University of São Paulo
"Luiz de Queiroz" College of Agriculture

Comparative analysis of the transcriptional profile of the sugarcane smut pathogen during the interaction with resistant and susceptible genotypes

Renato Gustavo Hoffmann Bombardelli

Dissertation presented to obtain the degree of Master in Science. Area: Genetics and Plant Breeding

Piracicaba
2020

Renato Gustavo Hoffmann Bombardelli
Agronomist Engineer

Comparative analysis of the transcriptional profile of the sugarcane smut pathogen during the interaction with resistant and susceptible genotypes

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. PhD. **CLAUDIA BARROS MONTEIRO VITORELLO**

Dissertation presented to obtain the degree of Master in Science. Area: Genetics and Plant Breeding

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Bombardelli, Renato Gustavo Hoffmann

Comparative analysis of the transcriptional profile of the sugarcane smut pathogen during the interaction with resistant and susceptible genotypes / Renato Gustavo Hoffmann Bombardelli. - - Versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2020.

81 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. RNA-Seq 2. *Saccharum* spp. 3. *Sporisorium scitamineum* I. Título

With all my love and afection,
To my beloved mother Lilian,
And my dearest grandparents Armindo and Arnilda.

ACKNOWLEDGMENTS

I would like to acknowledge the University of São Paulo, specially the “Luiz de Queiroz” College of Agriculture campus (ESALQ) for all the support.

I thank all the funding institutions that support me in this study. I thank the São Paulo Research Foundation (FAPESP) for the regular financial support and scholarships (Grant member: 2019/12788-8). I also thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 for previous scholarships.

I would like to specially acknowledge my advisor Prof. PhD. Claudia Barros Monteiro Vitorello, for welcoming me, for your advices and shared knowledge during those two years. I could not put into words all my affection for your figure as a brilliant and examplar scientist, strong woman and benevolent friend. Thank you for this amazing journey.

I thank my family for all the support, always being there for me, and encouragement for always keep the strength to carry on.

To all my lab partners: Hugo Rody, Gustavo Crestana, Mariana Marrafon, Juliana Calderan, Laina Oliveira, Marcella Ferreira, Jéssica Mendes, Gustavo Lima, Pedro Vilanova, Mariana Strabello, Natalia Teixeira, Andressa Bini, Clesivan Pereira, Deepak Sehgal, thank you for your patience, friendship and for making the work environment my second home.

To all my university friends: Gustavo Crestana, Hugo Rody, Leonardo Piacenti, Vinícius Nicoletti, Leandro Lemos, Pedro Berreta, Pedro Catto, Felipe Lieber, for sincere friendship and wonderful moments.

I thank all professors from the Department of Genetics, for their valuable lessons and contribution to my professional training.

I thank the examiners for accepting my invitaion, contributing to this work and my professional formation.

EPIGRAPH

“Na vida nem tudo são flores, muitas vezes são chicotes”.

“What we know is a drop, what we don't know is an ocean.”

— Isaac Newton

CONTENTS

RESUMO

Análise comparativa do perfil transcricional do fungo causador do carvão da cana-de-açúcar durante a interação com plantas resistentes e suscetíveis à doença

O carvão da cana-de-açúcar, *Sporisorium scitamineum*, é um fungo dimórfico e biotrófico, dependente do hospedeiro para sua reprodução sexual. A doença está espalhada por todos os países produtores da cana-de-açúcar, com exceção de Fiji, um grupo de ilhas vulcânicas na ilha Sul Pacífica da Oceania, e causa perdas econômicas com a redução da produtividade. Durante os últimos anos, o Grupo de Genômica da ESALQ/USP vem estudado extensivamente a interação carvão-cana, considerando seus vários aspectos, incluindo os mecanismos de ataque do patógeno, e resposta de defesa da planta. Os patógenos secretam moléculas conhecidas como efetores a fim de modular a fisiologia da planta, contra-atacar e proteger a si mesmo contra as barreiras impostas pela planta hospedeira. No presente trabalho, nós predizemos o secretoma e candidatos a efetores do *S. scitamineum*. Nós usamos a técnica do RNA-Seq para determinar o perfil de expressão e comparar a colonização de dois genótipos de cana-de-açúcar com contrastantes níveis de resistência (resistente, SP80-3280 e suscetível, IAC66-6 ao carvão) 48 horas após a inoculação. No primeiro capítulo, apresentamos uma breve revisão sobre o carvão da cana-de-açúcar. O segundo capítulo contém o estudo fornecendo a análise dos dados do transcriptoma do *S. scitamineum* quando infectando dois genótipos contrastantes de cana-de-açúcar às 48 horas após a inoculação e a comparação com dados já publicados do crescimento do *S. scitamineum* em meio de cultura. Nós encontramos resultados convincentes considerando-se o perfil de expressão dos genes do fungo infectando genótipos resistentes e suscetíveis de cana-de-açúcar. O patógeno expressa mais peroxidases, incluindo as catalases KatE e KatG, quando infectando o genótipo resistente, consistentemente com as respostas iniciais de defesa envolvendo a produção de espécies reativas de oxigênio (ROS) pelo genótipo SP80-3280. Nós mostramos a expressão variável de candidatos a efetores em cada tratamento, identificando promissores candidatos para futuros estudos funcionais. Explorando os genes relacionados ao cruzamento e crescimento filamentoso, uma fase crucial nos estágios iniciais da infecção, levando a formação da hifa dicariótica infectiva, nós mostramos a expressão de genes da via cAMP/PKA em *S. scitamineum*. Juntamente com o equilíbrio REDOX (produção/detoxificação do peróxido de hidrogênio), o fungo necessita da expressão de genes relacionados ao cruzamento para a infecção com suscesso. Nesse sentido, nós providenciamos intuições na infecção diferencial de genótipos de cana-de-açúcar resistente e suscetível pelo patógeno do carvão, sugerindo alguns candidatos para futuras caracterizações funcionais a fim de incrementar a compreensão desse patossistema.

Palavras-chave: RNA-Seq, *Saccharum* spp., *Sporisorium scitamineum*

ABSTRACT

Comparative analysis of the transcriptional profile of the sugarcane smut pathogen during the interaction with resistant and susceptible genotypes

The sugarcane smut pathogen, *Sporisorium scitamineum*, is a dimorphic and biotrophic fungus, host-dependent to its sexual reproduction. The disease spread across all the sugarcane producing countries, except in Fiji, the group of volcanic islands in the South Pacific island in Oceania, causes economic losses reducing the yield. During the later years, the Genomics Group at ESALQ/USP has been extensively studying the sugarcane-smut interaction, considering its various aspects, including the pathogen attack mechanisms and plant defense responses. The pathogens secrete molecules known as effectors to modulate plant physiology, counterattack, and protect themselves against the plant host's defense barriers. In this work, we predicted the secretome and candidate effectors of the *S. scitamineum.* We used the RNA-Seq technique to determine the expression profile and compare the colonization of two sugarcane genotypes with contrasting resistance levels (smut-resistant, SP80-3280 and -susceptible, IAC66-6) 48 hours after inoculation. In the first chapter, we provided a brief review of the sugarcane smut disease. The second chapter contains the study providing data analysis of the transcriptome of *S. scitamineum* when infecting two contrasting sugarcane genotypes 48 hours after inoculation and a comparison with previous data of *S. scitamineum* growth in axenic culture. We found compelling results considering genes' expression profile of the fungus infecting resistant and susceptible sugarcane genotypes. The pathogen expressed more peroxidases, including the catalases KatE and KatG, when infecting the resistant genotype, consistent with the early defense response involving reactive oxygen species (ROS) production by the SP80-3280 genotype. We revealed the variant expression of candidate effectors genes in each of the treatments, identifying promising candidates for further functional studies. Exploring the genes related to mating/filamentation growth, a crucial phase in the early infection stage, leading to the formation of the infective dikaryotic hyphae, we unveil the expression of genes of the cAMP/PKA pathway in *S. scitamineum*. Together with the REDOX equilibrium (production/detoxification of hydrogen peroxide), the fungus requires the expression of mating-related genes for successful infection. Herein we provided insights into the differential infection of resistant and susceptible sugarcane genotypes by the smut pathogen, suggesting some candidate genes for further functional characterization studies to increase this pathosystem's comprehension.

Keywords: RNA-Seq, *Saccharum* spp., *Sporisorium scitamineum*

CHAPTER 1: Sugarcane smut: the state of art

1. Sugarcane

Sugarcane is one of the most important crops cultivated globally, being the third most-produced commodity. Currently, around 90 countries cultivate sugarcane. Brazil is the leading producer, with a crop equivalent to 39.3% of global stock, followed by India, China, and Thailand (FAO, 2018). The 2018/19 Brazilian harvest generated a total of 620.4 million tons cultivated on an area of 8.6 million hectares in the country (CONAB, 2019). Besides the production of sugar and biofuels such as ethanol, the most used alternative fuel, the sugarcane crop has potential to produce other renewable products such as bioplastics, bio-hydrocarbons, and bio-electricity (Waclawovsky *et al.*, 2010), which has increased the interest in sugarcane utilization by its various producing countries (Lam *et al.*, 2009).

Modern sugarcane varieties (*Saccharum* spp.) come from a complex hybridization among *Saccharum* species, with significant contributions from the *S. spontaneum* (2n = 40 to 128) and *S. officinarum* (2n = 80) (D'Hont, 2005; Amalraj and Balasundaram, 2006). As a result of this interspecific hybridization, the sugarcane has an elevate polyploidy and aneuploidy degree with chromosomes number varying from 100 to 130 (2n) (D'Hont *et al.*, 1996), and its genome size estimate to approximately 10 Gbp (D'Hont and Glaszman, 2001). Various techniques as the sequencing of expressed sequence tags (ESTs) (Vettore *et al.*, 2001, 2003), transcriptome based on RNA-Seq (Cardoso-Silva *et al.*, 2014; Schaker *et al.*, 2016), the construction of a monoploid genome sequence of a commercial variety (R570) using Bacterial artificial chromosomes (BACs) (Garsmeur *et al.*, 2018), the sequencing of another commercial variety (SP80-3280) predicting 373 thousands of genes (Souza *et al.*, 2019), the partial genome sequence based on selected inserts cloned in BACs (de Setta *et al.*, 2014) and the sequencing of a monoploid/tetraploid version from the parental species *S. spontaneum* (Zhang *et al.*, 2018) are improving the knowledge of the complex sugarcane genome, as well as helping to uncover the behavior of sets of genes involved in essential processes such as the accumulation of sugar and resistance to plagues and diseases.

2. Smut disease

The fungi of *Ustilaginomycetes* class in the *Basidiomycota* phylum, known as smuts, including more than 1,650 species (Toh and Perlin, 2016) are host-specific pathogens that affect many angiosperm clades, being most common on the *Poaceae* family (Begerow *et al.*, 2014).

The only other pathogenic class with a more significant number of species is that of rust pathogens.

