *M. fructicola* is intrinsically high, or if the mutation rate identified in the present study is related to the geographical origin and fruit host in which they evolved. The genome size (~26 Mb) of both *M. fructicola* strains, as well as the rate of mutation, may suggest that *M. fructicola* could undergo genomic changes in order to adapt to plant surfaces, tolerate various environmental

stresses and survive under restricted nutritional resources. Its adaptation to plant environment can also be explained by the presence of a relatively large number of secondary metabolites clusters, YAP and CAZymes related genes in the genome.

Another interesting result was the discovery of 1,145 putative CAZymes in the *M. fructicola* genome. These genes could be the target of studies aimed to identify enzymes able to control fungal diseases *in vivo*, to evaluate their potential use as treatments for fruits and plants.

MATERIALS AND METHODS

DNA extraction

Metschnikowia fructicola, Strain 277, (Kurtzman et al., 2001) was grown in NYDP (nutrient broth (8 g l^{-1}), yeast extract (5 g l^{-1}), D-glucose (10 gl^{-1}) and chloramphenicol (250 mg l^{-1}). One ml of the yeast cell suspension was aseptically transferred from 24 h old starter culture to 250 ml Erlenmeyer flasks and place on an orbital shaker at 160 rpm for 24 h at 26°C. Yeast cells were pelleted by centrifugation at 6,000 rpm, washed twice with sterile distilled water, re-suspended in sterile water to initial volume and the cell suspension concentration was adjusted to 1 × 108 cells ml^{-1}.

M. fructicola strain AP47 was isolated from the carposphere of an apple grown in Piedmont, Northern Italy (Zhang et al., 2010). The strain was stored in tubes of Potato Dextrose Agar and 50 mg/L streptomycin at 4 °C. Suspensions of *M. fructicola* AP47 (5×10^5 cells/mL) were inoculated in 500 mL Potato Dextrose Broth (PDB, Difco) and incubated on a rotary shaker (180 rpm) at 24 °C for 4 days. Yeast mass was filtered from the culture, frozen in liquid nitrogen and DNA was extracted from 1 g frozen tissue. The final DNA preparation was incubated overnight at room temperature in 490 µl of Tris-EDTA (TE) buffer and 10 µl of DNase-free RNase (10 µg/ml), followed by phenol-chloroform extraction and isopropanol precipitation. Finally, DNA was

resuspended in 30µl TE buffer. DNA concentration and purity were checked by a spectrophotometer (Nanodrop 2000, Thermo Scientific, Wilmington, USA), and the DNA integrity was analysed by agarose gel electrophoresis (data not shown).

Sequencing

Strain 277 was sequenced on the Pacific Biosciences (PacBio) *RS* II Sequencer, as previously described (Hoffman et al., 2013; Pirone-Davies et al., 2015). Specifically, we prepared the library using 10 µg of genomic DNA, that was sheared to a size of 20kb fragments by g-tubes (Covaris, Inc., Woburn, MA) according to the manufacturer's instruction. The SMRTbell 20-kb template library was constructed using DNA Template Prep Kit 1.0 with the 20-kb insert library protocol (Pacific Biosciences; Menlo Park, CA, USA). Size selection was performed with BluePippin (Sage Science, Beverly, MA). The library was sequenced using the P6/C4 chemistry on 24 single-molecule real-time (SMRT) cells (8 with BluePippin and 16 without), with a 240-min collection protocol along with stage start.

The genome of *M. fructicola* AP47 was sequenced at the Genomics Platform of the Parco Tecnologico Padano using the Illumina MiSeq technology. Two paired-ends were prepared using Nextera XT DNA Sample Preparation Kit, following the manufacturer's instructions. Two paired-end (PE) libraries were prepared: PE1 with overlapping paired-end reads and PE2 with non-overlapping paired-end reads. One mate pair library was also prepared, using Nextera Mate Pair Sample Preparation Kit and following the manufacturer's instructions. Libraries were purified by AMPure XP beads and normalized to ensure equal library representation in the pools. Equal volumes of libraries were diluted in the hybridization buffer, heat denatured and sequenced. Standard phi X control library (Illumina) was spiked into the denatured HCT 116 library. The libraries and phi X mixture were finally loaded into a MiSeq 250 and MiSeq 300-Cycle v2 Reagent Kit (Illumina). Base calling was performed using the Illumina pipeline software. PE1 was composed of 2,1 Gb (330 mean insert

size, 43 % GC, 35% dupl level). PE2 was composed of 846 Mb (132 mean insert size, 45% GC, 12/ duplication level).

All the paired end sequences were trimmed with Trimmomatic v. 0.36 (Bolger et al., 2014) and cleaned with sickle v. 1.33 (Joshi and Fass, 2011) (**Table A.2**). The mate pair sequences were trimmed and cleaned with TrimGalore v. 0.4.2 (*https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/*).

The genome of *M. fructicola* AP47 was assembled at first with a *de novo* approach, using SPAdes (Bankevich et al., 2012), and then with a reference guided approach using IMR-DENOM (*http://mus.well.ox.ac.uk/19genomes/IMR-DENOM/*), with the strain 277 as the reference.

Assembly

Analysis of the sequence reads was implemented by using SMRT Analysis 2.3.0. The best *de novo* assembly was established with the PacBio Hierarchical Genome Assembly Process HGAP3.0 program (Chin et al., 2013) using the continuous-long-reads from the four SMRT cells, which contained the longest subreads, with a minimum subread length cutoff of 5000kb and target coverage of 20X. The resulting HGAP unique contigs (unitigs) were blasted against each other to identify smaller unitigs that show complete overlapping with other larger unitigs. These smaller unitigs were removed from the analysis. Afterwards the improved consensus sequence was uploaded in SMRT Analysis 2.3.0. and polished with Quiver using all 24 SMRT cells (Chin et al., 2013).

In total 24 SMRT cells were used, resulting in 93 contigs with 439X average genome coverage. The longest contig comprised 2,548,689 bp.

Transcriptome assembly, gene prediction and functional annotation

RNAseq from previous analysis (Hershkovitz et al., 2013) was used to assemble and predict transcribed regions in the Mf genome. Overall, 6,150 transcripts were

identified based on tophat, cufflinks and bowtie2 pipeline as described in (Langmead and Salzberg, 2012).

The transcriptome data, together with the transcripts and proteins sequences available on NCBI for *M. fructicola*, *M. biscuspidata*, *C. auris* and *C. lusitaniae*, were used to train the gene predictor SNAP (*http://korflab.ucdavis.edu/software.html*), following the suggested procedure (*http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/MAKER_Tutorial_for_GMOD_Online_Training_2014*). The augustus gene predictor (*http://augustus.gobics.de/*) was trained with the WebAUGUSTUS web service (Stanke and Morgenstern, 2005), using as data the sequence of the 6,150 transcripts identified with the RNA seq.

SNAP and augustus were then used as a part of the MAKER software (Cantarel et al., 2008) to conduct the gene prediction in the genome. The evidence used were the 6,150 transcripts discovered with the RNA seq and the transcripts and proteins sequences available on NCBI for *M. fructicola*, *M. biscuspidata*, *C. auris* and *C. lusitaniae*. The transcripts not coming from *M. fructicola* were included in the MAKER control files as "altest" evidence, which is specifically used for data from species related to the target genome and not from the target itself. The repeat library was constructed following the Basic protocol (*http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/Repeat_Library_Construction--Basic*), and MAKER was launched using the option "correct_est_fusion" in the control files and "-fix-nucleotides" in the command line. MAKER produced a gene coordinates gff3 file, which was used to extract the CDSs from the genome in order to translate them with BioPython (Cock et al., 2009) using the Alternative Yeast Nuclear Code, obtaining the protein sequences. Some of the predicted genes had putative CDSs, which did not start with a start codon and/or did not end with a stop one, and were therefore discarded, with the following exceptions: i) genes missing the stop codon, localized on the plus filament, which were the last gene of their contig; ii)

genes missing the stop codon, localized on the minus filament, which were the first gene of their contig; iii) genes missing the start codon, localized on the plus filament, which were the first gene of their contig; iv) genes missing the start codon, localized on the minus filament, which were the last gene of their contig. The genes of these categories were kept as partial genes.

The proteins were annotated with Blast2GO and Interproscan, using as blast database the fungal fraction of uniprot and swissprot databases (Uniprot Consortium, 2017)].

The CAT webservice was used to find Pfam modules (Robert et al., 2016) in the proteins and assign them CAZy families.

Proteinortho v. 5.16 was used to look for homologous proteins in the proteomes of *M. fructicola* 277, *C. auris* (BioProjects PRJNA342691 and PRJNA267757), *M. bicuspidata* (BioProject PRJNA207846) and *C. lusitaniae* (BioProject PRJNA12753).

Gene expression analysis

RNAseq analysis was done using RNAseq data from previous research (Hershkovitz et al., 2013). The RNAseq data number SRA054245 was download from SRA database in NCBI. The RNAseq data was mapped using bowtie (Langmead et al., 2009). Expression quantification was estimated using RSEM software (Liand Dewey, 2011). Differential expression analysis was done using edgeR Bioconductor package (Robinson et al., 2010).Clustering was done using K-mean cluster analysis (Basu et al., 2002) differentialy expressed genes threshold was FDR <0.05 (Benjamini et al., 1995) and log fold changes greater than 1 or smaller than -1.

Phylogenetic tree

All raw-data sequences of *Metschnikowia* species (Lachance et al., 2015) were downloaded from NCBI using SRAtoolkit (Leinonen et al., 2015) from BioProject ID PRJNA312754. The phylogenetic tree was constructed with an assembly and

alignment-free method of phylogeny reconstruction from next-generation sequencing data (Fan et al., 2015).

To place the whole-genome duplication event in the three, we downloaded the genomes of all the considered species, and we used them as databases to blast the full transcriptomes of *M. fructicola* and *M. bicuspidata* (**Table A.5**), using blastall v. 2.2.26 with default parameters. We then calculated the percentage of transcripts having a match, and, inside this fraction, the percentage of transcripts having a match on at least 2 contigs.

Genome comparison with *M. fructicola* strain AP47

A SNP calling approach was followed, using bwa mem (Li and Durbin ., 2009) to map Illumina reads of the strain AP47 of *M. fructicola* on the assembly of the strain 277. After using samtools view and samtools sort (Li et al., 2009) to obtain a sort.bam file, the following pipeline was used as described by Li et al. (2011)for the SNP calling:

samtools mpileup -guf reference.fa AP47.sort.bam | bcftools view -cg -| vcfutils.pl varFilter -D 200 -Q 20 - > file.vcf

The file AP47.sort.bam was obtained by merging the data from the two Illumina libraries with samtools merge.

The genome of the strain 277 and the gff3 and protein fasta files obtained with MAKER, were used to build a SnpEff (Cingolani et al., 2012) database, and the tool "snpeff eff" was used to evaluate the effect of the homozygous SNPs of the strain AP47. Since *M. fructicola* is a haploid organism, heterozygous SNPs were probably mistakes. The Alternative Yeast Nuclear Code was used to evaluate the effect of missense SNPs on protein sequences.

Analysis of the polymorphisms-related genes.

The number of high impact polymorphisms in genes characterized by gene onthology terms present in **supplementary File A.8** was calculated and compared to the genome

average. **Supplementary File A.8** was obtained by selecting all GO terms including the word "repair" or "mutation", and then removing manually undesired terms (e.g. "cell wall repair").

Analysis of the D1/D2 region

The primers NL-1 (GCATATCAATAAGCGGAGGAAAAG) and NL-4 (GGTCCGTGTTTCAAGACGG) (O'Donnel, 1993), used by Kurtzman and Robnett (1998) to amplify the D1/D2 region in *S. cerevisiae*, were blasted on the *M. pulcherrima* sequences available on NCBI, so to identify the D1/D2 region. The partial sequence of the large subunit ribosomal RNA gene of *M.pulcherrima* culture-collection CBS:2256 (GenBank: KY108498.1) was therefore downloaded, and blasted on the *M. fructicola* strain 277 genome. We then proceeded to identify the SNPs present in that region in the strains 277 and AP47, looking at both the homozygous and heterozygous SNPs. The blast version used was blastall v. 2.2.26.

Whole-genome duplication hypothesis

Proteinortho v. 5.16 was used to look for homologous proteins in the proteomes of *M. fructicola* 277, *C. auris* (BioProjects PRJNA342691 and PRJNA267757), *M. bicuspidata* (BioProject PRJNA207846) and *C. lusitaniae* (BioProject PRJNA12753). The variant rate in single-copy and homologous genes was calculated, and the same was done in their promoters.

The promoter analysis was performed considering as promoter the 1000 bases preceding the genes in the genome, or the 1000 bases following the genes when these were on the antisense strand.

YAP genes analysis

The protein sequence of various Yap genes was downloaded from www.yeastgenome.org, and analysed with Proteinortho v. 5.16 (Lechner et al., 2011),

looking for homologs in the proteins predicted for *M. fructicola* strain 277 and in the proteomes of *Candida albicans* (BioProjects PRJNA14005 and PRJNA10701), *C. auris* (BioProjects PRJNA342691 and PRJNA267757), *M. bicuspidata* (BioProject PRJNA207846) and *C. lusitaniae* (BioProject PRJNA12753).

Secondary metabolites cluster prediction

Secondary metebolites clustering was predicted using antiSMASH website (Weber et al., 2015)

Pulcherrimin gene cluster analysis

The proteins involved in pulcherrimin biosynthesis in *B. subtilis* (YVNB, YVNA, YVMC, YVMB, YVMA, CYPX; Randazzi et al., 2016) were downloaded from NCBI and used in a proteinortho v. 5.15 analysis with the MAKER predicted proteins of *M. fructicola*, with default parameters. The *B. subtilis* genes of interest were also blasted with blastp (blastall v. 2.2.26) against the predicted proteome of *M. fructicola*, using an e-value threshold of 10^{-5}.

ACKNOWLEDGMENTS

Work carried out with a contribution of the LIFE financial instrument of the European Union for the project "Low pesticide IPM in sustainable and safe fruit production" (Contract No. LIFE13 ENV/HR/000580). The Authors wish to thank Prof. Alberto Acquadro, University of Torino for his useful suggestion about bioinformatics analysis.

REFERENCES

- Amselem, J., Cuomo, C. A., Van Kan, J. A., Viaud, M., Benito, E. P., Couloux, A., ... & Fournier, E. (2011). Genomic analysis of the necrotrophic fungal pathogens

Sclerotinia sclerotiorum and *Botrytis cinerea*. *PLoS Genetics*, *7*(8), e1002230. https://doi.org/10.1371/journal.pgen.1002230

- Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., ... & Durbin, R. (2001). The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research*, *29*(1), 37-40. https://doi.org/10.1093/nar/29.1.37
- Banani, H., Spadaro, D., Zhang, D., Matic, S., Garibaldi, A., & Gullino, M. L. (2015). Postharvest application of a novel chitinase cloned from *Metschnikowia fructicola* and overexpressed in *Pichia pastoris* to control brown rot of peaches. *International Journal of Food Microbiology*, *199*, 54-61. https://doi.org/10.1016/j.ijfoodmicro.2015.01.002
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pyshkin, A. V. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455-477. https://doi.org/10.1089/cmb.2012.0021
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289-300. https://doi.org/10.2307/2346101
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2008). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research*, *37*(1), 233-238. https://doi.org/10.1093/nar/gkn663
- Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., ... & Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188-196. https://doi.org/10.1101/gr.6743907
- Cerveau, N., & Jackson, D. J. (2016). Combining independent *de novo* assemblies optimizes the coding transcriptome for nonconventional model eukaryotic organisms. *BMC Bioinformatics*, *17*(1), 525. https://doi.org/10.1186/s12859-016-1406-x

- Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., ... & Turner, S. W. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods*, *10*(6), 563. https://doi.org/10.1038/nmeth.2474
- Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Brown, L. C. W., Mavrommatis, K., ... & Birren, B. W. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell*, *158*(2), 412-421. https://doi.org/10.1016/j.cell.2014.06.034
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., ... & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80-92. https://doi.org/10.4161/fly.19695
- Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., ... & De Hoon, M. J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*(11), 1422-1423. https://doi.org/10.1093/bioinformatics/btp163
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674-3676. https://doi.org/10.1093/bioinformatics/bti610
- Droby, S., Wisniewski, M., Macarisin, D., & Wilson, C. (2009). Twenty years of postharvest biocontrol research: is it time for a new paradigm? *Postharvest Biology and Technology*, *52*(2), 137-145. https://doi.org/10.1016/j.postharvbio.2008.11.009
- Droby, S., Wisniewski, M., Teixidó, N., Spadaro, D., & Jijakli, M. H. (2016). The science, development, and commercialization of postharvest biocontrol products. *Postharvest Biology and Technology*, *122*, 22-29. https://doi.org/10.1016/j.postharvbio.2016.04.006
- Drori, E., Levy, D., Smirin-Yosef, P., Rahimi, O., & Salmon-Divon, M. (2017). CircosVCF: circos visualization of whole-genome sequence variations stored in VCF files. *Bioinformatics*, *33*(9), 1392-1393. https://doi.org/10.1093/bioinformatics/btw834
- Drozdova, P. B., Tarasov, O. V., Matveenko, A. G., Radchenko, E. A., Sopova, J. V., Polev, D. E., ... & Dobrynin, P. V. (2016). Genome sequencing and comparative

analysis of Saccharomyces cerevisiae strains of the Peterhof genetic collection. *PloS One*, *11*(5), e0154722. https://doi.org/10.1371/journal.pone.0154722

- Fan, H., Ives, A. R., Surget-Groba, Y., &Cannon, C. H. (2015). An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. *BMC Genomics*. 16, 522. https://doi.org/10.1186/s12864-015-1647-5
- Fares, M. A., Sabater-Munoz, B., &Toft, C. (2017). Genome Mutational and Transcriptional Hotspots Are Traps for Duplicated Genes and Sources of Adaptations. *Genome Biology and Evolution,* 9, 1229-1240. https://doi.org/10.1093/gbe/evx085.
- Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., ... & Gough, J. (2016). InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Research*, *45*(D1), 190-199. https://doi.org/10.1093/nar/gkw1107
- Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., ... & Salazar, G. A. (2015). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, *44*(1), 279-285. https://doi.org/10.1093/nar/gkv1344.
- Friel, D., Pessoa, N. M. G., Vandenbol, M., & Jijakli, M. H. (2007). Separate and combined disruptions of two exo-beta-1,3-glucanase genes decrease the efficiency of *Pichia anomala* (strain K) biocontrol against *Botrytis cinerea* on apple. *Molecular Plant-Microbe Interactions*, 20, 371-379. https://doi.org/10.1094/MPMI-20-4-0371
- Galhardo, R. S., Hastings, P. J., & Rosenberg, S. M. (2007). Mutation as a stress response and the regulation of evolvability. *Critical Reviews in Biochemistry and Molecular Biology*, 42, 399-435.
- Gao, Y., Honzatko, R. B., & Peters, R. J. (2012). Terpenoid synthase structures: a so far incomplete view of complex catalysis. *Natural Products Reports,* 29, 1153-1175. https://doi.org/10.1039/c2np20059g
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... & Chen, Z. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644. https://doi.org/10.1038/nbt.1883
- Hershkovitz, V., Ben-Dayan, C., Raphael, G., Pasmanik-Chor, M., Liu, J., Belausov,

E., & Droby, S. (2012). Global changes in gene expression of grapefruit peel tissue in response to the yeast biocontrol agent *Metschnikowia fructicola*. *Molecular Plant Pathology,* 13, 338-349. https://doi.org/10.1111/j.1364-3703.2011.00750.x

- Hershkovitz, V., Sela, N., Taha-Salaime, L., Liu, J., Rafael, G., Kessler, C., ... & Droby, S. (2013). De-novo assembly and characterization of the transcriptome of *Metschnikowia fructicola* reveals differences in gene expression following interaction with *Penicillium digitatum* and grapefruit peel. *BMC genomics*, *14*(1), 168. https://doi.org/10.1186/1471-2164-14-168
- Hirakawa, M. P., Martinez, D. A., Sakthikumar, S., Anderson, M. Z., Berlin, A., Gujja, S., ... & Berman, J. (2015). Genetic and phenotypic intra-species variation in *Candida albicans*. *Genome Research*, *25*(3), 413-425. https://doi.org/10.1101/gr.174623.114
- Hoffmann, M., Muruvanda, T., Allard, M. W., Korlach, J., Roberts, R. J., Timme, R., ... & Brown, E. W. (2013). Complete genome sequence of a multidrug-resistant *Salmonella enterica* serovar Typhimurium var. 5− strain isolated from chicken breast. *Genome Announcements.*, *1*(6), e01068-13. https://doi.org/10.1128/genomeA.01068-13
- Hyldgaard, M., Mygind, T., & Meyer, R. L. (2012). Essential oils in food preservation: mode of action, synergies, and interactions with food matrix components. *Frontiers in Microbiology*, *3*, 12. https://doi.org/10.3389/fmicb.2012.00012
- Jijakli, M. H., & Lepoivre, P. (1998). Characterization of an exo-beta-1,3-glucanase produced by *Pichia anomala* strain K, antagonist of *Botrytis cinerea* on apples. *Phytopathology*. 88, 335-343. https://doi.org/10.1094/PHYTO.1998.88.4.335
- Joshi, N., & Fass, J. (2011). Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. Available from: github com/najoshi/sickle
- Karabulut, O. A., Smilanick, J. L., Gabler, F. M., Mansour, M., & Droby, S. (2003). Near-harvest applications of *Metschnikowia fructicola*, ethanol, and sodium bicarbonate to control postharvest diseases of grape in central California. *Plant Disease,* 87, 1384-1389. https://doi.org/10.1094/PDIS.2003.87.11.1384
- Karabulut, O., Tezcan, H., Daus, A., Cohen, L., Wiess, B., & Droby, S. (2004). Control of preharvest and postharvest fruit rot in strawberry by *Metschnikowia fructicola*. *Biocontrol Science and Technology*. 14, 513-521. https://doi.org/10.1080/09583150410001682287

- Kolton, M., Sela, N., Elad, Y., & Cytryn, E. (2013). Comparative genomic analysis indicates that niche adaptation of terrestrial Flavobacteria is strongly linked to plant glycan metabolism. *PLoS One*. 8: e76704. https://doi.org/10.1371/journal.pone.0076704
- Kurtzman, C. P., & Droby, S. (2001). *Metschnikowia fructicola*, a new ascosporic yeast with potential for biocontrol of postharvest fruit rots. *Systematic and Applied Microbiology,* 24, 395-399. https://doi.org/10.1371/journal.pone.0076704
- Kurtzman, C. P., & Robnett, C. J. (1998). Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S). ribosomal DNA partial sequences. *Antonie van Leeuwenhoek,* 73, 331-371. https://doi.org/10.1023/A:1001761008817
- Lachance, M. A., Hurtado, E., & Hsiang, T. (2016). A stable phylogeny of the large-spored Metschnikowia clade. *Yeast*. 33, 261-275. https://doi.org/10.1002/yea.3163
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology,* 10, 25. https://doi.org/10.1186/gb-2009-10-3-r25
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods,* 9, 357-359. https://doi.org/10.1038/nmeth.1923
- Larralde-Corona, C. P., Ramírez-González, M. D. S., Pérez-Sánchez, G., Oliva-Hernández, A. A., & Narváez-Zapata, J. A. (2011). Identification of differentially expressed genes in the citrus epiphytic-yeast *Pichia guilliermondii* during interaction with *Penicillium digitatum*. *Biological Control*, 57, 208-214. doi 10.1016/j.biocontrol.2011.02.012
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: Detection of (Co-) orthologs in large-scale analysis. *BMC Bioinformatics.* 12, 124. https://doi.org/10.1186/1471-2105-12-124
- Leinonen, R., Sugawara, H., Shumway, M., & International Nucleotide Sequence Database Collaboration. (2010). The sequence read archive. *Nucleic Acids Research*, *39*(1), 19-21. https://doi.org/10.1093/nar/gkq1019
- Lenassi, M., Gostinčar, C., Jackman, S., Turk, M., Sadowski, I., Nislow, C., ... & Plemenitaš, A. (2013). Whole genome duplication and enrichment of metal cation transporters revealed by de novo genome sequencing of extremely halotolerant black yeast *Hortaea werneckii*. *PLoS One*, *8*(8), e71328. https://doi.org/10.1371/journal.pone.0071328.

- Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 12, 323. https://doi.org/10.1186/1471-2105-12-323
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 27, 2987-2993. https://doi.org/10.1093/bioinformatics/btr509
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 25, 1754-1760. https://doi.org/10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352
- Liu, J., Wisniewski, M., Droby, S., Vero, S., Tian, S., & Hershkovitz, V. (2011). Glycine betaine improves oxidative stress tolerance and biocontrol efficacy of the antagonistic yeast *Cystofilobasidium infirmominiatum*. *International Journal of Food Microbiology,* 146, 76-83. https://doi.org/10.1016/j.ijfoodmicro.2011.02.007
- Macarisin, D., Droby, S., Bauchan, G., & Wisniewski, M. (2010). Superoxide anion and hydrogen peroxide in the yeast antagonist–fruit interaction: A new role for reactive oxygen species in postharvest biocontrol? *Postharvest Biology and Technology,* 58, 194-202. https://doi.org/10.1016/j.postharvbio.2010.07.008
- Massart, S., Perazzolli, M., Höfte, M., Pertot, I., & Jijakli, M. H. (2015). Impact of the omic technologies for understanding the modes of action of Biol. Control agents against plant pathogens. *BioControl*. 60, 725-746. https://doi.org/10.1007/s10526-015-9686-z
- Namdeo, A. (2007). Plant cell elicitation for production of secondary metabolites: a review. *Pharmacognosy Reviews*, 1, 69-79.
- O'Donnell, K. (1993). Fusarium and its near relatives. *The Fungal Holomorph: Mitotic, Meiotic and Pleomorphic Speciation in Fungal Systematics*. Reynolds DR & Taylor JW (Eds). 225–233. CAB International, Wallingford, UK.
- Park, B. H., Karpinets, T. V., Syed, M. H., Leuze M. R., & Uberbacher, E. C. (2010). CAZymes Analysis Toolkit (CAT): Web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy

database. *Glycobiology*. 20, 1574–1584. https://doi.org/10.1093/glycob/cwq106

- Pirone-Davies, C., Hoffmann, M., Roberts, R. J., Muruvanda, T., Timme, R. E., Strain, E., ... & Tsai, Y. C. (2015). Genome-wide methylation patterns in *Salmonella enterica* subsp. enterica serovars. *PloS One*, *10*(4), e0123639. https://doi.org/10.1371/journal.pone.0123639
- Randazzo, P., Aubert-Frambourg. A., Guillot, A., & Auger, S. (2016). The MarR-like protein PchR (YvmB) regulates expression of genes involved in pulcherriminic acid biosynthesis and in the initiation of sporulation in *Bacillus subtilis*. *BMC Microbiol*ogy 16, 190. https://doi.org/10.1186/s12866-016-0807-3
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26, 139-140. https://doi.org/10.1093/bioinformatics/btp616
- Rodrigues-Pousada, C., Menezes, R. A., & Pimentel, C. (2010). The Yap family and its role in stress response. *Yeast*. 27, 245-258. https://doi.org/10.1002/yea.1752
- Saravanakumar, D., Ciavorella, A., Spadaro, D., Garibaldi, A., & Gullino M. L. (2008). *Metschnikowia pulcherrima* strain MACH1 outcompetes *Botrytis cinerea*, *Alternaria alternata* and *Penicillium expansum* in apples through iron depletion. *Postharvest Biology and Technology,* 49, 121-128. https://doi.org/10.1016/j.postharvbio.2007.11.006
- Sipiczki, M. (2006). Metschnikowia strains isolated from botrytized grapes antagonize fungal and bacterial growth by iron depletion. *Applied and Environmental Microbiology,* 72, 6716-6724. https://doi.org/10.1128/AEM.01275-06
- Spadaro, D., & Droby, S. (2016). Development of biocontrol products for postharvest diseases of fruit: the importance of elucidating the mechanisms of action of yeast antagonists. *Trends in Food Science & Technology*, 47, 39-49. https://doi.org/10.1016/j.tifs.2015.11.003
- Spadaro, D., Lorè, A., Garibaldi, A., & Gullino, M. L. (2013). A new strain of *Metschnikowia fructicola* for postharvest control of *Penicillium expansum* and patulin accumulation on four cultivars of apple. *Postharvest Biology and Technology* 75, 1-8. https://doi.org/10.1016/j.postharvbio.2012.08.001
- Spadaro, D., Micevska, T., & Simpson, S. (2008). Effect of nutrition on toxicity of contaminants to the epibenthic amphipod *Melita plumulosa*. *Archives of*

Environmental Contamination and Toxicology, 55: 593-602. https://doi.org/10.1007/s00244-008-9153-2

- Stanke, M., & Morgenstern, B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research,* 33, 465-467. https://doi.org/10.1093/nar/gki458
- Sugato, B., Banerjee, A., & Mooney, R. (2002). Semi-supervised clustering by seeding. *In Proceedings of 19th International Conference on Machine Learning,* 19-26.
- UniProt Consortium. (2017). UniProt: the universal protein knowledgebase. Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., ... & Breitling, R. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, *43*(1), 237-243. https://doi.org/10.1093/nar/gkv437
- Wisniewski, M., Droby, S., Norelli, J., Liu, J., & Schena, L. (2016). Alternative management technologies for postharvest disease control: the journey from simplicity to complexity. *Postharvest Biology and Technology,* 122, 3-10. https://doi.org/10.1016/j.postharvbio.2016.05.012
- Zhang, D., Spadaro, D., Garibaldi, & A., Gullino, M. L. (2010). Selection and evaluation of new antagonists for their efficacy against postharvest brown rot of peaches. *Postharvest Biology and Technology,* 55, 174-181. https://doi.org/10.1016/j.postharvbio.2009.09.007

SUPPLEMENTARIES

Supplementary file A.1. Fasta file of transcripts of *M. fructicola*.

Supplementary file A.2. Fasta file of CDSs of *M. fructicola*.

Supplementary file A.3. Fasta file of proteins of *M. fructicola*.

Supplementary file A.4. Gff file of *M. fructicola*.

Supplementary file A.5. Proteinortho analysis of *M. fructicola*, *M. bicuspidata*, *C. auris* and *C. lusitaniae*.

Supplementary file A.6. Annotation file of *M. fructicola*, produced by Blast2GO.

Supplementary file A.7. Vcf file, obtained by mapping the *M. fructicola* strain AP47 reads on the genome of strain 277.

Supplementary file A.8. List of GO terms related to the mutation or repair of the DNA sequence.

Supplementary table A.1. Annotation of *M. fructicola* transcripts.

Supplementary table A.2. CAZymes predicted in the *M. fructicola* 277 genome.

Supplementary table A.3. Fpkm expression data and statistical differences among conditions analysed with RNAseq.

All the supplementary material is available online at: https://drive.google.com/drive/folders/1MAF2Vag3B5sAgS9BPtPI4AOFZWItRAkB?usp=sharing

APPENDIX B: ELABORATED REGULATION OF GRISEOFULVIN BIOSYNTHESIS IN *PENICILLIUM GRISEOFULVUM* AND ITS ROLE ON CONIDIATION AND VIRULENCE

Silvia Valente [a†], Agnese Cometto [a], Edoardo Piombo [a], Giovanna Roberta Meloni [a], Ana-Rosa Ballester [b], Luis González-Candelas [b*] and Davide Spadaro [a*]

[a] Dept. Agricultural, Forest and Food Sciences (DISAFA), Università degli Studi di Torino, Largo Paolo Braccini 2, 10095 Grugliasco, Italy.
[b] IATA-CSIC – Instituto de Agroquímica y Tecnología de Alimentos. Calle Catedrático Agustín Escardino 7, Paterna, 46980-Valencia, Spain.

[*]Corresponding authors
E-mail addresses: davide.spadaro@unito.it, lgonzalez@iata.csic.es

ABSTRACT

Penicilium griseofulvum, the causal agent of apple blue mold, is able to produce *in vitro* and on apple a broad spectrum of secondary metabolites (SM), including patulin, roquefortine C and griseofulvin. Among them, griseofulvin is known for its antifungal and antiproliferative activity, and has received interest in many sectors, from medicine to agriculture. The biosynthesis of SM is finely regulated by filamentous fungi and can involve global regulators and pathway specific regulators, which are usually encoded by genes present in the same gene cluster as the backbone gene and tailoring enzymes. In the griseofulvin gene cluster, two putative transcription factors were previously identified, encoded by genes *gsfR1* and *gsfR2*, and their role has been investigated in the present work. Analysis of *P. griseofulvum* knockout mutants lacking either gene suggest that *gsfR2* forms part of a different

pathway and *gsfR1* exhibits many spectra of action, acting as regulator of griseofulvin and patulin biosynthesis and influencing conidia production and virulence on apple. The analysis of *gsfR1* promoter revealed that the regulation of griseofulvin biosynthesis is also controlled by global regulators in response to many environmental stimuli, such as carbon and nitrogen. The influence of carbon and nitrogen on griseofulvin production was further investigated and verified, revealing a complex network of response and confirming the central role of *gsfR1* in many processes in *P. griseofulvum*.

Keywords:
Transcription factor, knockout, gene cluster, regulation, secondary metabolites, apple blue mold

INTRODUCTION

Blue mold of pome fruit is one of the most important postharvest diseases of apple fruit. The major causal agent is *Penicillium expansum*, but many other *Penicillium* spp. can be found together causing blue mold, such as *P. griseofulvum*, *P. aurantiogriseum*, *P. crustosum, P. verrucosum, P. carneum, P. commune, P. brevicompactum, P. solitum* and *P. puberulum* (Moslem et al., 2010, 2013; Spadaro et al., 2011; Wu et al., 2019).

These species, as other *Penicillium* spp., can produce a wide range of secondary metabolites (SM), which are so called because they are not essential for primary metabolism and growth, but they play an essential role in the interactions, such as competition or communication with other microorganisms, and in many other cellular processes (Brakhage, 2013). Many of these SM have industrial value, such as antibiotics and antitumor compounds, while other metabolites can be toxic for humans and animals and are defined as mycotoxins.

In apples, the major concern is related to the presence of patulin, which is a mycotoxin with acute and chronic toxicity, characterized by stability in acid environment and during thermal processes, and therefore can be found in apples, apple juices and other derived products (Reddy et al., 2010).

Among causal agents of blue mold, *P. griseofulvum* is able to produce elevate amount of patulin, even higher compared to *P. expansum,* up to 4,500 µg/kg *in vivo*, while the European law limit is between 10 and 50 µg/kg depending on food commodities (Banani et al., 2016; Moslem et al., 2013; Spadaro et al., 2008).

P. griseofulvum can also produce other well-known mycotoxins such as roquefortine C and cyclopiazonic acid, and a really characteristic secondary metabolite called griseofulvin (Banani et al., 2016). Initially recognized as an antibiotic, this antifungal compound is classified as a potential carcinogenic for humans by the International Agency on Research on Cancer (World Health Organization, 2001). Nevertheless, griseofulvin is commonly found in many medical and veterinary products to treat cutaneous dermatophyte infections, and its role against cancer cells has been investigated by many researchers (Mauro et al., 2013; Rathinasamy et al., 2010). Thanks to its antiproliferative activity, griseofulvin and derived compounds have also been investigated for their potential use against fungal pathogens, such as *Botrytis cinerea*, *Sclerotinia sclerotiorum*, *Magnaporthe grisea*, *Corticium sasaki*, *Puccinia recondita*, *Blumeria graminis* f. sp. *hordei*, *Alternaria solani*, *Fusarium solani* and *Colletotrichum gloeosporioides*, showing efficacy both *in vitro* and *in vivo* (Bai et al., 2019; Ribeiro et al., 2018; Stierle and Stierle, 2015; Tang et al., 2015).

Due to the importance of griseofulvin in a wide range of applications, from medicine to agriculture, the griseofulvin biosynthetic genes were previously identified in *P. aethiopicum* (Chooi et al., 2010). The putative griseofulvin gene cluster consists of 13 genes, including a nonreducing polyketide synthase (PKS), tailoring enzymes and two putative transcription factors. The gene cluster was later determined in *P.*

griseofulvum and 10 genes of the cluster were found in the same order, while genes encoding for a reductase (*gsfK*), an hydrolase (*gsfH*) and a transcription factor (*gsfR2*) were not located in the same genomic region (Banani et al., 2016).