Several species in the smut group cause disease in important cereal crops such as maize – *Ustilago maydis* / *Sporisorium reilianum*; sorghum – *S. reilianum*; oat and barley – *U. hordei*; wheat – *U. tritici* and sugarcane – *S. scitamineum*. The first sugarcane smut report dated from 1877 in South Africa, describing the disease in a *Saccharum sinense* clone, known as "China cane" (Luthea *et al.*, 1940; Waller, 1969; Lee-Lovick, 1978; Bailey, 1979). In India, problems with the reduction of sugarcane yield in commercial fields reported back to the 1930s, but probably smut was already there, colonizing susceptible genotypes of *S. barberi* and the wild parent of sugarcane *S. spontaneum* (Ferreira and Comstock, 1989; Croft and Braithwaite, 2006). In the 1940s, the disease reached Argentina and soon spread to Brazil, Paraguay, and Bolivia (Bergamin Filho *et al.*, 1987; Rago *et al.*, 2009). In Kenya, the first epidemic smut rose around the same time (Waller, 1969). Then other reports back to the 1970s described the observation of the disease in Hawaii and the Caribbean Islands (Byther *et al.*, 1971; Lee-Lovick, 1978). Sugarcane smut reached Western Australia in 1998 (Croft and Braithwaite, 2006; Magarey *et al.*, 2010). The origin center of sugarcane (Papua, New Guinea) was considered smut free until the recent 2016 report of the disease infecting commercial crop fields (Tom *et al.*, 2017). The sugarcane smut is, currently, present in almost every sugarcane-growing country, except for Fiji, an isolated island in Oceania (Croft and Braithwaite, 2006; Sundar *et al.*, 2012; Tom *et al.*, 2017).

The damage caused by the sugarcane smut varies widely, from negligible to severe losses, depending on the environmental conditions, the tolerance level of the sugarcane variety and the aggressiveness of the prevalence races of the pathogen (Whittle, 1982; Hoy, 1986; Sundar *et al.*, 2012). The decrease of yield is related to an increase of tillering, reduced diameters of the canes, reduced sucrose and increased fiber content (Martinez *et al.*, 2000; Wada *et al.*, 2016).

S. scitamineum is a biotrophic fungus from the *Basidiomycota* phylum, *Ustilaginales* order, *Ustilaginomycetes* class, and *Ustilaginaceae* family (Hawksworth *et al.*, 1995). Its life cycle comprises two distinct phases: a monokaryotic, with saprophytic growth of non-infective haploid cells (n) and an infective dikaryotic (n + n) phase, in which karyogamy results in teliospores (2n) formation (Figure 1 - A). The infective hyphae originate from the fusion of two mating compatible haploid cells (sporidial cells). Mating compatibility in smut fungi depends on the products of two loci, *a* and *b* (Bölker, 2001). The *a* locus comprises genes encoding a pheromone (mfa) and a membrane receptor (pra) (Figure 1 - B). The compatible mating-types

have alleles that encode complementary proteins of the pair receptor and pheromone. The *b* locus has genes that encode two subunits of a heterodimeric transcript factor (bE and bW) that must be encoded by different alleles (bE1/bW2 and bE2/bW1) in order to produce a functional protein (Figure 1 - C) (Gillissen *et al.*, 1992; Kronstad and Staben, 1997; Bölker, 2001; Singh *et al.*, 2004). Compatible cells recognize each other and fuse resulting in the dikaryotic hyphae able to penetrate and infect the host tissues (Singh *et al.*, 2004; Peters *et al.*, 2017).

The sexual compatibility system of *S. scitamineum* is bipolar, regarding its linked mating loci *a* and *b* on chromosome 2 (Taniguti *et al.*, 2015). Bipolar systems, in which two alternative and distinct sets of genes and alleles at the mating loci, determine two mating-types, is the prevailing system in most fungi (Coelho *et al.*, 2010). However, in the *Basidiomycota* phylum, *Ustilaginomycotina* subphylum, most species of plant pathogens (for example, *U. maydis*) have a tetrapolar mating system, in which the two mating loci are unlinked (Coelho *et al.*, 2010). Some studies suggest the derivation of bipolar mating systems from the tetrapolar ones (Bakkeren and Kronstad, 1994; Lee *et al.*, 1999; Lengeler *et al.*, 2002; Hsueh *et al.*, 2008).

After penetrating the host surface, the fungus initiates an intense proliferation colonizing the meristematic tissues of the plant. In later stages of the infection, fungal sporogenesis results in the formation of a whip-like structure (the characteristic symptom of the disease). In the process of sporogenesis after karyogamy, the hyphae undergo fragmentation, followed by pigments deposition and cell wall reinforcement resulting in the release of billions of diploid teliospores (which ones have soot aspect and are the source of disease name). The dissemination of the teliospores occurs by wind, rain, or even during the harvest time, infecting other plants (Banuett and Herskowitz, 1996; Taniguti *et al.*, 2015).

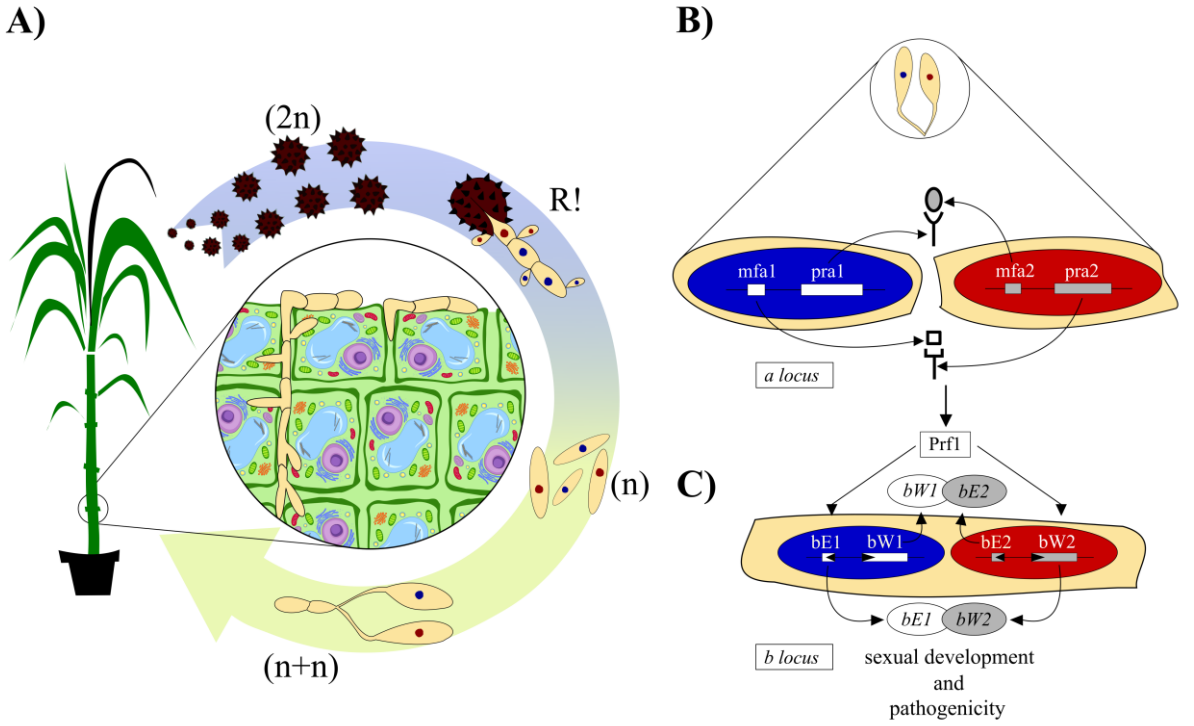


Figure 1 – A) Developmental stages of *S. scitamineum* life cycle: diploid teliospores (2n); germination and reductive meiosis (R!); resulting into haploid yeast-like sporidia (n); and anastomosis forming a dikaryotic infective hyphae (n+n), adapted from Taniguti *et al.* (2015). B) and C) schematic representation of the mating-type, sexual and pathogenic development: pheromone production (mfa) and receptor (pra) recognition encoded by the *a* locus leading to anastomosis B); formation of the heterodimeric functional proteins by the two complementary transcription factor subunits bE/bW, leading to the pathogenic development C), adapted from Bölker (2001).

3. Secretome, effectors, and plant host defense responses

The secretome is a fraction of the total proteome encoded by an organism secreted by a group of cells, as well as the machinery responsible for the secretion of these molecules (Tjalsma *et al.*, 2000). Agrawal *et al.* (2010) have revised this definition to include in the secretome only those proteins secreted to the extracellular medium.

The processes of pathogen recognition and defense response by the plant host immune system may have two phases. The first one involves the recognition of pathogen-associated molecular patterns (PAMPs) described as PTI (PAMP-triggered immunity) and the second associated with effectors recognition known as ETI (effector-triggered immunity) (Jones and Dangl, 2006). Receptors identified as PRRs (PAMP-recognition receptor) localized in the cellular membrane recognize PAMPs and trigger a signal transduction cascade activating the plant defense response (Hogenhout *et al.*, 2009). PTI is the most common plant defense mechanism, allowing fast and efficient responses to a large variety of pathogens (Roux *et al.*, 2014).

In contrast to the plant host defense responses, the adapted pathogens encode an arsenal of secreted proteins that can suppress the PTI, resulting in the susceptibility triggered by effectors (ETS – effector-triggered susceptibility) (Birch *et al.*, 2008). However, the plant hosts may have genes encoding resistance proteins (R genes) that recognize the pathogen effectors, in order to suppress the ETS response, resulting in resistance (Howden and Huitema, 2012). The recognition of the effectors by the products of R genes can be direct (gene-for-gene model) (Oßwald *et al.*, 2014) or indirect (guard model) (Jones and Dangl, 2006). The development of new effectors by the pathogen, in order to manipulate the plant host defense responses, followed by the development of new R genes by the plants, in order to combat these effectors is known as "arms race" in the plant-pathogen interaction (Coll *et al.*, 2011).

After the recognition of PAMPs or effectors, various signals may activate the plant immune system. Among them signaling based on changes of Ca^{+2} levels in the cytoplasm, rapidly production of reactive oxygen species (ROS), and signaling cascade via MAP-kinases (Wu *et al.*, 2013; Que *et al.*, 2014; Sánchez-Elordi *et al.*, 2015; Peters *et al.*, 2017). Regulatory hormones, such as salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) amplify the signs, resulting in the activation of transcription factors, defense genes, PR-proteins, phytoalexins, lignification of tissues, callose deposition and other cell wall reinforcement-related genes (Grant and Lamb, 2006).

In this context, understand the outcomes of critical components involved on the plant defense mechanisms, such as PAMPs, PRRs, effectors and R genes, can contribute to the development of new strategies for the disease management, avoiding significant yield and economic losses caused by diseases in the crops (Raffaele and Kamoun, 2012).

To study the function of the secreted effectors of smut fungi, usually, assays are conducted using the model for functional characterization of biotrophic fungi, *Ustilago maydis*. These effectors contribute to all steps of the disease development, since the early stages of the infection and penetration, and apoplastic, inter, and intracellular grown, until systemic dissemination. Therefore, there is a constant selection pressure from the host's immune system over those effectors, making them the fastest evolving genes in the pathogen genome (van der Linde *et al.*, 2012). As a consequence, the accumulation of mutations during the coevolution with the host immune system results in species or even race-specific effectors (Zuo *et al.*, 2019). Comparative genome studies revealed that the cereal smuts (*U. maydis*, *S. reilianum*, *S. scitamineum*, *U. hordei*, *U. trichophora*, and *U. tritici*) have similar numbers of genes encoding for secreted proteins. Whereas, *Melanopsichium pennsylvanicum*, the causal agent of the gall smut on the dicot host *Persicaria* spp., have a reduced set of effectors, potentially related to

jumping host from monocot to dicot plants (Sharma *et al.*, 2014; Benevenuto *et al.*, 2018; Schuster *et al.*, 2018).