The putative genes encoding the backbone enzyme (*gsfA*), *O*-methyltransferases (*gsfB, gsfC, gsfD*) halogenase (*gsfI*), dehydrogenase /reductase (*gsfE*) and cytochrome P450 (*gsfF*) were confirmed by producing deletion mutants (Cacho et al., 2013; Chooi et al., 2010), but the role of the putative transporter as well as the regulatory genes remains to be investigated.

Regulatory genes are an interesting target for mutagenesis because silencing or promoting their expression can provide a huge enhancement of compounds production, and therefore industrial value.

In this work, the regulation of griseofulvin biosynthesis was investigated by characterizing knockout mutants for the putative transcription factors encoded by *gsfR1* and *gsfR2* and comparing them with wild type and deletion mutants for the backbone gene of the cluster, *gsfA*.

However, the regulation of SM biosynthesis is incredibly complex and involves not only pathway-specific regulators, which are commonly encoded by genes located inside the gene cluster, but also global regulators which are expressed in response to various environmental stimuli (Brakhage, 2013). A second approach was therefore investigated, aimed to study the involvement of global regulators of the cluster, particularly by evaluating the effect of carbon and nitrogen sources on *P. griseofulvum* growth and on griseofulvin biosynthesis. The findings lead to a better comprehension of griseofulvin biosynthesis and of the role of this compound in the growth and virulence of *P. griseofulvum*.

MATERIALS AND METHODS

Fungal strain

Penicillium griseofulvum Dierckx strain PG3 was previously isolated from rotten apples in Piedmont (Spadaro et al., 2011) and grown on Potato Dextrose Agar (PDA, Merck KGaA, Darmstadt, Germany) with 50 µg/mL streptomycin (Merck) in the dark at 25 °C for 7-10 days to prepare conidial suspension, while transformed strains were grown on PDA containing 500 µg/mL of hygromycin B (ThermoFischer Scientific, Waltham, MA USA) in the same conditions. Conidial suspensions were obtained by adding 5 mL of sterile water with 0.01 % (w/v) Tween-20 and gently scraping the surface of fungal cultures grown on petri plates. The final conidia concentration was measured using hemocytometer and adjusted by diluting to different concentrations depending on each assay. Conidial suspension of wild type and transformed strains were maintained in glycerol at -80 °C.

Bioinformatic analysis

The promoter of the genes *gsfR1* and *gsfA* were extracted from the genome of *P. griseofulvum* PG3 (Genbank: GCA_001561935.1), and they were scanned for transcription factor (TF) binding using the CIS-BP database (http://cisbp.ccbr.utoronto.ca/TFTools.php) with default parameters. *P. chrysogenum* was used as model species when performing the analysis. Identified *P. chrysogenum* TFs were blasted against PG3 proteome to check for their presence in *P. griseofulvum*, and they were then identified by blasting against NCBI non-redundant protein database.

Obtainment of knockout mutants

Knockout mutants were obtained as previously described (Ballester et al., 2015), by amplifying 5' (promoter) and 3' (terminator) flanking regions of *gsfA*, g*sfR1* and *gsfR2* genes from the genomic DNA of *P. griseofulvum* PG3 with specific primers

pairs (O1/O2 and A3/A4, **Table SB.1**) containing a tail for USER cloning (Frandsen et al., 2008). Amplified fragments were mixed with pRFHU2 vector pre-digested with *PacI* and *NtBbVCI* (New England Biolabs, Ipswich, MA, USA) and treated with USER enzyme (New England Biolabs). Five µL of the USER mixture was directly used to transform chemically competent cells of *Escherichia coli* strain DH5α, with heat shock protocol, then bacterial cells were recovered and plated on Luria Bertani Agar (LB, Miller, Merck) supplemented with 25 µg/mL of kanamycin (ThermoFischer Scientific). Resistant transformants were screened by colony PCR and plasmidic DNA from positive colonies was obtained for both promoter and terminator using NucleoSpin® Plasmid EasyPure kit (Macherey-Nagel, Düren, Germany) following the manufacturer's instructions. The correct integration of promoter and terminator of genes was confirmed by sequencing the amplified fragments with primers pairs RF1/RF6 and RF2/RF5 (**Table SB.1**). Then, 10 ng of plasmid were introduced by electroporation into electro-competent *A. tumefaciens* AGL1 cells and *A. tumefaciens* mediated transformation of *P. griseofulvum* was conducted as described by Buron-Moles et al. (2012). Transformants were maintained on PDA with 500 µg/mL of hygromycin B, according to the results of a sensitivity assay performed on PDA for *P. griseofulvum*.

DNA and RNA extraction

DNA of *P. griseofulvum* PG3 and knockout mutants was obtained using E.Z.N.A.® Fungal DNA Mini Kit (Omega Bio-tek, Norcross, GA, USA) using the slightly modified protocol for isolation of high molecular weight DNA from fungal mycelium (Möller et al., 1992). Briefly, a conidial suspension (10^5 conidia/mL) was inoculated in 50 mL of Glucose Yeast Peptone medium (GPY; 4% w/v glucose, 0.5% w/v yeast extract and 0.5% w/v peptone) and incubated on a rotary shaker (200 rpm) at 24 °C for 2 days. Fungal mass was collected, dried and frozen in liquid nitrogen. DNA extraction from 0.5 g of frozen mycelium was performed in 50 mL

tubes containing 5 mL TES (100 mM Tris HCl, pH 8.0, 10 mM EDTA, 2% SDS), 25 µL proteinase K, and 50 µL ß-mercaptoethanol.

DNA concentration and purity were checked by spectrophotometer (Nanodrop 2000, Thermo Scientific, Wilmington, USA).

RNA extraction was performed according to Ballester and collaborators (2015) with some adjustments. First, 5 µL of a conidial suspension (10^6 conidia/mL) were inoculated on PDA with a cellophane membrane and they were incubated at 24 °C in the dark. One hundred mg of mycelium were collected at 5, 7 and 10 dpi and were frozen in liquid nitrogen. Two tungsten beads (diameter: 2.7 mm) were added to the mycelium that was crushed using TissueLyser (Qiagen, Hilden, Germany) for 1 min at 20.00 Hz speed. Then, 750 µL of extraction buffer (100 mM Tris-HCl pH 8.0; 100 mM lithium chloride, 10 mM EDTA pH 8.0; 1% SDS; 1% PVP-40; 1% β-mercaptoethanol) and 375 µL of phenol:chloroform:isoamyl alcohol (25:24:1, v:v:v) were added, then the mixture was vortexed and incubated at 65 °C for 10 min. After cooling, 375 µL of phenol:chloroform:isoamyl alcohol were added and the extract was vortexed. The homogenate was centrifuged at 13,800 x *g* for 10 min and the aqueous phase was re-extracted with 750 µL of phenol:chloroform:isoamyl alcohol. Nucleic acids were precipitated by adding 1/10 volume of 3 M sodium acetate pH 5.2 and 2 volumes of cold 100% ethanol and incubating at -20 °C for 60 min. The precipitated RNA was pelleted by centrifugation for 15 min at 13,800 x *g*. After washing with 500 µL of 70% ethanol, pellet was dissolved in 600 µL of TES buffer (10 mM Tris–HCl pH 8; 5 mM EDTA pH 8; 0.1% SDS) and precipitated overnight at −20 °C with 200 µL of 12 M lithium chloride. After centrifugation at 13,800 x *g* for 60 min, the pellet was washed with 500 µL of 70% ethanol. The pellet was re-extracted with 250 µL of 3 M sodium acetate, pH 6.0, to remove residual polysaccharides and washed with 70% ethanol. Finally, RNA extract was dissolved in 50 µL of DEPC-water. RNA concentration was measured spectrophotometrically and verified by gel electrophoresis.

DNase treatment was performed using TURBO DNA-free™ Kit (Thermo Fischer Scientific) with some changes. The samples were diluted 1:2 in a final volume of 50 µL. Five µL of 10X TURBO DNase™ Buffer and 0.5 µL TURBO DNase™ Enzyme were added to the RNA and gently mixed. The samples were incubated at 37 °C for 30 min. Further 0.5 µL of TURBO DNase™ Enzyme was added to the RNA and it was incubated again at 37 °C for 30 min. Finally, 5 µL of DNase Inactivation Reagent was added to the RNA and gently mixed. The samples, after incubation at room temperature for 5 min, were centrifuged at 13,000 x g for 2 min. The supernatant containing RNA was transferred to a new fresh tube.

First-strand cDNA was conducted with High Capacity cDNA Reverse Transcription Kit (Thermo Fischer Scientific) using 1 µg of total RNA. The 2X RT master mix consisted of 2 µL of 10X RT buffer, 0.8 µL of 25X dNTP Mix (100 mM), 2 µL of 10 X RT Random Primers, 1 µL MultiScribe® Reverse Transcriptase and nuclease-free water to a final volume of 20 µL. The thermal cycling conditions were 10 min at 25 °C, 120 min at 37 °C and 5 min at 85 °C.

PCR and qPCR

Upstream and downstream fragments of *gsfR1*, *gsfR2* and *gsfA* genes were amplified by PCR with Top-Taq DNA polymerase (Bioron Diagnostics GmbH, Römerberg, Germany) and a T100™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). The cycling conditions of the PCR were 5 min at 94 °C, followed by 35 cycles of 15 s at 94 °C, 20 s at 58 °C, 90 s at 72 °C and a final extension of 5 min at 72 °C. The PCR mixture contained 1X PCR Buffer, 0.2 mM dNTPs, 0.4 µM of each primer, 0.5 U of polymerase and 10 ng of genomic DNA of *P. griseofulvum*. Taq DNA Polymerase (Qiagen) was used to perform the other PCRs. To confirm the correct integration of T-DNA in the genome, several PCRs were conducted to amplify the hygromycin cassette and *gsfA*, *gsfR1* and *gsfR2* genes. The cycling conditions of the

PCRs were 3 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 45 s at 58 °C, 90 s at 72 °C and 5 min at 72 °C.

To verify the efficiency of the DNase treatment, a PCR was performed using ITS1 and ITS4 universal primers. The cycling conditions of the PCR were 3 min at 94 °C, followed by 40 cycles of 30 s at 94 °C, 30 s at 55 °C, 1 min at 72 °C and 5 min at 72 °C.

RT-qPCRs were performed with StepOne™ and StepOnePlus™ Real-Time PCR System with Power SYBR™ Green PCR Master Mix (ThermoFischer Scientific); cycling conditions were 5 min at 95 °C, followed by 45 cycles of 10 s at 95 °C, 30 s at 58 °C and 30 s at 72 °C. In order to determine the number of insertions in the genome of the transformants, the $2^{\Delta\Delta cq}$ method (Pfaffl, 2001) was used using genomic DNA of samples and comparing the amplification of β-tubulin gene or histone H3, with the amplification of the promoter of the corresponding gene target of deletion.

In order to verify the expression of *gsfA* and *gsfR1* genes in Δ*gsfR1* mutants, cDNA was added in the reaction instead of gDNA.

All primer sequences used in the PCR and qPCR reactions are listed in **Table SB.1**.

Characterization of mutants *in vitro*

Deletion and ectopic mutants were compared with wild type *P. griseofulvum* PG3 by inoculating 5 µl of spore suspension ($1*10^6$ conidia/mL) on PDA and incubating the plates at 25 °C in the dark. Some mutants were also inoculated on Malt Extract Agar (MEA, 2% malt extract, 2% glucose, 0.1 % peptone, 2% agar) and Meat Peptone Agar (MPA, 2.5% meat extract, 0.25% peptone, 2% agar) under the same conditions. Colony diameter (cm), number of asexual spores (conidia/plate) and griseofulvin production (µg/plate) were measured up to 10 days post inoculation (dpi). At least 5 plates were inoculated for each strain and every assay was performed three times.

Characterization of mutants *in vivo*

Apples 'Golden Delicious' were purchased from a local supermarket and stored at 4 °C until use. Fruits were surface disinfected with sodium hypochlorite (1% as chlorine), washed with deionized water and allowed to air dry before inoculation, then 21 fruits for each treatment were inoculated by wounding in three points at the equatorial region of the fruits with a sterile needle (3 mm depth, 3-4 mm wide) (Spadaro et al., 2013). Ten µl of conidial suspension ($1*10^8$ conidia/mL) of each strain were pipetted into each wound, while controls were inoculated with deionized water. Wounded apples were placed in plastic trays, covered with a transparent polyethylene film and stored at room temperature for 3-14 days.

Griseofulvin and patulin extraction

Griseofulvin and patulin produced *in vitro* were extracted simultaneously from the mycelium and the medium by washing the plates with 3 mL of methanol. The plates were gently scraped and then placed in rotary shaken at 450 rpm for 15 min. The liquid was then collected and centrifuged (13000 x *g*) at room temperature for 5 min. After that, the supernatant was filtered through a 0.45 µm syringe filter and analysed by HPLC.

Griseofulvin was extracted from rotten apples as previously described by Banani et al. (2016) with some changes. Briefly, the portion around the inoculation sites from 7 apples was collected and combined. Ten g of each sample were weighed and then placed in a centrifuge tube with 10 drops of pectinase enzyme solution and 10 mL of water were added. The mixture was left at 38 °C for 2 hours and then centrifuged at 3,600 *x g* for 5 min. Five mL of clear liquid were placed into a clean tube and griseofulvin was extracted with 5 mL of ethyl acetate (three times). The organic phase was evaporated to dryness using rotary evaporator and the residual dissolved

in 500 µL of MeOH:H_2O (1:1, v/v) transferred into a HPLC vial for HPLC-MS/MS analysis.

Chemical analyses

In order to characterize the production of griseofulvin and patulin on PDA, the analyses were carried out using the method previously described (Banani et al., 2016) with few modifications. A Waters HPLC system equipped with a 600 Consolider pump, a 717 plus Autosampler and a 2996 photodiode array detector was used. The analytical column used was a Kinetex 5 µm Biphenyl 100Å column (150 x 4.6 mm; Phenomenex, Torrance, CA, USA) coupled to a SecurityGuard ULTRA Cartridges UHPLC Biphenyl guard column (4.6 mm; Phenomenex). The chromatography conditions were: flow rate of 1 mL/min on a linear gradient of 5 to 95 % solvent B in 30 min followed by isocratic 95 % solvent B for another 10 min (solvent A: acid water pH 4.0 with acetic acid; solvent B: acetonitrile); the injection volume was 20 µL. The spectra were acquired between 200 nm and 600 nm wavelength, and patulin and griseofulvin quantifications was performed at 294 nm and 304 nm, respectively.

Patulin and griseofulvin were identified in the samples by comparing the retention time and UV-vis spectra with those of chemical standards.

Comparison of griseofulvin production on different growth media and *in vivo* was instead analysed by liquid chromatography coupled with mass spectrometry. The HPLC-MS/MS system consisted of a binary pump and a vacuum degasser (1260 Agilent Technologies, Santa Clara, CA, USA) connected to a Varian auto-sampler Model 410 Prostar (Hansen Way, CA, USA), equipped with a 20 µL loop and coupled with a Varian 310-MS TQ Mass Spectrometer. The sample was injected (10 µL) onto Luna C18(2) (150mm x 2mm i.d. 3µm, Phenomenex) and eluted under a flow of 200 µL/min. The mobile phase was an isocratic mixture of ACN:HCOOH 0.05% (60:40, v/v) for 5 min.

The mass spectrometer was equipped with an electrospray ionization (ESI) source operating in positive ion mode. Griseofulvin detection and quantitation were carried out using multiple reaction monitoring (MRM) mode by selecting three transitions: m/z 353 > 285 CE 18 eV (monitoring), m/z 353 > 215 CE 16 eV (quantification) and m/z 353 > 165 CE 16 eV (monitoring). The collision gas (Ar) pressure was set at 2 mbar for all experiments.

Quantification was performed by external calibration preparing a calibration curve for each mycotoxin. The calibration curves were prepared by dilution of standard solutions of griseofulvin and patulin.

Statistical analysis

All statistical analyses were performed with one-way ANOVA followed by Duncan test using IBM SPSS statistics software version 24 (SPSS Inc., Chicago, IL, USA), $p < 0.05$ was considered significant. The REST 2009 Software (Qiagen) was used for the statistical analysis of RT-qPCR results.

RESULTS

Obtainment of knockout mutants for gsfR1 and gsfR2 and role on griseofulvin biosynthesis

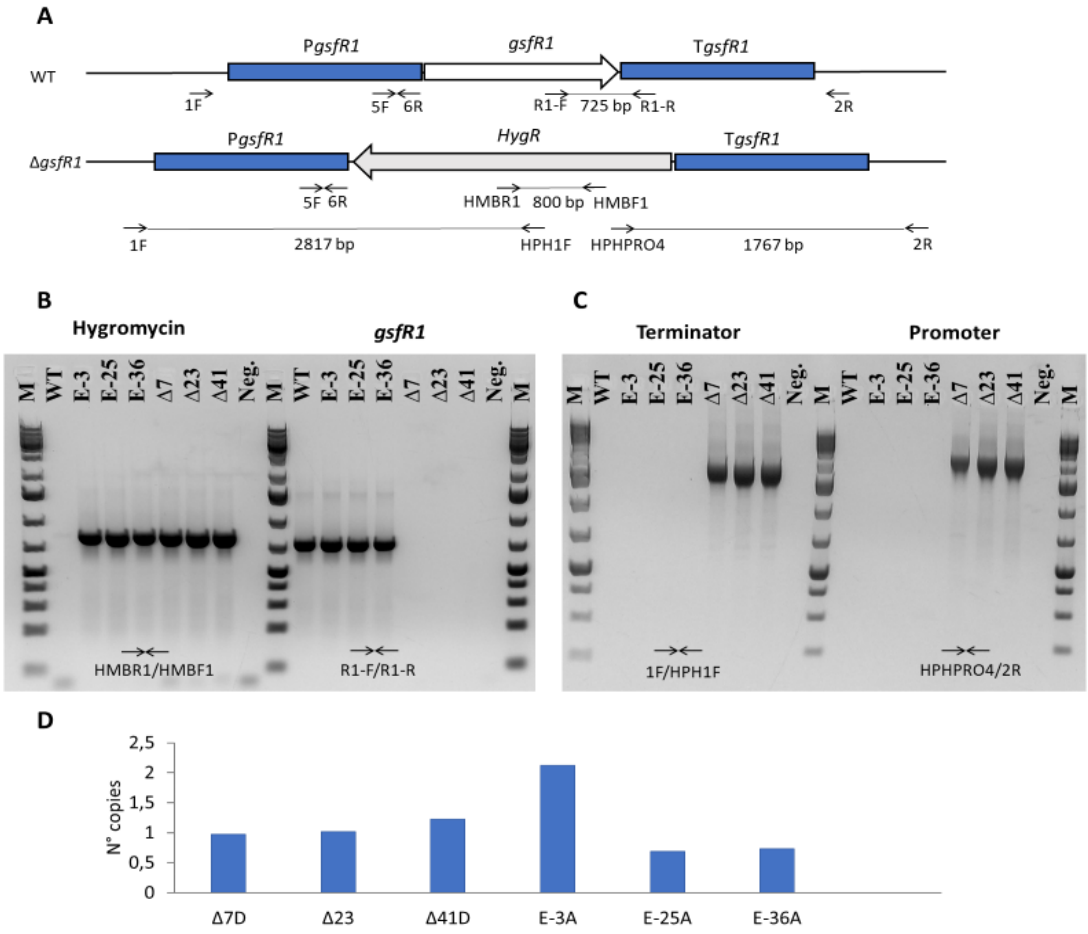

Figure B.1 – PCR analysis of *gsfR1* mutants. Schematic presentation of the *gsfR1* locus in the wild type and deletion mutants, primers used are presented as arrows (A). Amplification of hygromycin resistance cassette (*HygR*) and *gsfR1* gene (B); confirmation of orientation of inserted T-DNA (C); copies number of T-DNA inserted in *P. griseofulvum* (D) obtained through qPCR on gDNA using primers 5F/6R and histone H3 gene as reference gene. M= GeneRuler 1Kb Plus DNA ladder, WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfR1* and E = ectopic strains, Neg. = negative control (PCR mix without DNA).

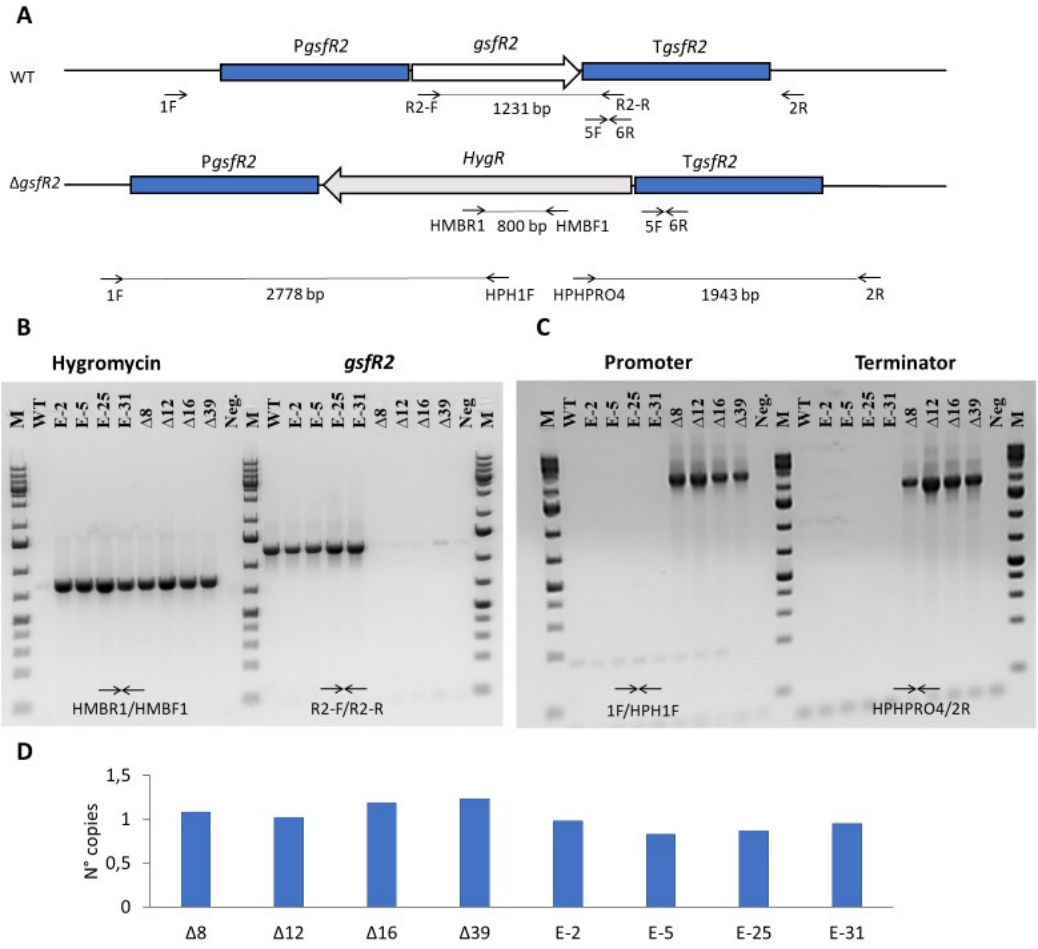

Figure B.2 – PCR analysis of *gsfR2* mutants. Schematic presentation of the *gsfR2* locus in the wild type and deletion mutants, primers used are presented as arrows (**A**). Amplification of hygromycin resistance cassette (*HygR*) and *gsfR2* gene (**B**); confirmation of orientation of inserted T-DNA (**C**); copies number of T-DNA inserted in *P. griseofulvum* (**D**), obtained through qPCR on gDNA using primers 5F/6R and histone H3 gene as reference gene. M= GeneRuler 1Kb Plus DNA ladder, WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfR2* and E = ectopic strains, Neg. = negative control (PCR mix without DNA).

The role of putative transcription factors encoded by *gsfR1* and *gsfR2* was assessed by obtaining deletion mutants and they were compared with the wild type strain and a deletion mutant for *gsfA*, the gene encoding the PKS.

The deletion of genes was mediated by *A. tumefaciens*, which delivered a T-DNA containing a hygromycin resistance cassette flanked by the promoter and terminator of target genes. Thanks to the presence of these flanking regions, homologous recombination led to replacement of genes with the antibiotic resistance cassette (**Figs. B.1-B.2**, **SB.1**).

The obtainment of knockout mutants was then assessed by PCR, by amplifying the hygromycin resistance cassette and target gene (**Figure B.1 B**, **B.2 B**, **SB.1 B**). As expected, the deleted genes were not amplified in knockout mutants, while hygromycin resistance cassette was only absent in the wild type PG3. Ectopic strains had a functional copy of the target gene and the antibiotic resistance cassette, due to non-target integration into the genome of the T-DNA.

To confirm the right insertion of the cassette, PCRs with primer pairs designed on the hygromycin resistance cassette and on 3'UTR or 5'UTR of deleted gene were performed (**Figure B.1 C**, **B.2 C**, **SB.1 C**). These primer pairs can only amplify if the T-DNA has been introduced by homologous recombination at the target sites in the 3'UTR and 5'UTR flanking regions. As expected, the PCR fragments were amplified in the knockout mutants, and failed amplification in the wild type PG3 and in the ectopic mutants.

The number of integrations in the genome was determined by qPCR by measuring the difference between the quantification cycle (Cq) of target and reference genes in the mutant and in the control wild type strain (Pfaffl, 2001) (**Figure B.1 D**, **B.2 D**, **SB.1 D**).

Three mutants with one single event of integration and two ectopic strains were selected for each knockout event and they were characterized *in vitro* both phenotypically and chemically.

Concerning the effect of gene deletion on phenotype, while the deletion of the polyketide synthase did not affect the phenotype (**Figure SB.2**), knockout mutants for *gsfR1* showed the same growth rate as the parental strain (**Figure B.3A**), but

exhibited colonies with a markedly less green color, caused by a significantly reduced conidiation *in vitro* (**Figure B.3 B, C**). On the other hand, knockout *gsfR2* mutants were similar to the wild type in both growth and sporulation (**Figure B.5**). The ability of knockout mutants to produce griseofulvin was assessed on PDA.

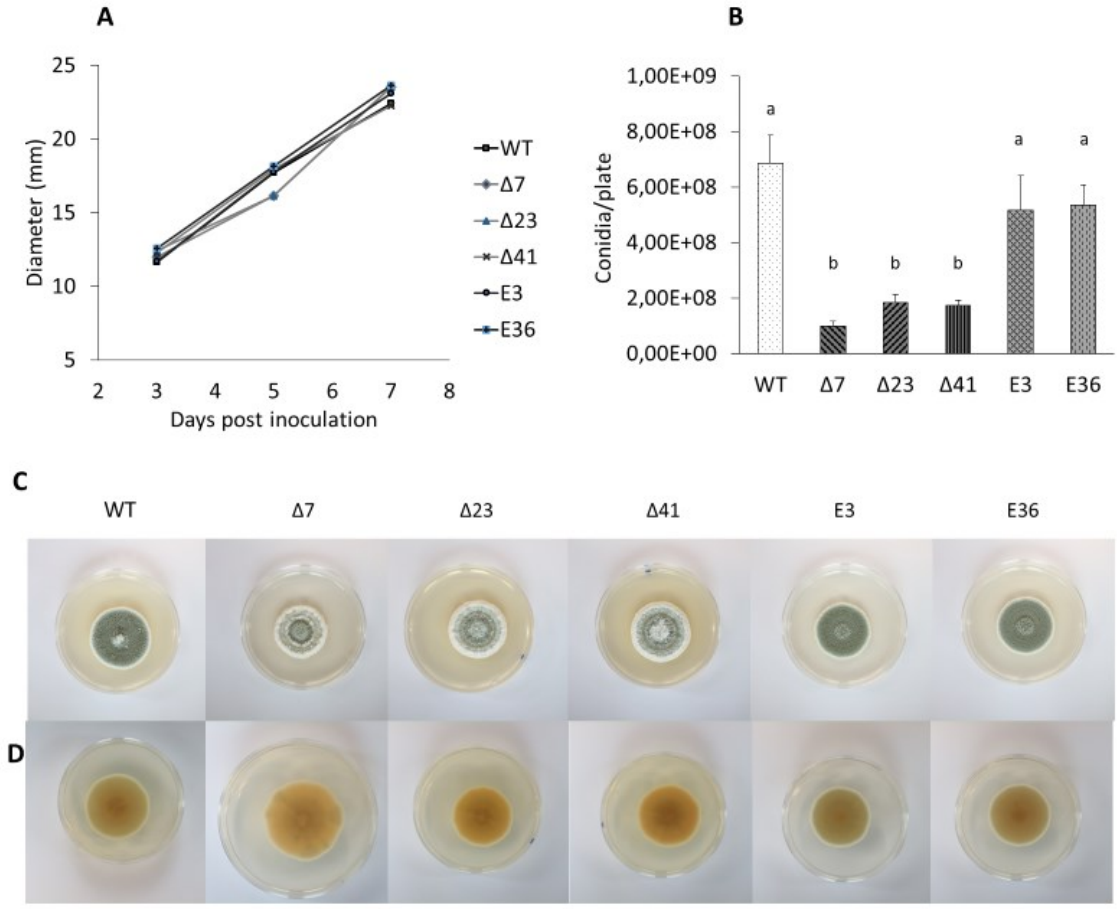

Figure B.3 – Effect of *gsfR1* deletion on *P. griseofulvum* growth *in vitro.* Colony diameter (A) and conidia production (B) from 3 to 7 dpi. Plate view (front C and reverse D) at 7 dpi on PDA. WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfR1* and E = ectopic strains. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

Mutants lacking the *gsfA* gene were no longer able to produce griseofulvin *in vitro* (**Figure SB.2 B**), while knockout mutants for the regulator gene *gsfR1* were able to produce significantly more griseofulvin compared to wild type strain, with an

average of about 30%, 50% and 110% higher production compared to wild type strain at 5, 7 and 10 dpi (**Figure B.4 A**, **Figure B.9**).

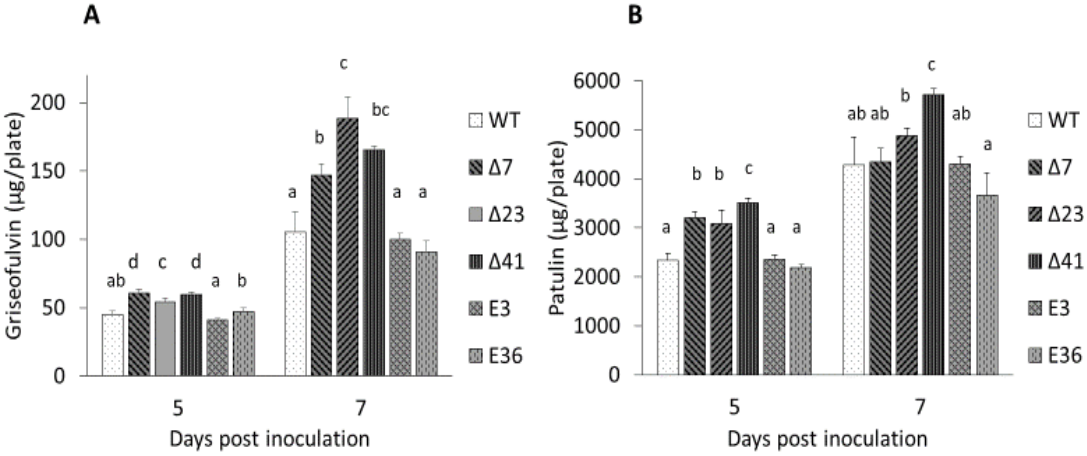

Figure B.4 - Effect of *gsfR1* deletion on griseofulvin and patulin production. Griseofulvin (A) and patulin (B) production were measured at 5 and 7 dpi. WT = wild type *P. griseofulvum*, Δ = deletion mutants and E = ectopic strains. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

The production of patulin was also investigated and deletion of *gsfR1* showed to influence also the regulation of this mycotoxin, especially at 5 dpi, reaching over 3000 µg/plate in all mutants compared to 2300 µg/plate produced by *P. griseofulvum* PG3, which was significantly different according to Duncan's statistical analysis ($p < 0.05$) (**Figure B.4 B**).

The knockout of *gsfR2* did not affect griseofulvin or patulin production, nor conidiation or growth rate, which were comparable between wild type and knockout and ectopic mutants (**Figure B.5, B.6**).

GsfR1 regulates negatively griseofulvin biosynthesis on PDA

The observation that the deletion of *gsfR1* led to a higher griseofulvin production *in vitro* permitted to hypothesize that this regulatory gene could encode a negative

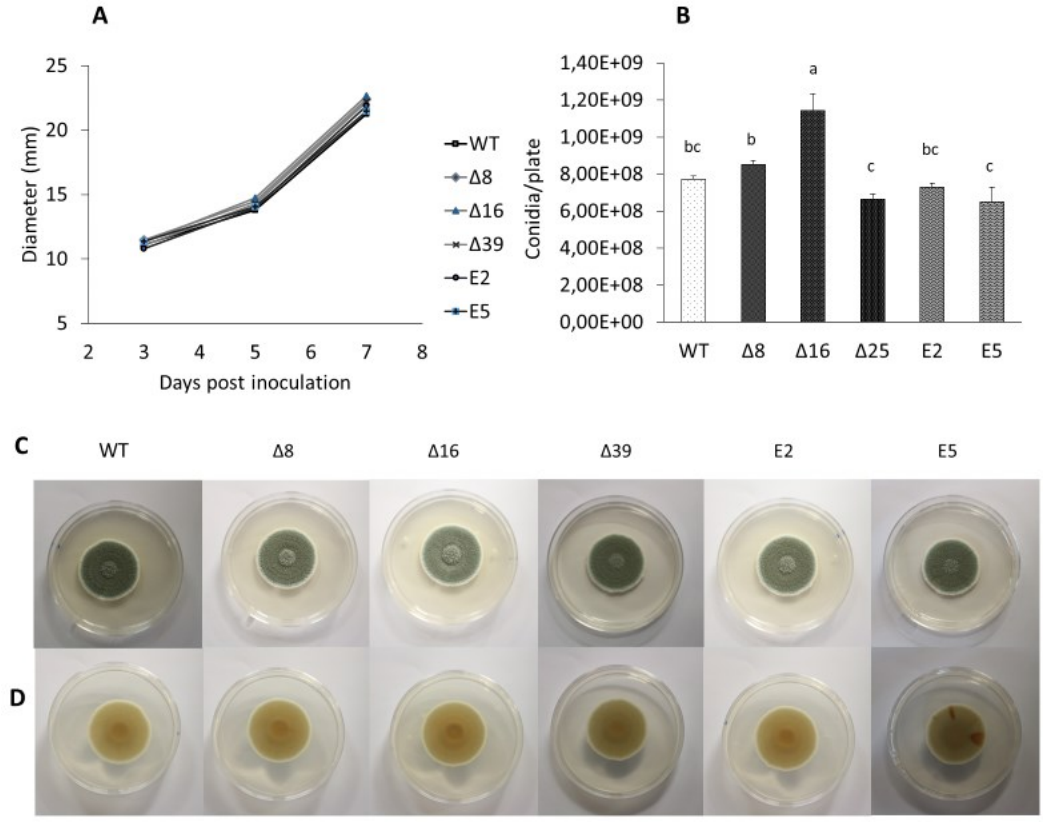

Figure B.5 – Effect of *gsfR2* deletion on *P. griseofulvum* growth *in vitro.* Colony diameter (A) and conidia production (B) from 3 to 7 dpi. Plate view (front C and reverse D) at 7 dpi on PDA. WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfR2* and E = ectopic strains. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

regulator of griseofulvin biosynthesis. To demonstrate this hypothesis, the relative expression of the polyketide synthase gene of three mutants and two ectopic strains was investigated through RT-qPCR. The Δ*gsfR1*deletion mutants showed an increased expression of *gsfA* compared to the wild type, both at 5 and 7 dpi, while ectopic mutants exhibited a similar gene expression (**Figure B.7**).