Various effectors genes were already functionally characterized in *U. maydis*. Some of the effectors common to all smuts contribute to counteracting the innate immune responses conserved among different host species and facilitates the pathogen infection and colonization (Zuo *et al.*, 2019). The protein encoded by the Pep1 gene is required for *U. maydis* to suppress the oxidative burst induced by the host apoplastic peroxidases. The function of Pep1 is essential for epidermal penetration and conserved among smut fungi infecting either monocot or dicot plant hosts (Doehlemann *et al.*, 2009; Hemetsberger *et al.*, 2015).

Moreover, *U. maydis* secretes several protective effectors that work as barriers from host-derived antifungal proteins (Zuo *et al.*, 2019). Among these effectors is the Rsp3 protein, which binds and shields the fungal cell wall, protecting the fungal mycelium from the antifungal activity of the maize mannose-binding proteins AFP1 and AFP2 (Ma *et al.*, 2018). Another protective effector is the UmFly1, which is a secreted fungal lysin metalloprotease and cleaves maize chitinase-A, reducing its catalytic activity (Ökmen *et al.*, 2018). UmFly1 is also required by *U. maydis* to activate endogenous chitinases by N-terminal processing necessary for the separation of cells in the yeast-like phase (Ökmen *et al.*, 2018). The opposite effects of UmFly1 on endogenous chitinases and host chitinases suggest a neofunctionalization of this effector toward virulence during evolutionary adaptation of *U. maydis* to maize (Zuo *et al.*, 2019). Pit2 is a secreted effector that inhibits maize papain-like cysteine proteases, which is crucial for the suppression of apoplastic host defense responses and successful infection and virulence of *U. maydis* (Doehlemann *et al.*, 2011; Mueller *et al.*, 2013).

Other effectors function reprogramming plant metabolic pathways when translocated into the host cell (Djamei *et al.*, 2011; Tanaka *et al.*, 2014; Redkar *et al.*, 2015a). An effector named Cmu1 prevents the production of the SA (the primary defense hormone) by sequestering its biosynthesis precursor chorismate (Djamei *et al.*, 2011; Djamei and Kahmann, 2012). The effector Tin2 redirects the lignin production pathway to anthocyanin production, through stabilization of the maize protein kinase ZmTTK1 (Tanaka *et al.*, 2014). It is known that the Tin2 effectors from *U. maydis* and *S. reilianum* target different kinase proteins paralogs in maize (Tanaka *et al.*, 2019). See1, another effector acting inside the host cell, prevents phosphorylation of the maize SGT1 protein involved in resistance and cell cycle regulation (Redkar *et al.*, 2015a). The sequence of See1 is conserved among the smut pathogens, however, the *U. hordei* See1 ortholog does not complement the *U. maydis* See1 knockout mutants

(Redkar *et al.*, 2015b), suggesting different host adaptation strategies of See1 orthologs in smut pathogens.

SAD1, functionally characterized in *S. reilianum*, is a species-specific effector causing suppression of apical dominance in maize plants and increase branching of female inflorescences and the production of *S. reilianum* teliospores (Ghareeb *et al.*, 2015). Effectors of *S. reilianum* contribute quantitatively to the virulence of the pathogen, as demonstrated recently by Ghareeb *et al.* (2019) in a study with deletion of a single gene and a gene cluster. Therefore, deleting a single gene may result in a slight reduction of virulence or no detectable phenotype (Ghareeb *et al.*, 2019). This fact complicates the studies for the functional characterization of a single effector gene.

4. Sugarcane-smut molecular interaction

Until now, there is no fungicide registered to control the sugarcane smut, only to prevent it in the form of bud treatment (Bayfidan EC – Triadimenol 250 g/L). The most efficient method to prevent the disease is the utilization of healthy seedlings free of pathogens and resistant varieties. However, the pathogen colonizes even resistant genotypes, occasionally producing the whip-like structure, but causing no economic losses in the field scale.

One of the best-characterized responses of sugarcane toward *S. scitamineum* is related to the oxidative burst. Genes encoding components of the reactive oxygen species (ROS) metabolism activated earlier after inoculation and have a different behavior according to the sugarcane genotypes (Menossi *et al.*, 2008; You-Xiong *et al.*, 2011; Schaker *et al.*, 2016; Peters *et al.*, 2017). Peters *et al.* (2017) demonstrated the accumulation of H_2O_2, and a reduction in antioxidant enzymes activities earlier in resistant sugarcane plants when compared to a susceptible genotype. The oxidative burst coincided with some phases of the fungal development: germination, formation of the appressorium, and colonization of the host tissues. Potentially, that would be the first layer in the sugarcane defense mechanism when the recognition of the pathogen occurs, leading to the activation of mechanisms related to the resistance.

Sugarcane plants resistant to the smut, accumulate glycoproteins during the infection (Fontaniella *et al.*, 2002) associated with the inhibition of the teliospores germination (Lloyd and Pillay, 1980; Lloyd and Naidoo, 1983; Martinez *et al.*, 2000; Blanch *et al.*, 2007). Detecting high levels of the polyamines conjugated with phenolic compounds in mature organs indicates the effect of the presence of the fungus (Legaz *et al.*, 1998; Piñon *et al.*, 1999). The phenolic

compounds produced in response to the infection inactivates the effect of polyamines through the conjugation inhibiting the germination of the fungus (Sundar *et al.*, 2012). The polyamines are essential for fungal growth and differentiation (Shapira *et al.*, 1989; Ruiz-Herrer, 1994).

The germination of the spores occurs on the base of new leaves, following hyphal growth and appressorium formation. In response to the infection, sugarcane produces β-1,3-glucanase and chitinase to attack components of the fungus cell wall (Blanch *et al.*, 2007). Also, the pathogen induces synthesis and activation of enzymes involved in the production and polymerization of monolignols involved in lignin synthesis, alters the expression pattern of transcription regulators, and that of genes related to ethylene and auxin pathways (Blanch *et al.*, 2007; LaO *et al.*, 2008; Menossi *et al.*, 2008; Schaker *et al.*, 2016; Peters *et al.*, 2017).

Studies involving the global responses of sugarcane to smut infection performed by different groups revealed a more comprehensive view of this particular plant-pathogen interaction. The authors used approaches such as transcriptomics (Thokoane and Rutherford, 2001; Borrás-Hidalgo *et al.*, 2005; LaO *et al.*, 2008; Wu *et al.*, 2013; Que *et al.*, 2014; Huang *et al.*, 2015; Taniguti *et al.*, 2015; Schaker *et al.*, 2016); proteomics (Que *et al.*, 2011; Barnabas *et al.*, 2016); and metabolomics (Schaker *et al.*, 2017). The first defense mechanism activated is based on PAMPs recognition by receptor proteins of the host. The resistant sugarcane variety N52/219 activates the expression of an RLK (S-receptor-like-kinase) protein involved in the perception of chitin, leading to the increase in expression levels of chitinases (Thokoane and Rutherford, 2001; Esh *et al.*, 2014; Que *et al.*, 2014). Other proteins associated with pathogenicity such as thaumatin, proteins with antifungal activity, are characteristic of plants resistant to smut (Heinze *et al.*, 2001; Que *et al.*, 2014; Su *et al.*, 2016). Resistant varieties also activate MAPK signaling pathways (Mitogen-Activated Protein Kinase), associated with defense mechanisms (Wu *et al.*, 2013; Que *et al.*, 2014; Sánchez-Elordi *et al.*, 2015). In susceptible genotypes, were observed an increase in the expression of an RGA (Resistant Gene Analog), homologous to the BAM1 (RGA482) (Schaker *et al.*, 2016). BAM1 is associated with meristematic identity suggesting an earlier association with whip development and meristematic functions. Genes encoding cytoplasmatic receptors NBS-LRR-like proteins (Nucleotide-Binding Site – Leucine-Rich Repeat) identified in both resistant (Borrás-Hidalgo *et al.*, 2005; LaO *et al.*, 2008; Que *et al.*, 2011) and susceptible varieties differentially expressed (Schaker *et al.*, 2016; Rody *et al.*, 2019).

Although various studies addressing sugarcane mechanisms in response to *S. scitamineum* infection are available, the mechanisms of fungal attack leading to disease have received little attention. Taniguti *et al.* (2015), using transcriptomic data, compared plants five

days after inoculation (dai) with ones after whip emission (200 dai). The authors identified approximately 13% of the *S. scitamineum* genes differentially expressed in the plant compared to *in vitro* growth. The predicted function of these genes suggested that the pathogen can overcome the plant defense through the detox of defense molecules (protease inhibitors, pisatin demethylase, benzoate 4-monooxygenase, chorismate mutase, superoxide dismutase, catalase) and the growing and colonization of the plant tissues using an arsenal of proteins related to nutrient absorption and cell wall degrading enzymes. Secreted proteins are of particular interest in understanding the molecular communication between the pathogen and its plant host. Using computational tools to predict the presence of signal peptide and absence of transmembrane domains and glycosylphosphatidylinositol (GPI) anchors, the authors proposed that *S. scitamineum* secretome comprise 305 proteins, including 54 carbohydrates active enzymes (CAZymes) and 70 effectors candidates (based on the predicted size of the proteins, and its cysteine content) (Taniguti *et al.*, 2015).

5. Dual transcriptomics in plant-fungal interaction: difficulties in the data collection and analysis

Even though the RNA-Seq technique provides more accurate detection of transcripts, including low-abundant ones, compared to other transcriptomics approaches, such as the microarrays chips, it still has a high cost when sequencing with high depth coverage (Kohler and Tisserant, 2014). Sequencing coverage reflects in the quantity and quality of the reads obtained, and consequently, the precision of detection. When conducting a dual transcriptomics assay aiming at the discovery of the plant and the pathogen components during the interaction, often the genome size and, consequently, the transcriptome size of the plant is significantly larger than that of the pathogen (Kohler and Tisserant, 2014). As plant growth is more expressive in the early stages of infection, the pathogen represents only a small percentage of the total RNA pool.

Unlike genome sequencing, predicting the number of sequences required to detect all transcripts is quite tricky, considering that the RNA concentration varies significantly between tissues and conditions (Wang *et al.*, 2009). For the *Escherichia coli* (4.6 – 5.3 Mb), Haas *et al.* (2012) estimated that sequencing two million reads cover about 80% of the transcriptome. Whereas, for covering 80% of the yeast (*Saccharomyces cerevisiae*, 12.1 Mb) transcriptome, it would be necessary to sequence four million reads (Nagalakshmi *et al.*, 2008; Sims *et al.*, 2014). On the plant side, for example, the model *Arabidopsis thaliana* (135 Mb), it would require a

sequence depth of 50 million reads to nearly saturate the coverage of expressed genes (Van Verk *et al.*, 2013).

Another crucial step into the dual transcriptome analysis is to correctly map the mixed sequenced reads to the respective reference genomes (if available). Generally, the genomes of plants and fungal are different enough to simply align the reads to the reference genomes without prefiltering (Kohler and Tisserant, 2014). When only one genome is available, usually the fungal, the separation can still be done, and the reads of the plant can be mapped to a set of transcripts, such as ESTs collections, or generated with high throughput sequencers.