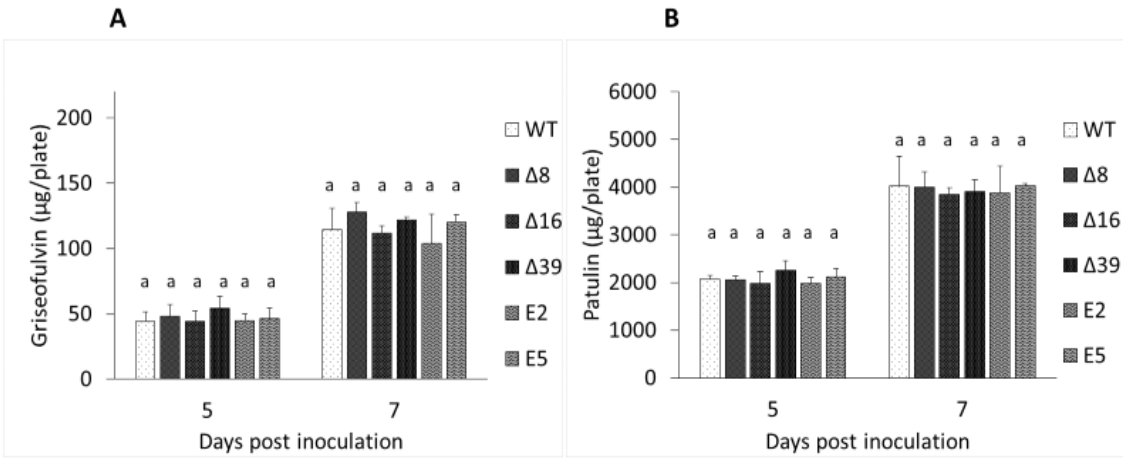

Figure B.6 - Effect of *gsfR2* deletion on griseofulvin and patulin production. Griseofulvin (**A**) and patulin (**B**) production were measured at 5 and 7 dpi. WT = wild type *P. griseofulvum*, Δ = deletion mutants and E = ectopic strains. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

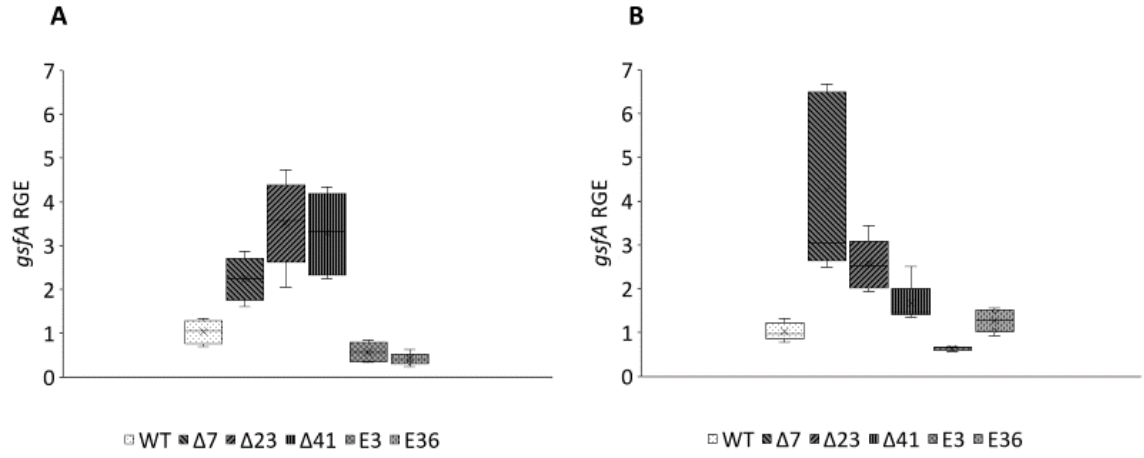

Figure B.7 – Relative gene expression (RGE) of *gsfA* gene, at 5 (**A**) and 7 (**B**) dpi on PDA. WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfR1* and E = ectopic strains. The expression is relative to the expression of the β-tubulin gene.

Virulence and griseofulvin production on apples

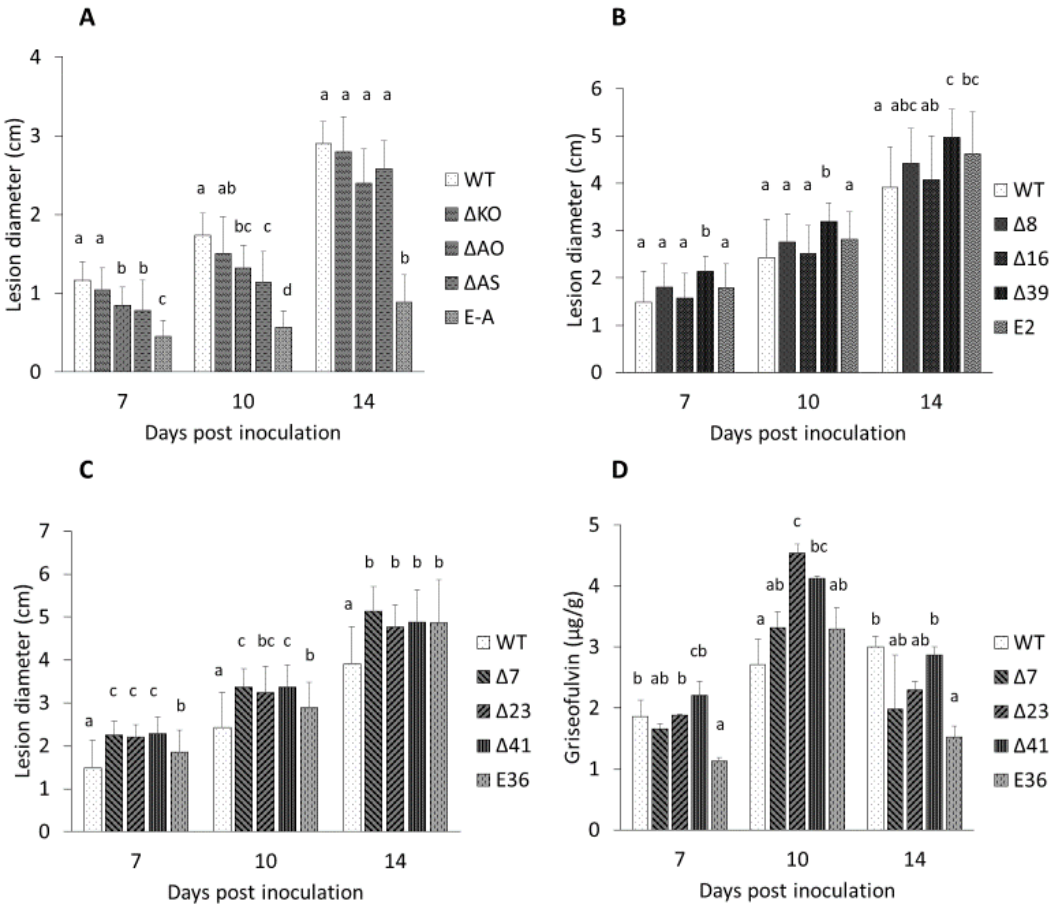

Figure B.8 - Effect of gene deletions on virulence and griseofulvin production *in vivo*. Lesion diameter caused by the wild type strain was compared with rot diameter of knockout ad ectopic strains for *gsfA* (A) *gsfR2* (B) and *gsfR1* (C). Griseofulvin production (µg/g of apple) was compared between wild type and Δ*gsfR1* strains (D). Measurements were taken at 7, 10 and 14 dpi. WT = wild type *P. griseofulvum*, Δ = deletion mutants and E = ectopic strains. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$).

In order to evaluate the effect of griseofulvin on the pathogenicity of *P. griseofulvum*, the virulence of knockout mutants was compared with that of the wild type strain and one ectopic mutant on apples (**Figure B.8**). *ΔgsfA* mutants, which were not able to produce griseofulvin, showed a slightly reduced virulence (**Figure B.8 A**). In this assay, apples infected with E-A ectopic mutant, which contains 4 insertions of the T-DNA, led to a strong reduction of virulence on apples. However, *gsfR2* mutants showed a similar virulence compared to the wild type strain (**Figure B.8 B**).

On the contrary, Δ*gsfR1* strains were slightly more virulent (**Figure B.8 C**) compared to the *P. griseofulvum* parental strain. Nevertheless, the ectopic mutant showed an enhanced virulence as well, although it was less virulent than the mutant in the first stages of the infections. For this reason, we cannot exclude that the enhanced virulence of the mutants was at least partly due to the transformation itself. Griseofulvin was then extracted from apples inoculated with *gsfR1* mutants, and a significant increased production of griseofulvin from knockout mutants was observed at 10 dpi (**Figure B.8 D**), while at 14 dpi *gsfR1* mutants produced less griseofulvin compared to wild type strain.

Global regulators in griseofulvin biosynthesis

By using information on the CIS-BP database, the promoters of *gsfR1* and *gsfA* were scanned in order to find putative binding sites for global regulators. Eighty-one and seventy-five transcription factors were predicted to bind the promoters of *gsfR1* and *gsfA*, respectively. Thirteen transcription factors seem to bind only the promoter of *gsfR1*, while 7 are typical of *gsfA* and 68 were predicted to regulate both genes, ensuing a complex scenario (**Table SB.2**).

Carbon and nitrogen have a strong effect on in vitro conidiation and griseofulvin production

Among the different stimuli, the availability of nutritional sources, especially nitrogen and carbon, is a key factor for the biosynthesis of many SMs and for the growth and differentiation of filamentous fungi. Because of this, it is not surprising that we found many putative TFs involved in carbon or nitrogen consumption, which could bind the promoter of *gsfR1* and *gsfA* (**Table SB.2**).

We focused our attention on the effect of carbon and nitrogen on griseofulvin production. To do that, two *gsfR1* deletion mutants and one ectopic strain were grown for ten days on three different media (PDA, MEA and MPA); one knockout mutant for the polyketide synthase gene was also included as control (denoted as KO). These three media are characterized by different C/N ratios: PDA is composed of potato extract and dextrose and has the highest carbon content, while on MEA and MPA, peptone was added, increasing the nitrogen content, which is ten times higher on MPA (thanks to the addition of meat extract) compared to MEA (Núñez et al., 2000). As expected, carbon and nitrogen showed a huge impact both on griseofulvin and conidia production (**Figure B.9**).

Conidiation of the wild type strain was reduced in both media containing peptone, with a stronger reduction on MEA (**Figure B.9A**), and a significant reduction of griseofulvin production was observed in MPA (23%) compared to PDA or MEA (**Figure B.9B**).

Conidiation of the *gsfA* knockout mutant, lacking the gene encoding the polyketide synthase, was similar to that observed for the wild type strain in the three media tested. However, the knockout mutant was not able to produce griseofulvin in any condition.

When focusing on *gsfR1* mutants, a more complex scenario was observed. Deletion of *gsfR1* gene led to a significant reduction of conidia on MEA and PDA compared to wild type, ectopic strain and *gsfA* mutants, suggesting that this gene could play a

role in the direct or indirect regulation of conidia production. Considering griseofulvin production, a significant increase of griseofulvin production was observed on PDA (**Figure B.4 and B.9B**). Instead, on the other two media, richer in nitrogen, there was a significantly reduced production compared to wild type.

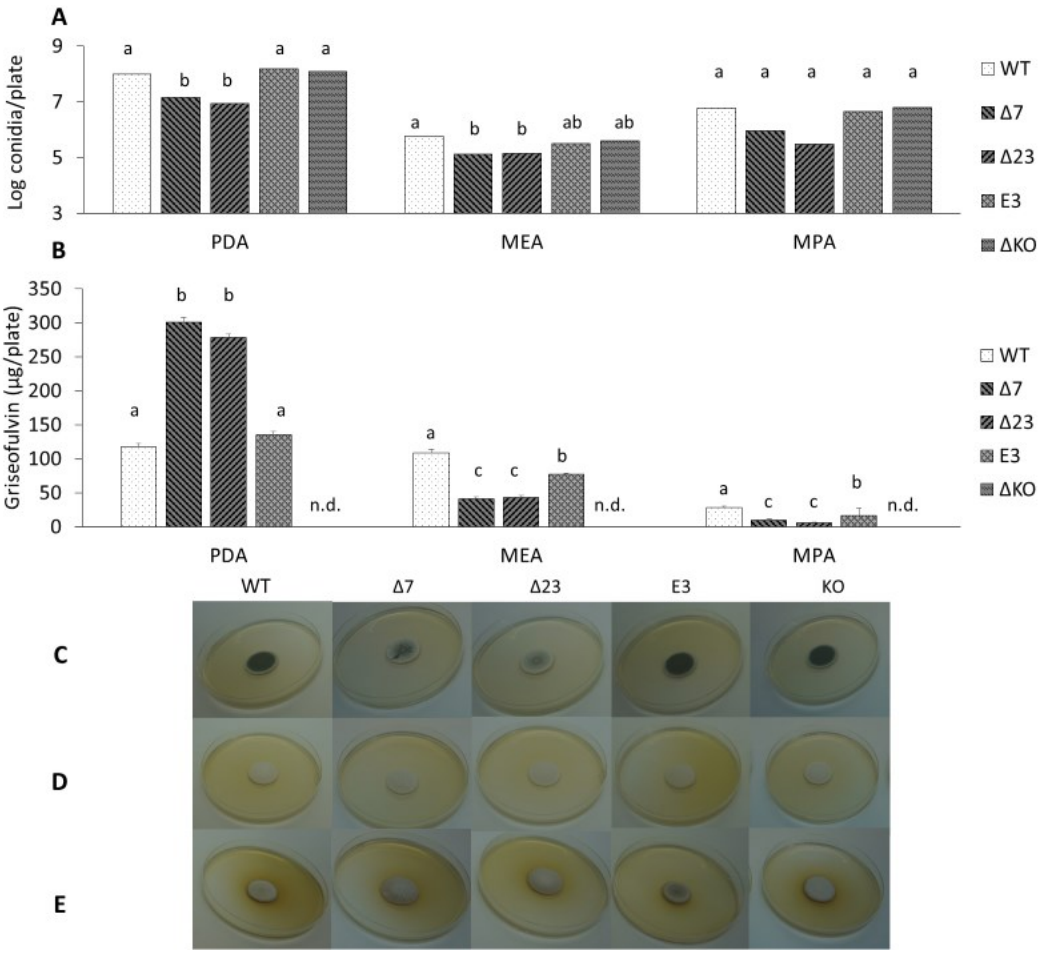

Figure 9 – Effect of carbon and nitrogen on conidiation and on griseofulvin production *in vitro*. Logarithm of conidia per plate (A) and griseofulvin produced (B) were measured at 10 dpi in three solid media (PDA, MEA and MPA). Colony view on PDA (C), MEA (D) and MPA (E). WT = wild type *P. griseofulvum*, Δ7, Δ23 = deletion mutants for *gsfR1*, E3 = ectopic strain and ΔKO = deletion mutant for *gsfA*. Values followed by the same letter are not statistically different by Duncan's multiple range test ($p < 0.05$); n.d. = non detectable.

DISCUSSION

Role of transcription factors *GsfR1* and *GsfR2* on griseofulvin biosynthesis

The griseofulvin gene cluster previously identified in *P. aethiopicum* includes two genes, *gsfR1* and *gsfR2*, which encode for putative transcription factors. The *gsfR2* gene in *P. griseofulvum* was located in a different genomic region of the biosynthetic gene cluster. It was previously reported how some TFs can regulate the expression of a specific gene cluster even when they are located outside. This is the case of both *fum21* and *zfr1*, which are located inside and outside the gene cluster respectively, and regulate fumonisin biosynthesis in *Fusarium verticilloides* (Brown et al., 2007; Flaherty and Woloshuk, 2004). The aim of our work was to elucidate the role of both putative transcription factor in griseofulvin biosynthesis by obtaining deletion mutants. Mutants for the *gsfA* core biosynthetic gene encoding a polyketide synthase were obtained as control because they are not able to produce griseofulvin (Chooi et al., 2010).

Our results show that the *gsfR1* gene encodes a putative transcription factor that not only acts on griseofulvin biosynthesis but also plays an essential role as an important regulator of *P. griseofulvum* development and secondary metabolism. *gsfR1* seems to regulate griseofulvin biosynthesis, acting as a negative regulator of the cluster on PDA medium and on apples. Indeed, the deletion of this regulatory gene led to an increase of griseofulvin production in these conditions and a higher expression of the *gsfA* gene *in vitro.* Conversely, on media supplemented with peptone, *gsfR1* gene positively regulates the griseofulvin biosynthesis, as all mutant strains were able to produce significantly less compound compared to wild type. Similar results were obtained studying the *mtfA* gene encoding the TF of aflatoxin in *Aspergillus flavus* and *Aspergillus parasiticus*, revealing that the regulatory activity of this class of enzymes is deeply dependent on environmental conditions (Zhuang et al., 2016).

A differential regulation of patulin production was also observed in knockout mutants, which displayed increased patulin production at 5 dpi *in vitro.* This result

suggests that *gsfR1* could be involved in regulation of other SMs produced by *P. griseofulvum*, revealing a multiplicity of downstream pathways that are regulated by the same transcription factor.

A crosstalk between different SM gene clusters has been largely recognized. For instance, in *Aspergillus nidulans* the putative transcription factor encoded by *scpR* controls the expression of *inpA* and *inpB*, two genes involved in the biosynthesis of a NRP-like compound, and the same TF is also involved in regulation of asperfuranone biosynthesis, a polyketide compound (Bergmann et al., 2010).

The interconnection of signals and responses is also testified by the complex pattern of TFs that can putatively bind the promoters of *gsfR1* and *gsfA*, triggering the activation or repression of griseofulvin biosynthesis and regulating at the same time other SM gene clusters and many aspects of the biology of *P. griseofulvum*, such as conidiation and virulence. Surprisingly, in the recently released genome of griseofulvin producer *Xylaria flabelliformis* (Mead et al., 2019) the gene *gsfR1* was missing, pointing to a possible regulation of griseofulvin biosynthesis independently of *gsfR1* action.

Considering the second putative transcription factor of the cluster, the deletion of *gsfR2* seems to have no effect on the parameters of *P. griseofulvum* considered, and it is likely that the gene is involved in a different biosynthetic pathway. Indeed, next to *gsfR2* is located the gene *gsfK*, which encodes for a putative reductase whose deletion in *P. aethiopicum* did not affect the production of griseofulvin (Cacho et al., 2013). These findings led to the hypothesis that both *gsfK* and *gsfR2* are probably involved in the biosynthesis of another SM. Moreover, in *X. flabelliformis* (Mead et al., 2019), *gsfJ* and *gsfG* genes were missing, suggesting that these two genes are not necessary for griseofulvin production.

Therefore, considering also that the knockout of *gsfR2* does not impair griseofulvin production, we suggest that the core genes for the biosynthesis of this metabolite are only 7 instead of 13.

Role of global regulators in griseofulvin biosynthesis

Secondary metabolism gene clusters are usually regulated by several global regulatory proteins in response to many environmental stimuli, such as light, carbon and nitrogen levels, pH and redox status, iron concentration and signaling from other organisms (Brakhage, 2013; Hoffmeister and Keller, 2007). Many of these global regulators are well characterized and their putative binding sites in the promoter regions of genes could be detected.