For the mapping step, various methods and softwares are available, some "unspliced" that do not allow alignments within significant gaps, and consequently do not support mapping spliced reads to the reference genome. The most common of these mappers are the ones based on the Burrows-Wheeler transform method, because of its more computational and time efficiencies, such as BWA (Li and Durbin, 2010) and Bowtie2 (Langmead and Salzberg, 2012). To mapping spliced reads to a reference genome, the most popular softwares are TopHat2 (Kim *et al.*, 2013) and Hisat2 (Kim *et al.*, 2015). First, one uses an unspliced method to align the reads to a reference, and then, the split of unmapped reads into shorter segments allow an appropriated mapping. Subsequently, one searches for the spliced sites in the neighboring genomic regions.

As some genomes can contain repetitive regions and large multigene families, reads can align to multiple locations (Kohler and Tisserant, 2014). Thus, the suggestion is to remove these alignments or treat them in a random way to subsequent analysis. For this purpose, the recommended approach is to allocate these multi mapped reads in proportion to the number of uniquely mapped ones (Mortazavi *et al.*, 2008).

After the mapping step, it is necessary to attribute the mapped reads to a gene or transcript, using an annotation reference or *de novo* approach to generate a count table. Some softwares are available for that, such as the FeatureCounts implemented in the Subread package (Liao *et al.*, 2013, 2014), HTSeq (Anders *et al.*, 2015), the R package Limma (Ritchie *et al.*, 2015), RSEM (Li and Dewey, 2011) and Cufflinks & Cuffmerge (Roberts *et al.*, 2011a, 2011b).

Another problem when dealing with dual transcriptome studies is to normalize the data correctly. Various robust methods are available to deal with transcriptomics data normalization in order to allow compare different treatments and assays. One of the most popular normalization schemes for expression data is the RPKM (Reads Per Kilobase reads Mapped) (Mortazavi *et al.*, 2008) used for single-end reads, and FPKM (Fragments) for paired-end sequencing. Another type of normalization used when exploring expression data is DESeq2

implemented in the R package (Love *et al.*, 2014) and EdgeR (Trimmed mean of M values - TMM) (Robinson *et al.*, 2010; McCarthy *et al.*, 2012). These methods use a scaling factor estimated from the mean of the median of reads counts ratio across the samples, assuming that the majority of the expressed genes are not differentially expressed (Kohler and Tisserant, 2014).

Studies using the dual-transcriptomics RNA-Seq approach with plant-fungus interaction are still scarce. Schaker *et al.* (2016) used a dual-transcriptomics RNA-Seq of sugarcane plants inoculated with *S. scitamineum* at 5 days after inoculation (dai) and 200 dai (whip emission). In that study, the mapping results of aligned reads to the fungal reference CDS region reported only 2% mapped reads at 5 dai, and an increase up to 18% at 200 dai. In another study with the pathosystem rice-blast fungus (*Oryza sativa* vs *Magnaporthe oryzae*), Kawahara *et al.* (2012) reported that at 24 hai, only 0.2-0.3% of the mixed reads (plant + fungus mRNA) mapped on the fungal reference genome. These results come from the early stage of infection, in which the plant growth and development is faster compared with the fungal colonization, which difficult the detection of fungal mRNA unless considering a deep sequencing.

Other studies focus on the beneficial interactions between fungus and plants, exploring mycorrhizal interactions. Kohler and Tisserant (2014) produced a review exploring dual transcriptomics related to these interactions. In the review, the authors emphasize the importance of the experimental design in order to obtain sufficient depth sequencing to detect both organisms which interact. A study considering dual-transcriptomic issues and how to address the problems was conducted by Naidoo *et al.* (2018), the authors highlight how to analyze the data produced by a dual-transcriptomic assay, exploring the methods developed for such task and future perspectives with the lower cost brought by the Next Generation Sequence (NGS).

The lacking of funding support, or need to outscoring sequencing services in many cases, limits the number of reads that can be obtained per sequencing, implicating in the reduced accuracy of transcript detection. However, even in those cases, with a reduced number of reads, the analysis can be cared out, especially if a high-quality reference genome is available (Naidoo *et al.*, 2018), with some precautions and later validation of differential expressed genes using more accurate techniques such as RT-qPCR. Therefore, it can still guide functional characterization studies for a more comprehensive understanding of the plant-pathogen interaction.

6. Hypothesis

The use of dual transcriptome data analysis can reveal candidate genes associated with *S. scitamineum* expression profile differentiation when infecting smut-resistant vs. -susceptible plants.

7. Objectives

7.1 Sequence and analyze data of a dual transcriptome RNAseq data obtained from a previous experiment (Peters, 2016);

7.2 Select all reads of *S. scitamineum* from the total data set based on the reference genome (Taniguti *et al.*, 2015);

7.3 Perform statistical analysis of differentially expressed genes;

7.4 Use annotation strategies to compile data of homologs functional predictions and that of experimental approaches;

7.5 Differentiate expression profile of *S. scitamineum* genes infecting resistant and susceptible plants.

REFERENCES

Agrawal, G. K., Jwa, N.-S., Lebrun, M.-H., Job, D., and Rakwal, R. (2010). Plant secretome: Unlocking secrets of the secreted proteins. *Proteomics* 10, 799–827. doi:10.1002/pmic.200900514.

Amalraj, V. A., and Balasundaram, N. (2006). On the Taxonomy of the Members of 'Saccharum Complex.' *Genet. Resour. Crop Evol.* 53, 35–41. doi:10.1007/s10722-004-0581-1.

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi:10.1093/bioinformatics/btu638.

Bailey, R. A. (1979). Possibilities for the control of sugarcane smut (Ustilago scitaminea) with fungicides. *Proc. South African Sugar Technol.*, 137–142.

Bakkeren, G., and Kronstad, J. W. (1994). Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7085–7089. doi:10.1073/pnas.91.15.7085.

Banuett, F., and Herskowitz, I. (1996). Discrete developmental stages during teliospore formation in the corn smut fungus, Ustilago maydis. *Development* 122, 2965–2976.

Barnabas, L., Ashwin, N. M. R., Kaverinathan, K., Trentin, A. R., Pivato, M., Sundar, A. R., *et al.* (2016). Proteomic analysis of a compatible interaction between sugarcane and *Sporisorium scitamineum*. *Proteomics* 16, 1111–1122. doi:10.1002/pmic.201500245.

Begerow, D., Schäfer, A. M., Kellner, R., Yurkov, A., Kemler, M., Oberwinkler, F., *et al.* (2014). "Ustilaginomycotina," in *Systematics and Evolution* (Springer Berlin Heidelberg), 295–329. doi:10.1007/978-3-642-55318-9_11.

Benevenuto, J., Teixeira-Silva, N. S., Kuramae, E. E., Croll, D., and Monteiro-Vitorello, C. B. (2018). Comparative Genomics of Smut Pathogens: Insights From Orphans and Positively Selected Genes Into Host Specialization. *Front. Microbiol.* 9, 1–17. doi:10.3389/fmicb.2018.00660.

Bergamin Filho, A., Amorim, L., Cardoso, C. O. N., Sanguino, A., Irvine, J. E., and Silva, W. M. (1987). "Carvão da cana-de-açúcar e sua epidemiologia," in *Boletim Tecnico Coopersucar-Edicao Especial*, 1–23.

Birch, P. R., Boevink, P. C., Gilroy, E. M., Hein, I., Pritchard, L., and Whisson, S. C. (2008). Oomycete RXLR effectors: delivery, functional redundancy and durable disease resistance. *Curr. Opin. Plant Biol.* 11, 373–379. doi:10.1016/J.PBI.2008.04.005.

Blanch, M., Legaz, M. E., Millanes, A. M., and Vicente, C. (2007). Glycoproteins of sugarcane plants facilitate the infectivity of Ustilago scitaminea and Xanthomonas albilineans, two sugarcane pathogens. *Commun. Curr. Res. Educ. Top. Trends Appl. Microbiol.*, 163–169.

Bölker, M. (2001). Ustilago maydis - A valuable model system for the study of fungal dimorphism and virulence. *Microbiology* 147, 1395–1401. doi:10.1099/00221287-147-6-1395.

Borrás-Hidalgo, O., Thomma, B. P. H. J., Carmona, E., Borroto, C. J., Pujol, M., Arencibia, A., *et al.* (2005). Identification of sugarcane genes induced in disease-resistant somaclones upon inoculation with Ustilago scitaminea or Bipolaris sacchari. *Plant Physiol. Biochem.* 43, 1115–1121. doi:10.1016/J.PLAPHY.2005.07.011.

Byther, R. S., Steiner, G. W., and Wismer, C. A. (1971). New Sugarcane diseases reported in Hawaii. *Sugarcane Pathol. Newsl.* 7, 18–21.

Cardoso-Silva, C. B., Costa, E. A., Mancini, M. C., Balsalobre, T. W. A., Canesin, L. E. C., Pinto, L. R., *et al.* (2014). De Novo Assembly and Transcriptome Analysis of Contrasting Sugarcane Varieties. *PLoS One* 9, e88462. doi:10.1371/journal.pone.0088462.

CONAB, Companhia Nacional de Abastecimento. Acompanhamento da safra Brasileira. Cana-de-açúcar – Quarto levantamento. Publicação Mensal, v.5, n.4, Safra 2018/19, p.1-75 Abril, 2019.

Coelho, M. A., Sampaio, J. P., and Gonçalves, P. (2010). A deviation from the bipolar-tetrapolar mating paradigm in an early diverged basidiomycete. *PLoS Genet.* 6, 19–21. doi:10.1371/journal.pgen.1001052.

Coll, N. S., Epple, P., and Dangl, J. L. (2011). Programmed cell death in the plant immune system. *Cell Death Differ.* 18, 1247–1256. doi:10.1038/cdd.2011.37.

Croft, B. J., and Braithwaite, K. S. (2006). Management of an incursion of sugarcane smut in Australia. *Australas. Plant Pathol.* 35, 113–122. doi:10.1071/AP05104.

D'Hont, A. (2005). Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet. Genome Res.* 109, 27–33. doi:10.1159/000082378.

D'Hont, A., and Glaszman, J. C. (2001). SUGARCANE GENOME ANALYSIS WITH MOLECULAR MARKERS : A FIRST DECADE OF RESEARCH. *Proc. Int. Soc. Sugar Cane Technol* 24, 556–559.

D'Hont, A., Grivet, L., Feldmann, P., Glaszmann, J. C., Rao, S., and Berding, N. (1996). Characterisation of the double genome structure of modern sugarcane cultivars (Saccharum spp.) by molecular cytogenetics. *MGG Mol. Gen. Genet.* 250, 405–413. doi:10.1007/BF02174028.

de Setta, N., Monteiro-Vitorello, C., Metcalfe, C., Cruz, G. M., Del Bem, L., Vicentini, R., *et al.* (2014). Building the sugarcane genome for biotechnology and identifying evolutionary trends. *BMC Genomics* 15, 540–557. doi:10.1186/1471-2164-15-540.

Djamei, A., and Kahmann, R. (2012). Ustilago maydis: Dissecting the Molecular Interface between Pathogen and Plant. *PLoS Pathog.* 8, e1002955. doi:10.1371/journal.ppat.1002955.

Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., *et al.* (2011). Metabolic priming by a secreted fungal effector. *Nature* 478, 395–398. doi:10.1038/nature10454.

Doehlemann, G., Reissmann, S., Aßmann, D., Fleckenstein, M., and Kahmann, R. (2011). Two linked genes encoding a secreted effector and a membrane protein are essential for Ustilago maydis-induced tumour formation. *Mol. Microbiol.* 81, 751–766. doi:10.1111/j.1365-2958.2011.07728.x.

Doehlemann, G., Van Der Linde, K., Aßmann, D., Schwammbach, D., Hof, A., Mohanty, A., *et al.* (2009). Pep1, a secreted effector protein of Ustilago maydis, is required for successful invasion of plant cells. *PLoS Pathog.* 5, e1000290. doi:10.1371/journal.ppat.1000290.

Esh, A. M. H., Guirgis, A. A., El-kholi, M. M. A., El-Absawy, E. A., Nasr, M. I., and Hassanien, E. H. (2014). The Activity of Pathogenesis Related Proteins in Smut Resistant and Susceptible Sugarcane (GT54-9) Mutants Induced by Gamma Radiation. *Adv. Plants Agric. Res.* 1, 1–12. doi:10.15406/apar.2014.01.00024.

FAO. FAOSTAT. Available at: <http://www.fao.org/faostat/en/#data/QC>. Access date: 20/04/2020.

Ferreira, S. A., and Comstock, J. C. (1989). "Smut," in *Diseases of sugarcane* (Elsevier), 211–229.

Fontaniella, B., Márquez, A., Rodríguez, C. W., Piñón, D., Solas, M. T., Vicente, C., *et al.* (2002). A role for sugarcane glycoproteins in the resistance of sugarcane to Ustilago scitaminea. *Plant Physiol. Biochem.* 40, 881–889. doi:10.1016/S0981-9428(02)01443-2.

Garsmeur, O., Droc, G., Antonise, R., Grimwood, J., Potier, B., Aitken, K., *et al.* (2018). A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nat. Commun.* 9, 2638–2648. doi:10.1038/s41467-018-05051-5.

Ghareeb, H., Drechsler, F., Löfke, C., Teichmann, T., and Schirawski, J. (2015). SUPPRESSOR OF APICAL DOMINANCE1 of Sporisorium reilianum modulates inflorescence branching architecture in maize and arabidopsis. *Plant Physiol.* 169, 2789–2804. doi:10.1104/pp.15.01347.

Ghareeb, H., Zhao, Y., and Schirawski, J. (2019). Sporisorium reilianum possesses a pool of effector proteins that modulate virulence on maize. *Mol. Plant Pathol.* 20, 124–136. doi:10.1111/mpp.12744.

Gillissen, B., Bergemann, J., Sandmann, C., Schroeer, B., Bölker, M., and Kahmann, R. (1992). A two-component regulatory system for self/non-self recognition in Ustilago maydis. *Cell* 68, 647–657. doi:10.1016/0092-8674(92)90141-X.

Grant, M., and Lamb, C. (2006). Systemic immunity. *Curr. Opin. Plant Biol.* 9, 414–420. doi:10.1016/J.PBI.2006.05.013.

Haas, B. J., Chin, M., Nusbaum, C., Birren, B. W., and Livny, J. (2012). How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* 13, 734. doi:10.1186/1471-2164-13-734.

Hawksworth, D. L., Kirk, P. M., Sutton, B. C., and Pegler, D. N. (1995). Ainsworth & Bisby's Dictionary of the Fungi. *Edn. Wallingford CAB Int.*, 616.

Heinze, B. S., Thokoane, L. N., Williams, N. J., Barnes, J. M., and Rutherford, R. S. (2001). The smut sugarcane interaction as a model system for the integration of marker discovery and gene isolation. *Proc S Afr Sug Technol Ass*, 88–93.

Hemetsberger, C., Mueller, A. N., Matei, A., Herrberger, C., Hensel, G., Kumlehn, J., *et al.* (2015). The fungal core effector Pep1 is conserved across smuts of dicots and monocots. *New Phytol.* 206, 1116–1126. doi:10.1111/nph.13304.

Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R., and Kamoun, S. (2009). Emerging Concepts in Effector Biology of Plant-Associated Organisms. *Mol. Plant-Microbe Interact.* 22, 115–122. doi:10.1094/MPMI-22-2-0115.

Howden, A. J. M., and Huitema, E. (2012). Effector-triggered post-translational modifications and their role in suppression of plant immunity. *Front. Plant Sci.* 3, 160. doi:10.3389/fpls.2012.00160.

Hoy, J. W. (1986). Incidence of Sugarcane Smut in Louisiana and Its Effect on Yield. *Plant Dis.* 70, 59–60. doi:10.1094/pd-70-59.

Hsueh, Y. P., Fraser, J. A., and Heitman, J. (2008). Transitions in sexuality: Recapitulation of an ancestral tri- and tetrapolar mating system in Cryptococcus neoformans. *Eukaryot. Cell* 7, 1847–1855. doi:10.1128/EC.00271-08.

Huang, X., Chen, M.-H., Yang, L.-T., Li, Y.-R., and Wu, J.-M. (2015). Effects of Exogenous Abscisic Acid on Cell Membrane and Endogenous Hormone Contents in Leaves of Sugarcane Seedlings under Cold Stress. *Sugar Tech* 17, 59–64. doi:10.1007/s12355-014-0343-0.

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi:10.1038/nature05286.

Kawahara, Y., Oono, Y., Kanamori, H., Matsumoto, T., Itoh, T., and Minami, E. (2012). Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS One* 7, e49423. doi:10.1371/journal.pone.0049423.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14, R36. doi:10.1186/gb-2013-14-4-r36.

Kohler, A., and Tisserant, E. (2014). “Exploring the transcriptome of mycorrhizal interactions,” in *Advances in Botanical Research* (Academic Press Inc.), 53–78. doi:10.1016/B978-0-12-397940-7.00002-1.

Kronstad, J. W., and Staben, C. (1997). MATING TYPE IN FILAMENTOUS FUNGI. *Annu. Rev. Genet.* 31, 245–276. doi:10.1146/annurev.genet.31.1.245.

Lam, E., Shine, J., Da Silva, J., Lawton, M., Bonos, S., Calvino, M., *et al.* (2009). Improving sugarcane for biofuel: engineering for an even better feedstock. *GCB Bioenergy* 1, 251–255. doi:10.1111/j.1757-1707.2009.01016.x.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923.

LaO, M., Arencibia, A. D., Carmona, E. R., Acevedo, R., Rodríguez, E., León, O., *et al.* (2008). Differential expression analysis by cDNA-AFLP of Saccharum spp. after inoculation with the host pathogen Sporisorium scitamineum. *Plant Cell Rep.* 27, 1103–1111. doi:10.1007/s00299-008-0524-y.

Lee-Lovick, G. (1978). Smut of sugarcane - Ustilago scitaminea. *Rev. Plant Pathol.* 57, 181–188.

Lee, N., Bakkeren, G., Wong, K., Sherwood, J. E., and Kronstad, J. W. (1999). The mating-type and pathogenicity locus of the fungus Ustilago hordei spans a 500-kb region. *Proc. Natl. Acad. Sci. U. S. A.* 96, 15026–15031. doi:10.1073/pnas.96.26.15026.

Legaz, M. E., Armas, R. d., Pinon, D., and Vicente, C. (1998). Relationships between phenolics-conjugated polyamines and sensitivity of sugarcane to smut (Ustilago scitaminea). *J. Exp. Bot.* 49, 1723–1728. doi:10.1093/jxb/49.327.1723.

Lengeler, K. B., Fox, D. S., Fraser, J. A., Allen, A., Forrester, K., Dietrich, F. S., *et al.* (2002). Mating-type locus of Cryptococcus neoformans: A step in the evolution of sex chromosomes. *Eukaryot. Cell* 1, 704–718. doi:10.1128/EC.1.5.704-718.2002.

Li, B., and Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323. doi:10.1186/1471-2105-12-323.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi:10.1093/bioinformatics/btp698.

Liao, Y., Smyth, G. K., and Shi, W. (2013). The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41. doi:10.1093/nar/gkt214.

Liao, Y., Smyth, G. K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/bioinformatics/btt656.

Lloyd, H. L., and Naidoo, G. (1983). Chemical assay potentially suitable for determination of smut resistance of sugarcane cultivars. *Plant Dis.* 67, 1103–1105.

Lloyd, H. L., and Pillay, M. (1980). The Development of an improved mehtod for evaluating sugarcane for resistance to smut. *Proc. South African Sugar Technol. Assoc.* 54, 168–172.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8.

Luthea, J. C., Sattar, A., and Sandhu, S. S. (1940). Experiments on the control of smut of Sugarcane (Ustilago scitaminea Syd.). *Proc. Natl. Acad. Sci. India* 12, 118–128.

Ma, L.-S., Wang, L., Trippel, C., Mendoza-Mendoza, A., Ullmann, S., Moretti, M., *et al.* (2018). The Ustilago maydis repetitive effector Rsp3 blocks the antifungal activity of mannose-binding maize proteins. *Nat. Commun.* 9, 1–15. doi:10.1038/s41467-018-04149-0.

Magarey, R. C., Denney, D., Sheahan, T., and Bull, J. I. (2010). The Australian sugarcane smut epidemic: epidemiological considerations and predictions for the final stages. in *Proceedings of the Australian Society of Sugar Cane TechnologistsTechnologists*, ed. R. C. Bruce (Mackay, Australia: Australian Society of Sugar Cane Technologists), 375–387.

Martinez, M., Medina, I., Naranjo, S., Rodríguez, C. W., Armas, R. de, Pinon, D., *et al.* (2000). Changes of some chemical parameters, involved in sucrose recovery from sugarcane juices, related to the susceptibility or resistance of sugarcane plants to smut (Ustilago scitaminea). *Int. sugar J.* 102, 445–448.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi:10.1093/nar/gks042.

Menossi, M., Silva-Filho, M. C., Vincentz, M., Van-Sluys, M.-A., and Souza, G. M. (2008). Sugarcane functional genomics: gene discovery for agronomic trait development. *Int. J. Plant Genomics*, 1–11. doi:10.1155/2008/458732.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi:10.1038/nmeth.1226.

Mueller, A. N., Ziemann, S., Treitschke, S., Aßmann, D., and Doehlemann, G. (2013). Compatibility in the Ustilago maydis-Maize Interaction Requires Inhibition of Host Cysteine Proteases by the Fungal Effector Pit2. *PLoS Pathog.* 9, e1003177. doi:10.1371/journal.ppat.1003177.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., *et al.* (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science (80).* 320, 1344–1349. doi:10.1126/science.1158441.

Naidoo, S., Visser, E. A., Zwart, L., Toit, Y. du, Bhadauria, V., and Shuey, L. S. (2018). Dual RNA-Sequencing to Elucidate the Plant-Pathogen Duel. *Curr. Issues Mol. Biol.* 27, 127–142. doi:10.21775/cimb.027.127.

Ökmen, B., Kemmerich, B., Hilbig, D., Wemhöner, R., Aschenbroich, J., Perrar, A., *et al.* (2018). Dual function of a secreted fungalysin metalloprotease in Ustilago maydis. *New Phytol.* 220, 249–261. doi:10.1111/nph.15265.