One of the transcription factors predicted to regulate *gsfR1* and not *gsfA* is KXG54396.1, a *P. griseofulvum* gene similar to *mtfA*, whose deletion or overexpression in *A. flavus* cause a reduction if aflatoxins production (Zhuang et al., 2016). Environmental conditions influence the regulatory activity of the gene in *A. flavus*, where the deletion increases conidiation *in vitro* and reduces it *in vivo* (Zhuang et al., 2016).

Other transcription factors putatively binding *gsfR1* and not *gsfA* included NsdD, which is necessary for sexual development and represses asexual one (Han et al., 2001), a quinic acid utilization activator, an homolog of the fluconazole resistance protein and *stuA*, a gene able to regulate cell pattern formation (Miller et al., 1992), carbon metabolism, effector expression and the synthesis of penicillin and several mycotoxins (IpCho et al., 2010; Sigl et al., 2011; Yang et al., 2018). Another transcription factor, KXG45700.1, was similar to LreA, a protein involved in conidiation and response to light (Igbalajobi et al., 2019), while KXG46600.1 was similar to Res2, necessary for meiosis (Zhu et al., 1997). KXG47124.1 has 60% identity with the transcription factor Msn2 of *A. parasiticus*, involved in conidiation, hyphal growth and on occasionally stress response in several fungi (Liu et al., 2013; Song et al., 2018; Tian et al., 2017). KXG48399.1 is homologous to another transcription factor related to asexual development, *flbC*, whose knock-out causes delayed conidiation in *A. nidulans* (Kwon et al., 2010). The *gsfR1* promoter was

predicted to be bound also by homologs of PAP1, SebA and NapA, transcription factors involved in oxidative stress response (Asano et al., 2007; Dinamarco et al., 2012; Ikner and Shiozaki, 2005). The presence of the binding sites for this group of transcription factors suggest a complex regulation of *gsfR1*, deeply influenced by sexual and asexual development and environmental conditions.

The promoter of *gsfA* had only 7 unique binding sites for transcription factor, including SreP, a repressor of siderophore biosynthesis in presence of high iron concentrations (Haas et al., 1997). Other transcription factors putatively regulating *gsfA* included two regulators of filamentous growth, a transcriptional activator of gluconeogenesis, a pathway necessary in *A. nidulans* for growing with certain carbon sources metabolized via Acetyl-CoA (Hynes et al., 2007), and a homolog of *gal4*, necessary in *Saccharomyces cerevisiae* for growth using galactose or lactose as only carbon sources (Riley et al., 1987). This data seems to suggest an increase in *gsfA* expression, and therefore griseofulvin production, in the presence of difficult carbon sources, but further experiments would be required to check this possibility.

Many transcription factors binding the promoters of both *gsfR1* and *gsfA* were uncharacterized. However, among them, there were homologs of the well-known global regulators AreA and CreA, which are respectively a positive and a negative transcription factor acting in response to nitrogen or carbon (Katz et al., 2008; Wilson and Arst, 1998). Moreover, the binding sites of NirA and the carbon catabolite repressor MIG1 were also found in the promoters of both genes, increasing the possibilities of production of griseofulvin being heavily influenced by carbon and nitrogen availability (Burger et al., 1991; Randhawa et al., 2018). Binding sites for genes regulating the utilization of different carbon and nitrogen sources are also present, such as *facB*, regulating acetate utilization (Todd et al., 1998), *argR*, regulating arginine metabolism (Dubois et al., 1987), *nirA*, regulating nitrate utilization (Burger et al., 1991), *amdR*, regulating amides, omega amino acids and lactams catabolism (Andrianopoulos and Hynes, 1990), and the quinic acid

utilization activator (Wheeler et al., 1996). Other two transcription factors, the cutinase transcription factor 1 beta and SKN7, respectively increase the expression of cutinases and the resistance to ROS, suggesting a role of the griseofulvin cluster in pathogenesis (Cao et al., 2009), as already indicated by the pathogenicity trials.
Beside the previous ones, other identified transcription factors were related to sexual and asexual development, heat shock, stress resistance, cellobiose response, sulphur catabolism and assimilation, iron homeostasis and adhesion, and hyphal growth (**Table SB.2**).
Both promoters of *gsfR1* and *gsfA* also contain binding sites similar to those recognized by Apa-2 (Chang et al., 1993), strengthening the hypothesis that the griseofulvin regulatory system is somewhat similar to the aflatoxins one, regulated by MtfA.

Global regulation of griseofulvin biosynthesis is triggered by carbon and nitrogen

Nitrogen is a central element for the life of living organisms, and many specific regulatory genes are expressed when nitrogen sources such as ammonium, glutamate or glutamine are scarce. These genes encode transcription factors that repress the utilizations of easy to assimilate nitrogen sources, in the so-called mechanism "Nitrogen Metabolism Repression" (NMR), and they are responsible for the activation of pathways required for the uptake of alternative nitrogen sources (Marzluf, 1997; Tudzynski, 2014). AreA is one of the main enzymes involved in NMR, and putatively regulate griseofulvin gene cluster, acting both on *gsfR1* and *gsfA*. Similarly, the global repressor in response to carbon, CreA, was found to be putatively involved in the regulation of griseofulvin biosynthesis, together with transcription factors that modulate the utilization of different carbon sources.
The effect of carbon and nitrogen sources on growth, conidiation and griseofulvin biosynthesis was investigated *in vitro*, inoculating the strains in three different

media, characterized by different C/N ratios. Núñez and collaborators (2000) used MEA and MPA to evaluate the effect of carbon and nitrogen on the production of verrucosidin, showing that MEA induce a higher production of this mycotoxin and that these growth media highly influence mycotoxin production. Similarly, the wild type showed a similar griseofulvin production on PDA and MEA while in MPA a significant reduction was observed, which is reasonable because carbon is essential to build the polyketide structure.

On the other hand, the deletion of *gsfR1* seems to have a differential effect on MEA and MPA, where the production of griseofulvin drops drastically in deletants compared to wild type conversely to what observed on PDA. The high nitrogen concentration of these media likely triggered the NMR, resulting in the activation of SM gene clusters, which could explain the higher production of griseofulvin in the wild type compared to deletion mutants for *gsfR1*. This result also indicates that GsfR1 can regulate griseofulvin biosynthesis both in a positive and negative manner, depending on external stimuli.

Regulation of griseofulvin biosynthesis and conidiation

Fungal development and secondary metabolism are deeply connected and many examples have elucidated that some compounds can be secreted at the time of conidiation (especially mycotoxins), while other SM can even induce sporulation themselves, such as linoleic-acid in *A. nidulans* or zearalenone in *Fusarium graminearum,* and some of them are required for formation of sexual ad asexual spores and for their survival, such as pigments (Calvo et al., 2002). Knockout mutants for *gsfR1* produce less conidia compared to the parental strain, but this behavior is not surprising as deletion of regulatory genes is often associated with a reduced conidiation, such as for deletion of *aflR* TF for aflatoxin biosynthesis (Wilkinson et al., 2004). The reduced conidiation at first was thought to be related with the antifungal properties of griseofulvin, which is highly produced on PDA

from knockout mutants. From this perspective, the role of *gsfR1* is to avoid an excessive production of griseofulvin, which is known to be an antifungal and could have toxic effect on the producer. Nevertheless, a strongly reduced production of conidia was observed in MEA and MPA, where the mutants exhibit a similar and lower griseofulvin production, respectively, compared to PDA. Consequently, the reduced conidiation of *gsfR1* mutants cannot be considered linked with the antifungal activity of griseofulvin, but instead seems to be the result of both the influence of the culture media and the effect of deletion of *gsfR1* itself.

Concerning the culture media, MEA and MPA media were supplemented with peptone. The malt extract added in MEA provides carbon and nitrogen content, while in MPA the meat extract contributes to richness in nitrogen. Studying the aflatoxin biosynthesis it was previously highlighted how simple sugars support fungal growth, sporulation and aflatoxin production, unlike complex sugars and peptone, which strongly repress aflatoxin production (Calvo et al., 2002; Yao et al., 2018).

The reduced number of conidia produced by all tested strains in MEA and MPA seems therefore to be related with the nutrient composition of these media. Instead, the reduced conidiation of knockout mutants in all tested media compared to wild type could be considered a consequence of a direct or indirect positive regulation of conidiation driven by g*sfR1*. Many TFs have a role in co-regulating spore production and mycotoxin formation. MtfA, RtfA, NsdC, RafA and StuA all regulate aflatoxin production, conidiation and sclerotia development in *A. flavus* (Yao et al., 2018), and GsfR1 could act similarly. Furthermore, the regulator of patulin biosynthesis was recently proposed to be connected with the development of *P. expansum*, even if the authors suggests that this is only a remote possibility which must be proved (Snini et al., 2016). Moreover, *gsfR1* is putatively bound by transcription factors involved in asexual and sexual development, which strongly supports this model.

Further studies are needed to investigate the role of *gsfR1* in regulation of sporulation or conidiation, together with already known global transcription factors.

Role of griseofulvin on pathogenicity of *P. griseofulvum*

Considering that SM are not essential for primary metabolism of fungi, but they confer several advantages to the producer organism, it is reasonable to think that they could have a role in pathogenicity of fungal pathogens. The relationship between SM production and pathogenicity has been extensively investigated previously, and many examples of SM and mycotoxins directly involved in pathogenesis are well documented (Scharf et al., 2014). Considering causal agents of blue molds, the main studies have investigated the role of patulin on incidence of *P. expansum*, but contrasting results emerged. For instance, Ballester et al. (2015) and Lie et al. (2015) pointed out that patulin produced by fungal pathogens is not essential for apple infection, as deletion of *patK*, *patL*, and *patN*, genes involved for patulin biosynthesis, did not affect the virulence . Conversely, Sanzani et al. (2016) and Barad et al. (2013) observed a direct correlation with virulence, because mutants which displayed a reduced production of patulin were less virulent on apples. These opposite results obtained studying the same pathosystem can be explained by the usage of different strains from different countries and by the great quantity of factors involved *in vivo* such as temperature, pH, cultivar sensitivity and storage conditions (Barad et al., 2016b). Among these factors, apple variety plays an essential role as demonstrated by Snini et al. (2016). Taking into account the different contributions, patulin could be considered a virulence factor (connected to severity of disease), instead of a pathogenicity one (linked to the ability to cause the disease), helping to establish the disease when the pathogen has already colonized the fruit (Barad et al., 2016b). Indeed, the addition of patulin directly on the wound completely restored the wild type phenotype in less virulent strains (Snini et al., 2016).

According to the presented results, griseofulvin may act similarly. Indeed, the deletion mutants for the *gsfA* gene are still able to induce the disease, even if they are not able to produce griseofulvin, but the infected apples exhibit reduced lesions. This supports the idea that griseofulvin alone should not be considered a pathogenicity factor but can act to enhance virulence.

ΔgsfR1 mutants were slightly more virulent compared to the wild type, while Δ*gsfR2* strains were comparable to the wild type both in griseofulvin production and in pathogenicity. Considering that Δ*gsfR2* mutants displayed no differences on virulence compared to the wild type, the changes on development of rot diameter of knockout mutants for *gsfA* and *gsfR1* genes could be related to griseofulvin production. The differences seen in the development of rot diameter on apples by Δ*gsfA* and Δ*gsfR1* mutants could suggest a role of griseofulvin biosynthesis in virulence of *P. griseofulvum*. Nevertheless, the ectopic strain showed increased virulence compared to the wild type and therefore further studies are necessary.

Concerning the griseofulvin produced *in vivo* by Δ*gsfR1* mutants compared to the wild type, a higher production was observed at 10 dpi (similarly to what observed on PDA plates) and a lower production at 14 dpi (similarly to what observed on MEA and MPA plates).

In the proposed scenario, when establishing the interaction with the host, *P. griseofulvum* will activate a series of well characterized responses, such as biosynthesis of gluconic acid and ammonium depletion, in order to improve the acidification of apples and the biosynthesis of polygalacturonases responsible for cell wall degradation of fruits (Barad et al., 2016a; Prusky et al., 2004). This reduced nitrogen availability, together with high availability of simple sugars, can trigger the activation of SM gene clusters, including griseofulvin, which is also enhanced in presence of ROS and cutinases according to promoter analysis. This hypothesis is supported by the fact that *gsfR1* mutants produced less griseofulvin compared to wild type after 14 days, so it is reasonable to think that positive regulation of

griseofulvin biosynthesis is activated at this point. Moreover, if GsfR1 acts as a negative regulator, it is reasonable to observe an earlier production of griseofulvin in the Δg*sfR1* knockout mutants, as we have observed in infected apples by day 10, when the production of griseofulvin is higher in the knockout mutants. To our knowledge, this is the first attempt to investigate the role of griseofulvin in virulence of *P. griseofulvum* and further investigation is necessary to prove its involvement.

CONCLUSIONS

This work led to a better understanding of the complex regulation of griseofulvin biosynthesis. The role of putative transcription factors was investigated. From our results, *gsfR2* is not involved in this pathway and it is probably part of another gene cluster, together with the putative reductase *gsfK*, while *gsfG* and *gsfJ* are putatively not essential for griseofulvin biosynthesis due to their lack in the genome of the producer species *X. flabelliformis*. Therefore, the griseofulvin gene cluster should be considered composed of 7 genes, instead of 13, as previously described.

gsfR1, on the other hand, is involved in griseofulvin biosynthesis, acting as a negative regulator of the cluster on PDA and on apples. However, in different culture conditions, characterized by nitrogen and complex sugars richness, *gsfR1* could act as a positive regulator of griseofulvin biosynthesis. These findings suggest that *gsfR1* can trigger different responses depending on external stimuli, especially nitrogen and carbon availability. The deletion of *gsfR1* has a huge impact on many aspects of *P. griseofulvum,* ranging from secondary metabolism to virulence and conidiation. It is remarkable to notice that *gsfR1* seems to be involved in the regulation of patulin biosynthesis, which *P. griseofulvum* can produce in high amounts on apples. This finding could lead to new strategies to limit the production of patulin on apples.

Concerning pathogenicity of *P. griseofulvum*, the deletion mutants for *gsfA* were less virulent than the wild type strain, while the deletion mutants for *gsfR1* were

slightly more virulent *in vivo* suggesting a direct role of griseofulvin on virulence of *P. griseofulvum*. Until now, this is the first study on the role of griseofulvin in pathogenicity, and further investigations are needed to confirm this data and to highlight the mechanisms of action involved.

gsfr1 is also involved in the asexual multiplication of *P. griseofulvum*, as a reduced conidiation was observed compared to the wild type for all *in vitro* conditions tested. In conclusion, g*sfR1* represents a nice example of how the regulation of SM biosynthesis and fungal development can be complex and extremely interconnected.

ACKNOWLEDGMENTS

We thank Prof. Maria Lodovica Gullino for reviewing and editing of the manuscript and Dr. Houda Banani for setting up the knockout on *P. griseofulvum*. This work was supported by Fondazione Cassa di Risparmio di Cuneo (progetto SMART APPLE – Innovative and SMART technologies for sustainable apple production) and by the Spanish Ministry of Science, Innovation and Universities (AGL2017-88120-R, AEI/FEDER, UE).

REFERENCES

- Andrianopoulos, A., & Hynes, M. J. (1990). Sequence and functional analysis of the positively acting regulatory gene amdR from *Aspergillus nidulans*. *Molecular and Cellular Biology*, *10*(6), 3194-3203. https://doi.org/10.1128/mcb.10.6.3194
- Asano, Y., Hagiwara, D., Yamashino, T., & Mizuno, T. (2007). Characterization of the bZip-type transcription factor NapA with reference to oxidative stress response in *Aspergillus nidulans*. *Bioscience, Biotechnology, and Biochemistry*, *71*(7), 1800-1803. https://doi.org/10.1271/bbb.70133
- Bai, Y. B., Gao, Y. Q., Nie, X. D., Tuong, T. M., Li, D., & Gao, J. M. (2019). Antifungal Activity of Griseofulvin Derivatives against Phytopathogenic Fungi in Vitro and in Vivo and Three-Dimensional Quantitative Structure-Activity Relationship Analysis. *Journal of Agricultural and Food Chemistry*, *67*(22), 6125-6132. https://doi.org/10.1021/acs.jafc.9b00606

- Ballester, A. R., Marcet-Houben, M., Levin, E., Sela, N., Selma-Lázaro, C., Carmona, L., ... & Gabaldón, T. (2015). Genome, transcriptome, and functional analyses of *Penicillium expansum* provide new insights into secondary metabolism and pathogenicity. *Molecular Plant-Microbe Interactions*, *28*(3), 232-248. https://doi.org/10.1094/MPMI-09-14-0261-FI
- Banani, H., Marcet-Houben, M., Ballester, A. R., Abbruscato, P., González-Candelas, L., Gabaldón, T., & Spadaro, D. (2016). Genome sequencing and secondary metabolism of the postharvest pathogen *Penicillium griseofulvum*. *BMC Genomics*, *17*(1), 19. https://doi.org/10.1186/s12864-015-2347-x
- Barad, S., Espeso, E. A., Sherman, A., & Prusky, D. (2016). Ammonia activates pacC and patulin accumulation in an acidic environment during apple colonization by *Penicillium expansum*. *Molecular Plant Pathology*, *17*(5), 727-740.https://doi.org/10.1111/mpp.12327
- Barad, S., Sionov, E., & Prusky, D. (2016). Role of patulin in post-harvest diseases. *Fungal Biology Reviews*, *30*(1), 24-32. https://doi.org/10.1016/j.fbr.2016.02.001
- Bergmann, S., Funk, A. N., Scherlach, K., Schroeckh, V., Shelest, E., Horn, U., ... & Brakhage, A. A. (2010). Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. *Applied and Environmental Microbioogy*, *76*(24), 8143-8149. https://doi.org/10.1128/AEM.00683-10
- Bibbins, M., Crepin, V., Cummings, N., Mizote, T., Baker, K., Mellits, K., & Connerton, I. (2002). A regulator gene for acetate utilisation from *Neurospora crassa*. *Molecular Genetics and Genomics*, *267*(4), 498-505. https://doi.org/10.1007/s00438-002-0682-5
- Brakhage, A. A. (2013). Regulation of fungal secondary metabolism. *Nature Reviews Microbiology*, *11*(1), 21. https://doi.org/10.1038/nrmicro2916
- Brown, D. W., Butchko, R. A., Busman, M., & Proctor, R. H. (2007). The *Fusarium verticillioides* FUM gene cluster encodes a Zn (II) 2Cys6 protein that affects FUM gene expression and fumonisin production. *Eukaryotic Cell*, *6*(7), 1210-1218. https://doi.org/10.1128/EC.00400-06
- Burger, G., Strauss, J., Scazzocchio, C., & Lang, B. F. (1991). nirA, the pathway-specific regulatory gene of nitrate assimilation in *Aspergillus nidulans*, encodes a putative GAL4-type zinc finger protein and contains four introns in highly conserved

regions. *Molecular and Cellular Biology*, *11*(11), 5746-5755. https://doi.org/10.1128/mcb.11.11.5746