Oßwald, W., Fleischmann, F., Rigling, D., Coelho, A. C., Cravador, A., Diez, J., *et al.* (2014). Strategies of attack and defence in woody plant- *Phytophthora* interactions. *For. Pathol.* 44, 169–190. doi:10.1111/efp.12096.

Peters, L. P. (2016). A more detailed view of reactive oxygen species metabolism in the sugarcane and Sporisorium scitamineum interaction. Available at: https://www.teses.usp.br/teses/disponiveis/11/11137/tde-23012017-094942/en.php

Peters, L. P., Carvalho, G., Vilhena, M. B., Creste, S., Azevedo, R. A., and Monteiro-Vitorello, C. B. (2017). Functional analysis of oxidative burst in sugarcane smut-resistant and -susceptible genotypes. *Planta* 245, 749–764. doi:10.1007/s00425-016-2642-z.

Piñon, D., de Armas, R., Vicente, C., and Legaz, M. E. (1999). Role of polyamines in the infection of sugarcane buds by Ustilago scitaminea spores. *Plant Physiol. Biochem.* 37, 57–64. doi:10.1016/S0981-9428(99)80067-9.

Que, Y., Su, Y., Guo, J., Wu, Q., and Xu, L. (2014). A Global View of Transcriptome Dynamics during Sporisorium scitamineum Challenge in Sugarcane by RNA-seq. *PLoS One* 9, e106476. doi:10.1371/journal.pone.0106476.

Que, Y., Xu, L., Lin, J., Ruan, M., Zhang, M., and Chen, R. (2011). Differential Protein Expression in Sugarcane during Sugarcane-Sporisorium scitamineum Interaction Revealed by 2-DE and MALDI-TOF-TOF/MS. *Comp. Funct. Genomics*, 1–10. doi:10.1155/2011/989016.

Raffaele, S., and Kamoun, S. (2012). Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* 10, 417–430. doi:10.1038/nrmicro2790.

Rago, A. M., Casagrande, M. V., and Massola Júnior, N. S. (2009). Variabilidade patogênica de ustilago scitaminea no Estado de São Paulo. *Summa Phytopathol.* 35, 93–97. doi:10.1590/s0100-54052009000200002.

Redkar, A., Hoser, R., Schilling, L., Zechmann, B., Krzymowska, M., Walbot, V., *et al.* (2015a). A Secreted Effector Protein of Ustilago maydis Guides Maize Leaf Cells to Form Tumors . *Plant Cell* 27, 1332–1351. doi:10.1105/tpc.114.131086.

Redkar, A., Villajuana- Bonequi, M., and Doehlemann, G. (2015b). Conservation of the *Ustilago maydis* effector See1 in related smuts. *Plant Signal. Behav.* 10, e1086855. doi:10.1080/15592324.2015.1086855.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., *et al.* (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007.

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011a). Identification of novel transcripts in annotated genomes using RNA-seq. *Bioinformatics* 27, 2325–2329. doi:10.1093/bioinformatics/btr355.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L., and Pachter, L. (2011b). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.* 12, R22. doi:10.1186/gb-2011-12-3-r22.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.

Rody, H. V. S., Bombardelli, R. G. H., Creste, S., Camargo, L. E. A., Van Sluys, M.-A., and Monteiro-Vitorello, C. B. (2019). Genome survey of resistance gene analogs in sugarcane: genomic features and differential expression of the innate immune system from a smut-resistant genotype. *BMC Genomics* 20, 809. doi:10.1186/s12864-019-6207-y.

Roux, F., Voisin, D., Badet, T., Balagué, C., Barlet, X., Huard-Chauveau, C., *et al.* (2014). Resistance to phytopathogens *e tutti quanti*: placing plant quantitative disease resistance on the map. *Mol. Plant Pathol.* 15, 427–432. doi:10.1111/mpp.12138.

Ruiz-Herrer, A. (1994). Polyamines, DNA methylation, and fungal differentiation. *Crit. Rev. Microbiol.* 20, 143–150.

Sánchez-Elordi, E., Ríos, L. M. de los, Vicente, C., and Legaz, M.-E. (2015). Sugar cane arginase competes with the same fungal enzyme as a false quorum signal against smut teliospores. *Phytochem. Lett.* 14, 115–122. doi:10.1016/J.PHYTOL.2015.09.013.

Schaker, P. D. C., Palhares, A. C., Taniguti, L. M., Peters, L. P., Creste, S., Aitken, K. S., *et al.* (2016). RNAseq Transcriptional Profiling following Whip Development in Sugarcane Smut Disease. *PLoS One* 11, e0162237. doi:10.1371/journal.pone.0162237.

Schaker, P. D. C., Peters, L. P., Cataldi, T. R., Labate, C. A., Caldana, C., and Monteiro-Vitorello, C. B. (2017). Metabolome Dynamics of Smutted Sugarcane Reveals Mechanisms Involved in Disease Progression and Whip Emission. *Front. Plant Sci.* 8, 1–17. doi:10.3389/fpls.2017.00882.

Schuster, M., Schweizer, G., and Kahmann, R. (2018). Comparative analyses of secreted proteins in plant pathogenic smut fungi and related basidiomycetes. *Fungal Genet. Biol.* 112, 21–30. doi:10.1016/J.FGB.2016.12.003.

Shapira, R., Altman, A., Henis, Y., and Chet, I. (1989). Polyamines and Ornithine Decarboxylase Activity during Growth and Differentiation in Sclerotium rolfsii. *Microbiology* 135, 1361–1367. doi:10.1099/00221287-135-5-1361.

Sharma, R., Mishra, B., Runge, F., and Thines, M. (2014). Gene Loss rather than gene gain is associated with a host jump from monocots to dicots in the smut fungus melanopsichium pennsylvanicum. *Genome Biol. Evol.* 6, 2034–2049. doi:10.1093/gbe/evu148.

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., and Ponting, C. P. (2014). Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* 15, 121–132. doi:10.1038/nrg3642.

Singh, N., Somai, B., and Pillay, D. (2004). Smut disease assessment by PCR and microscopy in inoculated tissue cultured sugarcane cultivars. *Plant Sci.* 167, 987–994. doi:10.1016/J.PLANTSCI.2004.05.006.

Souza, G. M., Van Sluys, M. A., Lembke, C. G., Lee, H., Margarido, G. R. A., Hotta, C. T., *et al.* (2019). Assembly of the 373k gene space of the polyploid sugarcane genome reveals reservoirs of functional diversity in the world's leading biomass crop. *Gigascience* 8, 1–18. doi:10.1093/gigascience/giz129.

Su, Y., Wang, Z., Xu, L., Peng, Q., Liu, F., Li, Z., *et al.* (2016). Early Selection for Smut Resistance in Sugarcane Using Pathogen Proliferation and Changes in Physiological and Biochemical Indices. *Front. Plant Sci.* 7, 1–10. doi:10.3389/fpls.2016.01133.

Sundar, A. R., Barnabas, E. L., Malathi, P., and Viswanathan, R. (2012). A mini-review on smut disease of sugarcane caused by Sporisorium scitamineum. *InTech Bot.*, 107–128.

Tanaka, S., Brefort, T., Neidig, N., Djamei, A., Kahnt, J., Vermerris, W., *et al.* (2014). A secreted Ustilago maydis effector promotes virulence by targeting anthocyanin biosynthesis in maize. *Elife* 3, 1–27. doi:10.7554/elife.01355.

Tanaka, S., Schweizer, G., Rössel, N., Fukada, F., Thines, M., and Kahmann, R. (2019). Neofunctionalization of the secreted Tin2 effector in the fungal pathogen Ustilago maydis. *Nat. Microbiol.* 4, 251–257. doi:10.1038/s41564-018-0304-6.

Taniguti, L. M., Schaker, P. D. C., Benevenuto, J., Peters, L. P., Carvalho, G., Palhares, A., *et al.* (2015). Complete Genome Sequence of Sporisorium scitamineum and Biotrophic Interaction Transcriptome with Sugarcane. *PLoS One* 10, e0129318. doi:10.1371/journal.pone.0129318.

Thokoane, L. N., and Rutherford, R. S. (2001). cDNA-AFLP differential display of sugarcane (Saccharum spp. hybrids) genes induceb by challenge with the fungal pathogen Ustilago scitaminea (sugarcane smut). *Proc S Afr Sug Technol Ass* 75, 104–107.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., and van Dijl, J. M. (2000). Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* 64, 515–547. doi:10.1128/MMBR.64.3.515-547.2000.

Toh, S. S., and Perlin, M. H. (2016). Resurgence of less-studied smut fungi as models of phytopathogenesis in the omics age. *Phytopathology* 106, 1244–1254. doi:10.1094/PHYTO-02-16-0075-RVW.

Tom, L., Braithwaite, K., and Kuniata, L. S. (2017). Incursion of sugarcane smut in commercial cane crops at Ramu Papua New Guinea. *39th Conf. Aust. Soc. Sugar Cane Technol. ASSCT 2017*, 377–384.

van der Linde, K., Hemetsberger, C., Kastner, C., Kaschani, F., van der Hoorn, R. A. L., Kumlehn, J., *et al.* (2012). A maize cystatin suppresses host immunity by inhibiting apoplastic cysteine proteases. *Plant Cell* 24, 1285–1300. doi:10.1105/tpc.111.093732.

Van Verk, M. C., Hickman, R., Pieterse, C. M. J., and Van Wees, S. C. M. (2013). RNA-Seq: Revelation of the messengers. *Trends Plant Sci.* 18, 175–179. doi:10.1016/j.tplants.2013.02.001.

Vettore, A. L., da Silva, F. R., Kemper, E. L., Souza, G. M., da Silva, A. M., Ferro, M. I. T., *et al.* (2003). Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res.* 13, 2725–2735. doi:10.1101/gr.1532103.

Vettore, A. L., da Silva, F. R., Kemper, E. L., and Arruda, P. (2001). The libraries that made SUCEST. *Genet. Mol. Biol.* 24, 1–7. doi:10.1590/S1415-47572001000100002.

Waclawovsky, A. J., Sato, P. M., Lembke, C. G., Moore, P. H., and Souza, G. M. (2010). Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. *Plant Biotechnol. J.* 8, 263–276. doi:10.1111/j.1467-7652.2009.00491.x.

Wada, A., Anaso, A., and Bassey, M. (2016). Sugar Cane Whip Smut (Sporisorium scitamineum Syd) Caused Field Sucrose and Juice Quality Losses of Two Sugar Cane Varieties in Nigeria. *Int. J. Plant Soil Sci.* 10, 1–11. doi:10.9734/ijpss/2016/24566.

Waller, J. M. (1969). Sugarcane smut (Ustilago scitaminea) in Kenya: I. Epidemiology. *Trans. Br. Mycol. Soc.* 52, 139–151. doi:10.1016/S0007-1536(69)80167-1.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi:10.1038/nrg2484.

Whittle, A. M. (1982). Yield loss in sugar-cane due to culmicolous smut infection. *Trop. Agric.* 59, 239–242.

Wu, Q., Xu, L., Guo, J., Su, Y., and Que, Y. (2013). Transcriptome Profile Analysis of Sugarcane Responses to. *Biomed Res. Int.*, 1–9. doi:http://dx.doi.org/10.1155/2013/298920.

You-Xiong, Q., Jian-Wei, L., Xian-Xian, S., Li-Ping, X., and Ru-Kai, C. (2011). Differential gene expression in sugarcane in response to challenge by fungal pathogen Ustilago scitaminea revealed by cDNA-AFLP. *J. Biomed. Biotechnol.*, 1–10. doi:10.1155/2011/160934.

Zhang, J., Zhang, X., Tang, H., Zhang, Q., Hua, X., Ma, X., *et al.* (2018). Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. *Nat. Genet.* 50, 1565–1573. doi:10.1038/s41588-018-0237-2.

Zuo, W., Ökmen, B., Depotter, J. R. L., Ebert, M. K., Redkar, A., Misas Villamil, J., *et al.* (2019). Molecular Interactions Between Smut Fungi and Their Host Plants. *Annu. ev. Phytopathol.* 57. doi:10.1146/annurev-phyto-082718-100139.

CHAPTER 2: The smut pathogen *Sporisorium scitamineum* uses different strategies to infect resistant and susceptible sugarcane genotypes

ABSTRACT

Sporisorium scitamineum, the causal agent of sugarcane smut disease, is a dimorphic and biotrophic fungus, host-dependent to its sexual reproduction and teliospores formation. The Genomics Group has been extensively studying the sugarcane-smut interaction during the past few years, considering various aspects of the interaction, including pathogen attack mechanisms and plant defense responses. In this study, we present a transcriptomic profile of fungal genes expressed in sugarcane plants resistant and susceptible to smut 48 hours after inoculation, and a comparison with previously obtained data of *S. scitamineum* in axenic culture (Taniguti *et al.*, 2015). Furthermore, we predicted the putative fungal secretome, including candidate effectors, and in association with secreted proteases and hydrolases, revealed their expression patterns among the three treatments. We suggested that different strategies are used by the pathogen to colonize genotypes with different genetic backgrounds considering resistance to smut. *S. scitamineum* infecting resistant plants induced proteases that may be responsible for triggering the host immune system. The success of *S. scitamineum* in reproducing in sugarcane susceptible plants may rely on sets of effector candidates to ensure the biotrophic growth. We unveiled the expression of some genes that may help to develop strategies for the fungal control, for instance, the glucose oxidase, used in several studies as an agent of fungal control. Last, we present some of the antioxidant enzymes (KatE, KatG and Prx1) potentially involved in the colonization of resistant plants.

Keywords: RNA-Seq; *Saccharum* spp.; Transcriptome

1. Introduction

Sugarcane (*Saccharum* spp.) is a major crop used for sugar and biofuels production. The world production of sugarcane in 2018 was 1.9 billion tons, cultivated in an area of 26.3 million hectares (FAO, 2018). In Brazil, the 2018 harvest resulted in 746.8 million tons (which corresponds to 39.3% of the global production), cultivated in 10 million hectares (FAO, 2018). Considering the crescent usage of sugarcane and its subproducts, alongside the great concern in the rational use of land, it is now imperative to seek increased production without increasing the agricultural frontier. Climate changes and the green harvest practices recently implemented in Brazil may influence the incidence of pests and diseases in sugarcane fields (Zhao and Li, 2015). Thus, a possible alternative is the control of damages caused by pests and diseases. Higher temperatures (25-30 ºC) and drier conditions usually favor the propagation and spread of diseases (Zekarias *et al.*, 2010), and for sugarcane, ratooning practices induce the

development of symptoms in infected plants (Akalach and Touil, 1996; Croft and Braithwaite, 2006).

Among the diseases that attack sugarcane crops, the smut is one of the most important as it can cause severe damage to the crop, reducing its yield and affecting the region's agriculture economy (Sundar *et al.*, 2012). This disease's causal agent is a biotrophic fungus from the *Ustilaginales* order, *Sporisorium scitamineum* (Syd.) M. Piepenbring, M. Stoll & Oberw (Piepenbring *et al.*, 2002). An elongated whip-like structure covered with spores resembling soot is the main symptom of the sugarcane smut disease. In this whip-like structure occurs fungal reproduction and teliospores production.

The selection of resistant genotypes is still the most efficient way to control smut disease. While there is no immunity to the fungus colonization, the number of whips developed in a given population of infected plants determine resistance or susceptibility. Sugarcane genotypes present different resistance levels to smut disease where less than 15% of whip formation, the genotype is considered resistant (Latiza *et al.*, 1980; Lemma *et al.*, 2015).

Understanding the transcriptional profile of the fungus *S. scitamineum* during the early infection of sugarcane can lead to the identification of fungal genes involved in modulating the plant metabolism and pathogenicity determinants. Among the strategies used by pathogens to penetrate and establish colonization is the secretion of enzymes and effectors. Herein, we investigated the fungal transcriptome 48 hours after inoculation (hai) of two contrasting sugarcane genotypes focusing on secreted proteins and candidate effectors.

2. Material and Methods

2.1 Biological material

We used two sugarcane genotypes with different levels of resistance to smut. Previously, Peters *et al.* (2017) inoculated single-bud sugarcane sets of 10-month-old healthy plants of the SP80-3280 (resistant) and IAC66-6 (susceptible) genotypes using the SSC39 strain of *S. scitamineum*. They then collected tissues containing the meristematic region of twenty buds from each genotype at time-points of 6, 12, 24, 48 and 72 hours after inoculation (hai) and performed a quantitative real-time PCR for fungal DNA quantification in infected tissues (Peters, 2016). The experiment was conducted in biological triplicates and entirely randomized for each treatment. In this work, we used the material collected 48 hai to conduct the RNA-Seq experiments.

2.2 RNA extraction, library preparation, and sequencing

Tissues collected 48 hai, were immediately frozen and grounded to a powder in liquid nitrogen. We used the commercial Kit Direct-zol^{TM} RNA MiniPrep (Zymo Research) according to the manufacturer's instructions to conduct the RNA extraction. The samples were treated with DNAse (Sigma), checked for RNA quality and quantity, and the RNA-Seq paired-end library prepared with the TruSeq RNA Sample Prep v2 Low Throughput (LT) kit (Illumina). Sequencing was performed in the HiSeq2500 Illumina System (2x100bp).

2.3 Preprocessing and mapping

First, we used the FastQC v0.11.5 (Andrews, 2010) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to verify the quality of the reads obtained in each file relative to the biological triplicates. We then conducted the trimming and quality filtering using the Cutadapt v1.18 (Martin, 2011), removing the adapters and keeping only the reads with no "N" bases and Phred quality score higher than 20 on average. Using Hisat2 v2.1.0 (Kim *et al.*, 2015), we mapped with default parameters the reads to the reference genome of *S. scitamineum* SSC39B (Taniguti *et al.*, 2015) available at the GenBank (NCBI) under the accession number GCA 001010845.1. For the *in vitro* transcriptome analysis, we used the data collected by Taniguti *et al.* (2015), publicly available under the BioProject number PRJNA275890.

2.4 Expression analysis, correction, and comparisons

To visualize each treatment's gene expression patterns, we used Circos Software v0.69-9 (Krzywinski *et al.*, 2009) to present the data after transformation. We used Fragment Per Kilobase of Gene per Million of Reads Mapped (FPKM) normalization of the counts and calculated the average among the biological triplicates following the formula:

$$FPKM = \frac{\text{Number of fragments mapped to a gene}}{\frac{\text{gene length in bp}}{10^3}\,\frac{\text{Total number of mapped reads}}{10^6}}$$

This normalization allowed us to compare the general expression profile among the treatments analyzed: *S. scitamineum* at 48 hai in the SP80-3280 (resistant) and IAC66-6 (susceptible) genotypes, with those publicly available of the fungus growing overnight in axenic culture (*in vitro*).

2.5 Differential expression analysis

We used FeatureCounts v1.6.0 from the Subread package (Liao *et al.*, 2013, 2014) to generate the count table. From the GFF3 (General Feature Format) file obtained in the NCBI (BioProject PRJNA275631), we generated a GTF (Gene Transfer Format) file of the reference genome, using the gffread package from Cufflinks v2.2.1 software (Trapnell *et al.*, 2010). To detect differentially expressed genes (DEGs), we used the EdgeR v3.30.3 from Bioconductor (BiocManager v3.11) R v4.0.2 (R Core Team, 2020) package (Robinson *et al.*, 2010; McCarthy *et al.*, 2012) with default parameters. We considered Genes with $P < 0.05$ as differentially expressed (DEGs) to compare the fungus infecting both contrasting genotypes, and $FDR < 0.05$ for the comparisons with the fungus growing in axenic culture.

2.6 Secretome and candidate effector genes prediction

To perform the *S. scitamineum* secretome prediction, we used the approach described in Taniguti *et al.* (2015). First, SignalP v4.1 (Petersen *et al.*, 2011) was used to predict signal peptides in the encoded proteins. Then, we verified the absence of transmembrane domains using TMHMM v2 (Krogh *et al.*, 2001), as well as the absence of GPI-anchors using the PredGPI platform (Pierleoni *et al.*, 2008) (http://gpcr.biocomp.unibo.it/predgpi/). We considered proteins as part of the secretome if they had signal peptide, and transmembrane domains and GPI-anchors were absent. Then we submitted the predicted secretome (305 proteins) to EffectorP v2.0 (Sperschneider *et al.*, 2018) to identify effector candidates. To complete the set of candidate effectors, we included the orthologs identified previously by Benevenuto *et al.* (2018) as part of smut fungi's core effectors, and other recent studies (Table 1).

Table 1 – Additional candidate effectors in *S. scitamineum*, orthologs of functionally characterized effectors from OrthoGroups in Benevenuto *et al.* (2018) and other recent studies.

***S. scitamineum* ID**	**Organism of functional characterization**	**Reference organism ID**	**Functional name**	**Reference functional characterization**
g6307_chr21_Ss	*U. maydis*	umag_05731	Cmu1	(Kämper *et al.*, 2006; Djamei *et al.*, 2011)
g674_chr01_Ss	*U. maydis*	umag_02475	Stp1	(Kämper *et al.*, 2006; Schipper, 2009; Liang, 2012)
g1843_chr03_Ss	*U. maydis*	umag_02011	ApB73	(Stirnberg and Djamei, 2016)
g1816_chr03_Ss	*U. maydis*	umag_01987	Pep1	(Doehlemann *et al.*, 2009; Hemetsberger *et al.*, 2012)
g2337_chr05_Ss	*U. maydis*	umag_01375	Pit2	(Doehlemann *et al.*, 2011; Mueller *et al.*, 2013)
g4911_chr14_Ss	*U. maydis*	umag_05302	Tin2	(Brefort *et al.*, 2014; Tanaka *et al.*, 2014)
g4906_chr14_Ss	*U. hordei*	uhor_10022	UhAvr1	(Linning *et al.*, 2004; Ali *et al.*, 2014)
g3970_chr10_Ss	*U. maydis*	umag_03274	Rsp3	(Ma *et al.*, 2018)
g6535_chr22_Ss	*U. maydis*	umag_06098	Fly1	(Ökmen *et al.*, 2018)

2.7 Annotation and Gene Ontology

We used Blast2GO v5.2.5 (Conesa *et al.*, 2005) to attribute Gene Ontology (GO) terms to the predicted proteins encoded by genes expressed and defined as DEGs using EdgeR v3.30.3 pipeline. Then we performed the Enrichment of GO terms using ErmineJ v3.1.2 (Lee *et al.*, 2005; Ballouz *et al.*, 2017), including all the three GO domains (Biological Process, Molecular Function, and Cellular Component) in the Over-Representation Analysis (ORA), with the parameter's minimum and maximum genes size set to 5 and 1000, respectively.