- Buron-Moles, G., López-Pérez, M., González-Candelas, L., Viñas, I., Teixidó, N., Usall, J., & Torres, R. (2012). Use of GFP-tagged strains of *Penicillium digitatum* and *Penicillium expansum* to study host-pathogen interactions in oranges and apples. *International Journal of Food Microbiology*, *160*(2), 162-170. https://doi.org/10.1016/j.ijfoodmicro.2012.10.005
- Cacho, R. A., Chooi, Y. H., Zhou, H., & Tang, Y. (2013). Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. *ACS Chemical Biology*, *8*(10), 2322-2330. https://doi.org/10.1021/cb400541z
- Calvo, A. M., Wilson, R. A., Bok, J. W., & Keller, N. P. (2002). Relationship between secondary metabolism and fungal development.Microbiology and Molecular Biology Reviews, *66*(3), 447-459. https://doi.org/10.1128/mmbr.66.3.447-459.2002
- Cao, C., Liu, W., & Li, R. (2009). *Penicillium marneffei* SKN7, a Novel Gene, could Complement the Hypersensitivity of S. cerevisiaeskn7 Disruptant Strain to Oxidative Stress. *Mycopathologia*, *168*(1), 23-30. https://doi.org/10.1007/s11046-009-9192-x
- Chang, P. K., Cary, J. W., Bhatnagar, D. E. E. P. A. K., Cleveland, T. E., Bennett, J. W., Linz, J. E., ... & Payne, G. A. (1993). Cloning of the *Aspergillus parasiticus* apa-2 gene associated with the regulation of aflatoxin biosynthesis. *Applied and Environmental Microbioogy*, *59*(10), 3273-3279.
- Chooi, Y. H., Cacho, R., & Tang, Y. (2010). Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chemistry & Biology*, *17*(5), 483-494. https://doi.org/10.1016/j.chembiol.2010.03.015
- Cramer, R. A., Perfect, B. Z., Pinchai, N., Park, S., Perlin, D. S., Asfaw, Y. G., ... & Steinbach, W. J. (2008). Calcineurin target CrzA regulates conidial germination, hyphal growth, and pathogenesis of *Aspergillus fumigatus*. *Eukaryotic Cell*, *7*(7), 1085-1097. https://doi.org/10.1128/EC.00086-08
- Damveld, R. A., Arentshorst, M., Franken, A., VanKuyk, P. A., Klis, F. M., Van Den Hondel, C. A., & Ram, A. F. (2005). The *Aspergillus niger* MADS-box transcription factor RlmA is required for cell wall reinforcement in response to cell wall stress. *Molecular Microbiology*, *58*(1), 305-319. https://doi.org/10.1111/j.1365-2958.2005.04827.x

- Dinamarco, T. M., Almeida, R. S., de Castro, P. A., Brown, N. A., dos Reis, T. F., Ramalho, L. N. Z., ... & Goldman, G. H. (2012). Molecular characterization of the putative transcription factor SebA involved in virulence in *Aspergillus fumigatus*. *Eukaryotic Cell*, *11*(4), 518-531. https://doi.org/10.1128/EC.00016-12
- Dubois, E., Bercy, J., & Messenguy, F. (1987). Characterization of two genes, ARGRI and ARGRIII required for specific regulation of arginine metabolism in yeast. *Molecular and General Genetics MGG*, *207*(1), 142-148. https://doi.org/10.1007/BF00331501
- Flaherty, J. E., & Woloshuk, C. P. (2004). Regulation of fumonisin biosynthesis in *Fusarium verticillioides* by a zinc binuclear cluster-type gene, ZFR1. *Applied and Environmental Microbioogy*, *70*(5), 2653-2659. https://doi.org/10.1128/aem.70.5.2653-2659.2004
- Frandsen, R. J., Andersson, J. A., Kristensen, M. B., & Giese, H. (2008). Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. *BMC Molecular Biology*, *9*(1), 70. https://doi.org/10.1186/1471-2199-9-70
- Haas, H., Angermayr, K., & Stöffler, G. (1997). Molecular analysis of a *Penicillium chrysogenum* GATA factor encoding gene (sreP) exhibiting significant homology to the *Ustilago maydis* urbs1 gene. *Gene*, *184*(1), 33-37. https://doi.org/10.1016/s0378-1119(96)00570-7
- Han, K. H., Han, K. Y., Yu, J. H., Chae, K. S., Jahng, K. Y., & Han, D. M. (2001). The nsdD gene encodes a putative GATA-type transcription factor necessary for sexual development of *Aspergillus nidulans*. *Molecular microbiology*, *41*(2), 299-309. https://doi.org/10.1046/j.1365-2958.2001.02472.x
- Hoffmeister, D., & Keller, N. P. (2007). Natural products of filamentous fungi: enzymes, genes, and their regulation. *Natural Product Reports*, *24*(2), 393-416. https://doi.org/10.1039/b603084j
- Hogues, H., Lavoie, H., Sellam, A., Mangos, M., Roemer, T., Purisima, E., ... & Whiteway, M. (2008). Transcription factor substitution during the evolution of fungal ribosome regulation. *Molecular Cell*, *29*(5), 552-562. https://doi.org/10.1016/j.molcel.2008.02.006
- Hynes, M. J., Szewczyk, E., Murray, S. L., Suzuki, Y., Davis, M. A., & Sealy-Lewis, H. M. (2007). Transcriptional control of gluconeogenesis in *Aspergillus nidulans*. *Genetics*, *176*(1), 139-150. https://doi.org/10.1534/genetics.107.070904

- Igbalajobi, O., Yu, Z., & Fischer, R. (2019). Red-and Blue-Light Sensing in the Plant Pathogen *Alternaria alternata* Depends on Phytochrome and the White-Collar Protein LreA. *mBio*, *10*(2), e00371-19. https://doi.org/10.1128/MBIO.00371-19
- Ikner, A., & Shiozaki, K. (2005). Yeast signaling pathways in the oxidative stress response. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *569*(1-2), 13-27. https://doi.org/10.1016/j.mrfmmm.2004.09.006
- IpCho, S. V., Tan, K. C., Koh, G., Gummer, J., Oliver, R. P., Trengove, R. D., & Solomon, P. S. (2010). The transcription factor StuA regulates central carbon metabolism, mycotoxin production, and effector gene expression in the wheat pathogen *Stagonospora nodorum*. *Eukaryotic Cell*, *9*(7), 1100-1108. https://doi.org/10.1128/EC.00064-10
- Katz, M. E., Bernardo, S. M., & Cheetham, B. F. (2008). The interaction of induction, repression and starvation in the regulation of extracellular proteases in *Aspergillus nidulans*: evidence for a role for CreA in the response to carbon starvation. *Current Genetics*, *54*(1), 47. https://doi.org/10.1007/s00294-008-0198-6
- Kellner, N., Heimel, K., Obhof, T., Finkernagel, F., & Kämper, J. (2014). The SPF27 homologue Num1 connects splicing and kinesin 1-dependent cytoplasmic trafficking in *Ustilago maydis*. *PLoS Genetics*, *10*(1), e1004046. https://doi.org/10.1371/journal.pgen.1004046
- Kovács, Z., Szarka, M., Kovács, S., Boczonádi, I., Emri, T., Abe, K., ... & Pusztahelyi, T. (2013). Effect of cell wall integrity stress and RlmA transcription factor on asexual development and autolysis in *Aspergillus nidulans*. *Fungal Genetics and Biology*, *54*, 1-14. https://doi.org/10.1016/j.fgb.2013.02.004
- Kwon, N. J., Garzia, A., Espeso, E. A., Ugalde, U., & Yu, J. H. (2010). FlbC is a putative nuclear C2H2 transcription factor regulating development in *Aspergillus nidulans*. *Molecular Microbiology*, *77*(5), 1203-1219. https://doi.org/10.1111/j.1365-2958.2010.07282.x
- Lee, K., Sung, C., Kim, B. G., & Hahn, J. S. (2013). Activation of Aro80 transcription factor by heat-induced aromatic amino acid influx in *Saccharomyces cerevisiae*. *Biochemical and Biophysical Research Communications*, *438*(1), 43-47. https://doi.org/10.1016/j.bbrc.2013.07.019
- Li, B., Zong, Y., Du, Z., Chen, Y., Zhang, Z., Qin, G., ... & Tian, S. (2015). Genomic characterization reveals insights into patulin biosynthesis and pathogenicity in Penicillium species. *Molecular Plant-Microbe Interactions*, *28*(6), 635-647.

https://doi.org/10.1094/MPMI-12-14-0398-FI

- Lints, R., Davis, M. A., & Hynes, M. J. (1995). The positively acting amdA gene of *Aspergillus nidulans* encodes a protein with two C2H2 zinc-finger motifs. *Molecular Microbiology*, *15*(5), 965-975. https://doi.org/10.1111/j.1365-2958.1995.tb02365.x
- Liu, Q., Ying, S. H., Li, J. G., Tian, C. G., & Feng, M. G. (2013). Insight into the transcriptional regulation of Msn2 required for conidiation, multi-stress responses and virulence of two entomopathogenic fungi. *Fungal Genetics and Biology*, *54*, 42-51. https://doi.org/10.1016/j.fgb.2013.02.008
- Marzluf, G. A. (1997). Genetic regulation of nitrogen metabolism in the fungi. *Microbiology and Molecular Biology Reviews*, *61*(1), 17-32.
- Mauro, V., Carette, D., Pontier-Bres, R., Dompierre, J., Czerucka, D., Segretain, D., ... & Pointis, G. (2013). The anti-mitotic drug griseofulvin induces apoptosis of human germ cell tumor cells through a connexin 43-dependent molecular mechanism. *Apoptosis*, *18*(4), 480-491. https://doi.org/10.1007/s10495-012-0800-8
- Mead, M. E., Raja, H. A., Steenwyk, J. L., Knowles, S. L., Oberlies, N. H., & Rokas, A. (2019). Draft Genome Sequence of the Griseofulvin-Producing Fungus Xylaria flabelliformis Strain G536. *Microbiology Resource Announcements*, *8*(38), e00890-19. https://doi.org/10.1128/MRA.00890-19
- Miller, K. Y., Wu, J., & Miller, B. L. (1992). StuA is required for cell pattern formation in Aspergillus. *Genes & Development*, *6*(9), 1770-1782. https://doi.org/10.1101/gad.6.9.1770
- Möller, E. M., Bahnweg, G., Sandermann, H., & Geiger, H. H. (1992). A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Research*, *20*(22), 6115. https://doi.org/10.1093/nar/20.22.6115
- Moslem, M., Abd-Elsalam, K., Yassin, M., & Bahkali, A. (2010). First morphomolecular identification of *Penicillium griseofulvum* and *Penicillium aurantiogriseum* toxicogenic isolates associated with blue mold on apple. *Foodborne Pathogens and Disease*, *7*(7), 857-861. https://doi.org/10.1089/fpd.2009.0507
- Moslem, M. A., Yassin, M. A., El-Rahim, A., El-Samawaty, M. A., Sayed, S. R., & Amer, O. E. (2013). Mycotoxin-producing Penicillium species Involved in Apple Blue Mold. *Journal of Pure and Applied Microbiology*, *7*(1), 187-193.

- Nobile, C. J., & Mitchell, A. P. (2005). Regulation of cell-surface genes and biofilm formation by the *C. albicans* transcription factor Bcr1p. *Current Biology*, *15*(12), 1150-1155. https://doi.org/10.1016/j.cub.2005.05.047
- Nunez, F., Diaz, M. C., Rodriguez, M., Aranda, E., Martin, A., & Asensio, M. A. (2000). Effects of substrate, water activity, and temperature on growth and verrucosidin production by *Penicillium polonicum* isolated from dry-cured ham. *Journal of Food Protection*, *63*(2), 231-236. https://doi.org/10.4315/0362-028X-63.2.231
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), e45-e45. https://doi.org/10.1093/nar/29.9.e45
- Prusky, D., McEvoy, J. L., Saftner, R., Conway, W. S., & Jones, R. (2004). Relationship between host acidification and virulence of *Penicillium spp*. on apple and citrus fruit. *Phytopathology*, *94*(1), 44-51. https://doi.org/10.1094/phyto.2004.94.1.44
- Randhawa, A., Ogunyewo, O. A., Eqbal, D., Gupta, M., & Yazdani, S. S. (2018). Disruption of zinc finger DNA binding domain in catabolite repressor Mig1 increases growth rate, hyphal branching, and cellulase expression in hypercellulolytic fungus *Penicillium funiculosum* NCIM1228. *Biotechnology for biofuels*, *11*(1), 15. https://doi.org/10.1186/s13068-018-1011-5
- Rathinasamy, K., Jindal, B., Asthana, J., Singh, P., Balaji, P. V., & Panda, D. (2010). Griseofulvin stabilizes microtubule dynamics, activates p53 and inhibits the proliferation of MCF-7 cells synergistically with vinblastine. *Bmc Cancer*, *10*(1), 213. https://doi.org/10.1186/1471-2407-10-213
- Reddy, K. R. N., Spadaro, D., Lore, A., Gullino, M. L., & Garibaldi, A. (2010). Potential of patulin production by *Penicillium expansum* strains on various fruits. *Mycotoxin Research*, *26*(4), 257-265. https://doi.org/10.1007/s12550-010-0064-5
- Ribeiro, A. I., Costa, E. S., Thomasi, S. S., Brandão, D. F. R., Vieira, P. C., Fernandes, J. B., ... & da Silva, M. F. D. G. F. (2018). Biological and Chemical Control of *Sclerotinia sclerotiorum* using *Stachybotrys levispora* and Its Secondary Metabolite Griseofulvin. *Journal of Agricultural and Food Chemistry*, *66*(29), 7627-7632. https://doi.org/10.1021/acs.jafc.7b04197
- Riley, M. I., Hopper, J. E., Johnston, S. A., & Dickson, R. C. (1987). GAL4 of

Saccharomyces cerevisiae activates the lactose-galactose regulon of Kluyveromyces lactis and creates a new phenotype: glucose repression of the regulon. *Molecular and Cellular Biology*, *7*(2), 780-786. https://doi.org/10.1128/mcb.7.2.780

- Scharf, D. H., Heinekamp, T., & Brakhage, A. A. (2014). Human and plant fungal pathogens: the role of secondary metabolites. *PLoS pathogens*, *10*(1), e1003859. https://doi.org/10.1371/JOURNAL.PPAT.1003859
- Sigl, C., Haas, H., Specht, T., Pfaller, K., Kürnsteiner, H., & Zadra, I. (2011). Among developmental regulators, StuA but not BrlA is essential for penicillin V production in *Penicillium chrysogenum*. *Applied and Environmental Microbioogy*, *77*(3), 972-982. https://doi.org/10.1128/AEM.01557-10
- Snini, S. P., Tannous, J., Heuillard, P., Bailly, S., Lippi, Y., Zehraoui, E., ... & Puel, O. (2016). Patulin is a cultivar-dependent aggressiveness factor favouring the colonization of apples by *Penicillium expansum*. *Molecular plant pathology*, *17*(6), 920-930. https://doi.org/10.1111/mpp.12338
- Song, Z., Yang, J., Xin, C., Xing, X., Yuan, Q., Yin, Y., & Wang, Z. (2018). A transcription factor, MrMsn2, in the dimorphic fungus *Metarhizium rileyi* is essential for dimorphism transition, aggravated pigmentation, conidiation and microsclerotia formation. *Microbial biotechnology*, *11*(6), 1157-1169. https://doi.org/10.1111/1751-7915.13302
- Spadaro, D., Garibaldi, A., & Gullino, M. L. (2008). Occurrence of patulin and its dietary intake through pear, peach, and apricot juices in Italy. *Food Additives and Contaminants*, *1*(2), 134-139. https://doi.org/10.1080/02652030802363790
- Spadaro, D., Lorè, A., Amatulli, M. T., Garibaldi, A., & Gullino, M. L. (2011). First Report of *Penicillium griseofulvum* causing blue mold on stored apples in Italy (Piedmont). *Plant Disease*, *95*(1), 76-76. https://doi.org/10.1094/PDIS-08-10-0568
- Spadaro, D., Lorè, A., Garibaldi, A., & Gullino, M. L. (2013). A new strain of *Metschnikowia fructicola* for postharvest control of *Penicillium expansum* and patulin accumulation on four cultivars of apple. *Postharvest Biology and Technology*, *75*, 1-8. https://doi.org/10.1016/j.postharvbio.2012.08.001
- Stierle, A. A., & Stierle, D. B. (2015). Bioactive secondary metabolites produced by the fungal endophytes of conifers. *Natural Product Communications*, *10*(10), Tang, H. Y., Zhang, Q., Li, H., & Gao, J. M. (2015). Antimicrobial and allelopathic metabolites produced by *Penicillium brasilianum*. *Natural Product Research*, *29*(4),

345-348. https://doi.org/10.1080/14786419.2014.940347

- Tian, L., Yu, J., Wang, Y., & Tian, C. (2017). The C2H2 transcription factor VdMsn2 controls hyphal growth, microsclerotia formation, and virulence of *Verticillium dahliae*. *Fungal Biology*, *121*(12), 1001-1010. https://doi.org/10.1016/j.funbio.2017.08.005
- Todd, R. B., Andrianopoulos, A., Davis, M. A., & Hynes, M. J. (1998). FacB, the *Aspergillus nidulans* activator of acetate utilization genes, binds dissimilar DNA sequences. *The EMBO Journal*, *17*(7), 2042-2054. https://doi.org/10.1093/emboj/17.7.2042
- Tudzynski, B. (2014). Nitrogen regulation of fungal secondary metabolism in fungi. *Frontiers in Microbiology*, *5*, 656. https://doi.org/10.3389/fmicb.2014.00656
- Vienken, K., Scherer, M., & Fischer, R. (2005). The Zn (II) 2Cys6 putative *Aspergillus nidulans* transcription factor repressor of sexual development inhibits sexual development under low-carbon conditions and in submersed culture. *Genetics*, *169*(2), 619-630. https://doi.org/10.1534/genetics.104.030767
- Wheeler, K.A., Lamb, H.K., & Hawkins, A.R. (1996). Control of metabolic flux through the quinate pathway in *Aspergillus nidulans*. Biochemical Journal, 315, 195–205. https://doi.org/10.1042/bj3150195
- Wilkinson, H. H., Ramaswamy, A., Sim, S. C., & Keller, N. P. (2004). Increased conidiation associated with progression along the sterigmatocystin biosynthetic pathway. *Mycologia*, *96*(6), 1190-1198. https://doi.org/10.1080/15572536.2005.11832867
- Wilson, R. A., & Arst, H. N. (1998). Mutational analysis of AREA, a transcriptional activator mediating nitrogen metabolite repression in *Aspergillus nidulans* and a member of the “streetwise” GATA family of transcription factors. *Microbiology and Molecular Biology Reviews*, *62*(3), 586-596.
- Wolfers, S., Kamerewerd, J., Nowrousian, M., Sigl, C., Zadra, I., Kürnsteiner, H., ... & Bloemendal, S. (2015). Microarray hybridization analysis of light-dependent gene expression in *Penicillium chrysogenum* identifies bZIP transcription factor PcAtfA. *Journal of Basic Microbiology*, *55*(4), 480-489. https://doi.org/10.1002/jobm.201400588
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, International Agency for Research on Cancer, & World Health Organization. (2001). *Some thyrotropic agents* (Vol. 79). World Health Organization.

- Wu, G., Jurick II, W. M., Lichtner, F. J., Peng, H., Yin, G., Gaskins, V. L., ... & Bennett, J. W. (2019). Whole-genome comparisons of *Penicillium spp.* reveals secondary metabolic gene clusters and candidate genes associated with fungal aggressiveness during apple fruit decay. *PeerJ*, *7*, e6170. https://doi.org/10.7717/peerj.6170
- Yang, W., Wu, H., Wang, Z., Sun, Q., Qiao, L., & Huang, B. (2018). The APSES gene MrStuA regulates sporulation in *Metarhizium robertsii*. *Frontiers in Microbiology*, *9*, 1208. https://doi.org/10.3389/fmicb.2018.01208
- Yao, G., Yue, Y., Fu, Y., Fang, Z., Xu, Z., Ma, G., & Wang, S. (2018). Exploration of the regulatory mechanism of secondary metabolism by comparative transcriptomics in *Aspergillus flavus*. *Frontiers in Microbiology*, *9*, 1568. https://doi.org/10.3389/fmicb.2018.01568
- Zhou, X., Liu, W., Wang, C., Xu, Q., Wang, Y., Ding, S., & Xu, J. R. (2011). A MADS-box transcription factor MoMcm1 is required for male fertility, microconidium production and virulence in *Magnaporthe oryzae*. *Molecular Microbiology*, *80*(1), 33-53. https://doi.org/10.1111/j.1365-2958.2011.07556.x
- Zhu, Y., Takeda, T., Whitehall, S., Peat, N., & Jones, N. (1997). Functional characterization of the fission yeast Start-specific transcription factor Res2. *The EMBO journal*, *16*(5),
- Zhuang, Z., Lohmar, J., Satterlee, T., Cary, J., & Calvo, A. (2016). The master transcription factor mtfA governs aflatoxin production, morphological development and pathogenicity in the fungus *Aspergillus flavus*. *Toxins*, *8*(1), 29. https://doi.org/10.3390/toxins8010029

SUPPLEMENTARIES

Supplementary Ffgure B.1. PCR analysis of *gsfA* mutants. Schematic presentation of the *gsfA* locus in the wild type and deletion mutants, primers used are presented as arrows (A). Amplification of hygromycin resistance cassette (*HygR*) and *gsfA* gene (B); confirmation of orientation of inserted T-DNA (C); copies number of T-DNA inserted in *P. griseofulvum* (D), obtained through qPCR on gDNA using primers 5F/6R and β-tubulin gene as reference gene. M= GelPilot Wide Range

Ladder, WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfA* and E = ectopic strains, Neg. = negative control (PCR mix without DNA).

Supplementary figure B.2. Effect of *gsfA* deletion on *P. griseofulvum.* Colony diameter (**A**) and griseofulvin production (**B**) from 3 to 7 dpi. Plate view (front **C** and reverse **D**) at 7 dpi on PDA. WT = wild type *P. griseofulvum*, Δ = deletion mutants for *gsfA* and E = ectopic strains, n.d. = non detectable.

Supplementary table B.1. List of primers used in this work.

Supplementary table B.2. Transcription factors putatively binding the promoter of *gsfR1* and *gsfA.* It is reported the name of homolog and the regulated processes in other organisms with references when available.

All the supplementary material is available online at: https://drive.google.com/drive/folders/1MAF2Vag3B5sAgS9BPtPI4AOFZWItRAkB?usp=sharing

ACKNOWLEDGMENTS

Wow, the long road is finally over! It is impossible to list all the people that made this possible, but I will try!

I would have to start with **my parents**, being born was an huge advantage in my reasearch activity and I would never have completed my thesis without it. Also, the people responsible with spawning me were very patient and kind in withstanding my endless complaining before and during the PhD. That is done, I will quit complaining now, after all I got my PhD, so there is nothing left to whine about.

My parents did endure my complaining the longest, but **my friends** did it the most, so I have to acknowledge them too. There are many of them, but I think **Alessandra Guerrieri** deserves a special mention for listening to me regularly for more than three years while being abroad. She complained a lot too, though, so I'm not really the bad guy here.

It's colleagues turn! And also crossover time, because I was lucky enough to meet at work people that became close friends and companions to me. The first among them is **Silvia Valente**, my lab-bro, who was fondamental in her moral, scientific and bureacratic support (I hate bureacracy and I kept forgetting modules and deadlines and other evil stuff). She once offered to give me biscuits, I said “yes” and she answered “too bad there aren’t any!”, **passing me an empy biscuit box** with a **villainous laughter**… And I still feel like I have to thank her here, so that should tell you how much I owe her! Other very important people are **Slavica Matić**, my lab mother, which kept me fed during the long working hours, and **Alexandra Cucu**, my lab sister, which kept me from eating too much of Slavica’s food during the long working hours. **Simona Prencipe**, lab big sister, gifted with wisdom beyond human comprehension, was instrumental in keeping me doing research, and not just throwing science at walls to see if it stuck. **Betta the chemistry person** was as importanto to me as Panoramix is to Asterix: she knew mysterious and esoteric chemistry things, and was kind enough to share them with me. **Domenico Bertetti**

and **Renata Luongo**, lab technicians of infinite patience, turned me from a good-for-nothing ignorant child into a good-for-something ignorant child. **Davide Spadaro**, my most gracious supervisor, source of divine truths and divine duties, was the person that somehow thought I would have made a good PhD student. I hope he was right, for he was a good supervisor: I learnt so much in these years, and I feel like I became exactly the sort of scientist I wanted to be, so thanks! **Maria Lodovica Gullino**, the director of Agroinnova, for her leadership. She was able to push me to work more and to work harder. People tell me I should also thank **the Italian Ministry of Education, University and Research**, for giving me money. Ok I wrote so much and there are still so many people! I wish to thank, in no particular order: **Martina Pellicciaro**, **Giulia Tabone**, **Andrea Masino**, **Alberto Acquadro**, **Toni Gabaldòn**, **Samir Droby**, **Antonio Biasi**, **Alessio Bellini**, **Matteo Schiavinato**, **Marina Marcet Houben**, **Houda Banani**, **Fabiano Sillo.**
Finally, I wish to thank **Giulia Tettamanti**, you are the strength of my life.

CURRICULUM VITAE

Edoardo Piombo was born on November 17th, 1992 in Ivrea (TO). He performed there the studies and in July 2011 he obtained the high school diploma in classical studies at "Liceo classico statale Carlo Botta".

He enrolled at the Bachelor in Biotechnologies at the University of Turin in September 2011, and in his thesis he studied the topic of cadmium contaminated soil phytoremediation, under the supervision of Professor Andrea Schubert. He graduated in July 2014 with 110 cum laude/110.

After graduation, he attended the Master in "Plant Biotechnologies" at the University of Turin, where he graduated in 2016 with 110 cum laude/110 and special mention. His research thesis was titled "Comparative genomic and phenotypic analysis of pathways involved in secondary metabolites production in *Fusarium fujikuroi*" and was supervised by Professor Alberto Acquadro and Professor Davide Spadaro.

In October 2016 he started the PhD in Biology and Applied Biotechnology at the University of Turin, tutored by Professor Davide Spadaro. His research topic was called "Study of the molecular plant-fungal pathogen interactions". He performed the experiments and data analysis in Grugliasco at the Centre of Competence for the innovation in the agro-environmental field (Agroinnova) and he was founded by Italian Ministry of Education, University and Research (MIUR).

From August 2018 to October 2018 he spent three months as visiting student at Centre for Genomic Regulation (CRG) of Barcelona, following a project on secondary metabolites prediction and comparative genomics on *Penicillium spp.*, supervised by Professor Toni Gabaldón. From January 2019 to March 2019 he worked instead as a visiting student at the Volcani Center in Israel, under the

supervision of Professor Samir Droby. His work there focused on the study of the apple microbiome by both amplicon-based and *de novo* metagenomics.

PUBLICATIONS

ISI Journals

- Piombo, E., Bosio, P., Acquadro, A., Abbruscato, P., and Spadaro, D. (2019). Different phenotypes, similar genomes: three newly sequenced *Fusarium fujikuroi* strains induce different symptoms in rice depending on temperature. *Phytopathology*, (ja). https://doi.org/10.1094/PHYTO-09-19-0359-R
- Piombo, E., Sela, N., Wisniewski, M., Hoffmann, M., Gullino, M. L., Allard, M. W., Levin, E., Spadaro, D. and Droby, S. (2018). Genome sequence, assembly and characterization of two *Metschnikowia fructicola* strains used as biocontrol agents of postharvest diseases. *Frontiers in Microbiology* 9: 593. https://doi.org/10.3389/fmicb.2018.00593

Non-ISI Journals

- Banani, H., Piombo, E., Siciliano, I., Acquadro, A., Garibaldi, A., Gullino, M. L., and Spadaro, D. (2017) Genome mining, secondary metabolism and pathogenicity of three strains of *Fusarium fujikuroi*. *Protezione delle Colture* 10: 25.
- Piombo, E., Acquadro, A., Siciliano, I., Gullino, M. L., Garibaldi, A., and Spadaro, D. (2019) Analisi dell'effetto della temperatura sui sintomi del bakanae del riso. *Protezione delle Colture* 12: 68-69.
- Piombo, E., Gullino, M. L., Garibaldi, A., and Spadaro, D. (2019) Analisi genomica di *Fusarium fujikuroi* per l'identificazione di geni chiave per la patogenesi su riso. *Protezione delle Colture* 12: 69
- Piombo, E., Sela, N., Olari, S., Wisniewski, M., Hoffmann, M., Gullino, M. L., Wallard, M., Levin, E., Spadaro, D., and Droby, S. (2018) Sequenziamento, assemblaggio e caratterizzazione di due ceppi di *Metschnikowia fructicola* usati nella lotta biologica alle malattie postraccolta. *Protezione delle Colture* 11: 40
- Spadaro, D., Prencipe, S., Valente, S., Piombo, E., Garibaldi, A., and Gullino M. L. (2019) Presenza di funghi micotossigeni e gestione del rischio di contaminazione da micotossine nella frutta secca. *Protezione delle Colture* 12: 17-25.
- Valente, S., Piombo, E., Prencipe, S., Meloni, R. G., Gullino, M. L., Garibaldi, A., Gabaldòn, T., and Spadaro D. (2019) Genomica e HPLC-MS/MS per esplorare il potenziale micotossigeno di *Penicillium* spp. *Protezione delle Colture* 12: 84-85.

Abstracts at Congresses

- Abdelfattah, A., Danino, Y., Piombo, E., Raphael, G., Feygenberg, O., Wisniewski, M., and Droby, S. (2019) High-throughput sequencing gives new insights on dates microbiome: the influence of developmental stage, tissue and growing season. *Book of Abstracts of the International Workshop The Fruit Microbiome: A New Frontier*, 9-13 September 2019, Leesburg, USA, 41.
- Aragona, M., Piombo, E., Campos, L., Spadaro, D., Gullino M. L. San Segundo De Los Mozos, B., and Infantino, A. (2018) New insights into the infection process of Fusarium fujikuroi in rice using a GFP expressing isolate. *Book of abstracts of the BSPP 2018 Presidential Conference: "Imaging Plant Microbe Interactions"*, 10-11 December 2018 Warwick, United Kingdom.
- Freilich, S., Vetcos, M., Malik, A., Medina, S., Piombo, E., Cohen, M., Ofek-Lalzar, M., Faiganbion-Doron, A., Mazzola, M. (2019) Computational approaches for deciphering the functions of microbial communities. *Book of Abstracts of the International Workshop The Fruit Microbiome: A New Frontier*, 9-13 September 2019, Leesburg, USA, 22.
- Piombo, E., Banani, H., Siciliano, I., Abbruscato, P., Acquadro, A., Gullino, M. L., and Spadaro, D. (2017) Genome mining, pathogenicity and secondary metabolism of three strains of *Fusairum fujikuroi*, the causal agent of bakanae disease on rice. *Abstract Book of the 29th Fungal Genetics Conference*, 14-19 March 2017, Pacific Grove, USA, 157.
- Piombo, E., Banani, H., Siciliano, I., Abbruscato, P., Acquadro A., Gullino, M. L., and Spadaro, D. (2016) Genome mining, pathogenecity and secondary metabolism of three strains of *Fusarium fujikuroi*, the causal agent of bakane disease on rice. *Journal of Plant Pathology* 98: 59.
- Piombo, E., Campos, L., Spadaro, D., Gullino, M. L., San Segundo De Los Mozos, B., Infantino, A., and Aragona, M. (2018) New insights into the infection process of *Fusarium fujikuroi* in rice using a GFP expressing isolate. *Journal of Plant Pathology* 100: 644.
- Piombo, E., Rosati, M., Gullino, M. L., and Spadaro, D. (2019) Sequencing of avirulent strains of *Fusarium fujikuroi* contributes to clarify the pathogenicity mechanism on rice. *Journal of Plant Pathology* 101: in print.
- Piombo, E., Rosati, M., and Spadaro, D. (2020) Understanding the bakanae disease: looking for disease-related genes through the study of avirulent strains of *Fusarium fujikuroi*. *Book of Abstracts of the 15th European Conference on Fungal Genetics*, 17-20 February 2020, Rome. In print.

- Piombo, E., Sela, N., Wisniewski, M., Hoffmann, M., Gullino, M. L., Wallard, M., Levin, E., Spadaro, D., and Droby, S. (2018) Exploring the genome of *Metschnikowia fructicola*, a biocontrol yeast effective against postharvest diseases. *Abstracts of the 11th International Congress of Plant Pathology (ICPP 2018)*, 29 July-3 August 2018, Boston, USA. *Phytopathology*, 108 (10S), S1.114
- Piombo, E., Sela, N., Wisniewski, M., Hoffman, M., Gullino, M. L., Allard, M. W., Levin, E., Spadaro, D., and Droby, S. (2018) Genome and transcriptome of two strains of *Metschnikowia fructicola* provide new insight on the biocontrol mechanisms against postharvest diseases. *Journal of Plant Pathology* 100: 644.
- Piombo, E., Sela, N., Wisniewski, M., Hoffmann, M., Gullino, M. L., Allard, M., Levin, E., Spadaro, D., and Droby, S. (2018) *Metschnikowia fructicola*, a biocontrol yeast against postharvest diseases: genome sequence, assembly and characterization. *Book of Abstracts of the 14th European Conference on Fungal Genetics*, 25-28 February 2018, Haifa, Israel, 229.
- Spadaro, D., and Piombo, E., (2019) Towards metagenomics-based diagnostics to detect and prevent plant pathogens. *Book of Abstracts of the International Workshop The Fruit Microbiome: A New Frontier*, 9-13 September 2019, Leesburg, USA, 33.
- Spadaro, D., Piombo, E., Valente, S., Banani, H., Prencipe, S., Marcet-Houben, M., Gullino, M. L., and Gabald ón, T. (2019) Scanning genomes to identify secondary metabolite production by postharvest pathogens. Book of *Abstracts V International Symposium on Postharvest Pathology*, 19-24 May 2019, Liège, Belgium. S-III-O1, 50.
- Spadaro, D., Prencipe, S., Valente, S., Piombo, E., Garibaldi, A., and Gullino, M. L. (2020) Innovative management strategies for Aspergillus spp. and Penicillium spp. on nuts. *Plant Pathology in the 21th century: Contribution to the 11th International ISPP congress. Postharvest Pathology: Next Generation Solutions to Reducing Losses and Enhancing Safety.* (Spadaro, D., Droby, S., Gullino, M.L.), Springer, Dordrecht, The Netherlands, in print
- Valente, S., Cometto, A., Meloni, G. R., Piombo, E., Ballester, A. R., Gonzalez Candelas, L., Gullino, M. L., and Spadaro, D. (2019) Regulation of griseofulvin biosynthesis and role on growth and virulence of *Penicillium griseofulvum,* causal agent of apple blue mold. *Journal of Plant Pathology* 101: in print.
- Valente, S., Cometto, A., Piombo, E., Ballester, A. R., Gullino, M. L., Gonzalez-Candelas, L., and Spadaro, D. (2019) New insights into

griseofulvin biosynthesis by *Penicillium griseofulvum*, an agent of blue mould on apples. *Book of Abstracts V International Symposium on Postharvest Pathology*, 19-24 May 2019, Liège, Belgium. P015, 61.

- Valente, S., Cometto, A., Meloni, G. R., Piombo, E., Ballester, A. R., Gonzalez-Candelas, L., Gullino, M. L., and Spadaro, D. (2020) Role of GsfR1 and global regulators on griseofulvin and other secondary metabolites biosynthesis and on growth and virulence of *Penicillium griseofulvum*. *Book of Abstracts of the 15th European Conference on Fungal Genetics,* 17-20 February 2020, Rome. In print