2.8 Obtaining domains with InterPro numbers (IPR)

We downloaded the InterProScan (IPR) v5.39-77.0 database and ran it locally with all the 6,677 protein sequences encoded by the *S. scitamineum* genome to search for conserved domains. Then used Python3 custom scripts to select and add to a table containing the summarized information of the data analyzed, the IPR numbers, and domain names of each protein-encoded gene identifier.

2.9 Retrieving information of Enzyme Commission Numbers and EuKaryotic Othologous Groups (KOG)

Using Blast KOALA (Kanehisa *et al.*, 2016) in the KEGG platform (Kanehisa and Goto, 2000; Kanehisa *et al.*, 2017), we retrieved information about pathways and Enzyme Commission (EC) number of enzymes. We used the parameters taxonomy group – "Fungi"; and KEGG genes database set to "family_eukaryotes", to include all 6,677 proteins of the *S. scitamineum* in a single run. The Blast KOALA search retrieved KEGG Orthology (KO) and EC numbers, when available. We performed an EggNOG mapping using the EggNOG v5.0 database and the EggNOG-mapper v2 with default parameters (Huerta-Cepas *et al.*, 2019) to get KOG categories (Tatusov *et al.*, 2003).

3. Results

3.1 RNA-Seq reads quality filtering and mapping

The Illumina data of two sugarcane genotypes (susceptible and resistant) infected with *S. scitamineum* SSC39 48 hai and respective control samples (12 libraries) generated 211,940,408 reads. After trimming and quality filtering, 207,043,218 high-quality reads remained in the analysis. This study used six libraries, including three replicates composed of inoculated plants of the resistant genotype and three replicates of inoculated plants of the susceptible genotype. These libraries generated a total of 97,988,470 reads. After adapters removal and quality filtering, we maintained 95,822,110 of high-quality reads, approximately 0.13% aligned to the *S. scitamineum* SSC39 reference (Table 2). We also re-analyzed the raw data (64,234,488 reads) of three libraries sequenced of *S. scitamineum* growing in axenic culture previously obtained (Taniguti *et al.*, 2015). After trimming, we maintained 59,989,160 high-quality reads, of which 58,707,848 mapped to the reference (approximately 97.8%) (Table 2).

Table 2 – Transcriptional profiling of *S. scitamineum* infecting smut-resistant and -susceptible plants of sugarcane 48 hours after inoculation.

Library	Sample	Total number of reads after trimming	Mapped *S. scitamineum* paired reads
Susceptible	1	5,188,522	51,148
Susceptible	2	14,595,736	21,362
Susceptible	3	16,225,566	22,794
Resistant	1	17,793,274	46,562
Resistant	2	22,524,820	52,256
Resistant	3	19,494,192	81,364
Axenic culture*	1	18,728,470	18,364,234
Axenic culture*	2	22,405,608	21,902,169
Axenic culture*	3	18,855,082	18,441,445

*Data obtained from Taniguti *et al.* (2015).

3.2 *Sporisorium scitamineum* differentially expressed genes when interacting with sugarcane

We first analyzed the differentially expressed genes (DEGs), considering the infected plants' libraries of the resistant and susceptible genotypes. Because the number of reads of the pathogen in both libraries was low (considering the pool of plant-pathogen tissues), we analyzed only coding sequences (CDSs) having at least one count per million (CPM) mapped reads in all three biological replicates in both treatments (preferentially expressed genes). These data entered the analysis of DEGs in the EdgeR v3.30.3 pipeline (Robinson *et al.*, 2010; McCarthy *et al.*, 2012). The approach generated 1,076 preferentially expressed genes, in which 113 were DEGs with a $P < 0.05$. Out of these genes, 63 were down-regulated and 50 up-regulated in the resistant genotype compared to the susceptible one (Supplementary Table 1). The average distribution of mapped reads to *S. scitamineum* chromosomes demonstrates that the fungus preferentially expressed some genes in each of the treatments, *i.e.*, during plant tissues colonization and *in vitro* growth (Figure 1). The proteins encoded by these 113 genes were further examined using GO annotation enrichment analysis and KOG functional categorization.

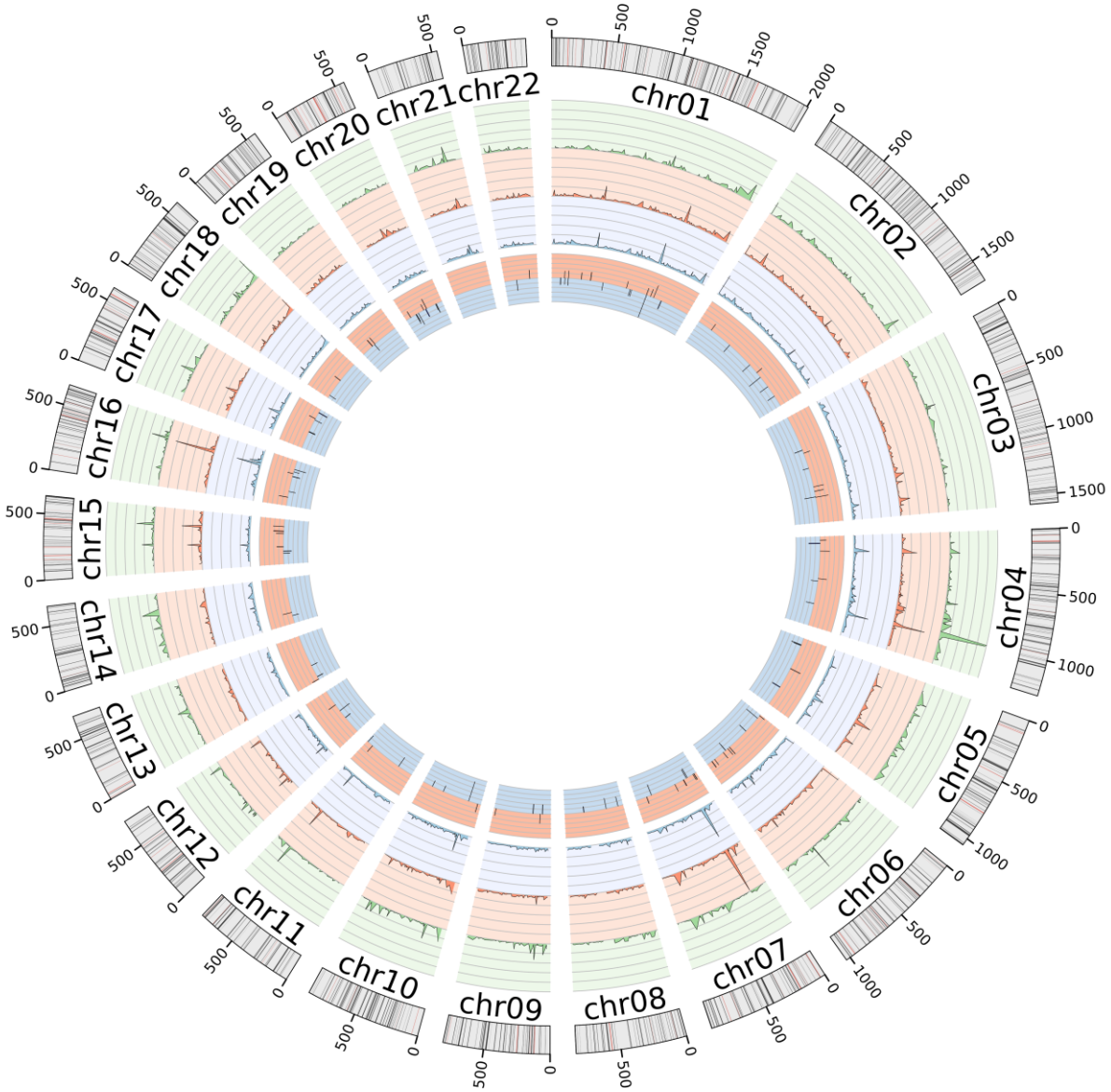


Figure 1 – Distribution of transcriptome reads mapped to *S. scitamineum* chromosomes. The first outer circle represents 22 of 26 *S. scitamineum* chromosomes, in which we detected expressed genes. Each chromosome has the length represented in kbp (placed outside the circle). The outermost circle exhibits black lines representing chromosome location of preferentially expressed genes (in the Resistant – Susceptible comparison at 48 hai), and red lines representing DEGs. The next three circles exhibit genes expressed (average of FPKM normalized) of *S. scitamineum*: 1) growth in axenic culture (green); 2) infecting resistant plants (SP80-3280) (red); and 3) infecting susceptible plants (IAC66-6) (light blue), respectively. The innermost circle exhibits the log2FC for DEGs (red-up-regulated and blue-down-regulated), comparing gene expressions during the infection of resistant versus susceptible plants.

3.3 Functional categorization of DEGs (incompatible vs. compatible interaction)

The genes resulting from our first approach, *i.e.*, those that were differentially expressed when we compared the experiments of *S. scitamineum* infecting resistant plants versus *S.*

scitamineum infecting susceptible plants, were submitted to two functional enrichment analysis. To provide an overview of the molecular events modulated in the process of infection of both the resistant and susceptible plants, we used a KOG analysis, assigning classes according to their expression profiles (up or down-regulated) (Figure 2). Most of the genes were either not assigned to any KOG category (no hit) or classified as unknown functions (category S). Generally, most of the genes differentially expressed in *S. scitamineum* infecting resistant plants were down-regulated (63). The three categories with the most DEGs were amino acid transport and metabolism (E); lipid transport and metabolism (I); and posttranslational modification, protein turnover, chaperones (O). From the 50 DEGs induced in *S. scitamineum* infecting resistant plants, 20 genes fell in categories related to energy production and conversion (C), carbohydrate transport and metabolism (G), protein modification and turnover (O), secondary metabolism (Q), and signal transduction (T). In contrast, all other biological processes had most of the DEGs induced during susceptible plants' infection.

The three GO terms identified as enriched using ErmineJ v3.1.2 (Lee *et al.*, 2005; Ballouz *et al.*, 2017) software ($P < 0.05$) were DNA metabolic processes (GO:0006259), protein binding (GO:0005515), and cellular response to a stimulus (GO:0051716). Five DEGs were related to the cellular response to a stimulus, three of them induced in the resistant plant infection (g1075_chr02_Ss, catalase KatE; g4614_chr13_Ss, catalase-peroxidase KatG; g2922_chr06_Ss, related to UV-damaged DNA-binding protein) and two repressed (g5571_chr17_Ss, spore specific protein 1; g5947_chr19_Ss, histone chaperone FACT POB3). The GO term GO:0005515 also revealed three other genes repressed related to chromatin in resistant plant infections (g2845_chr06_Ss, histone H4; g2846_chr06_Ss, histone H3.2; g5434_chr16_Ss, histone H3.1) and a PITH-domain containing protein, which is a general proteasome-interacting module (PS51532).

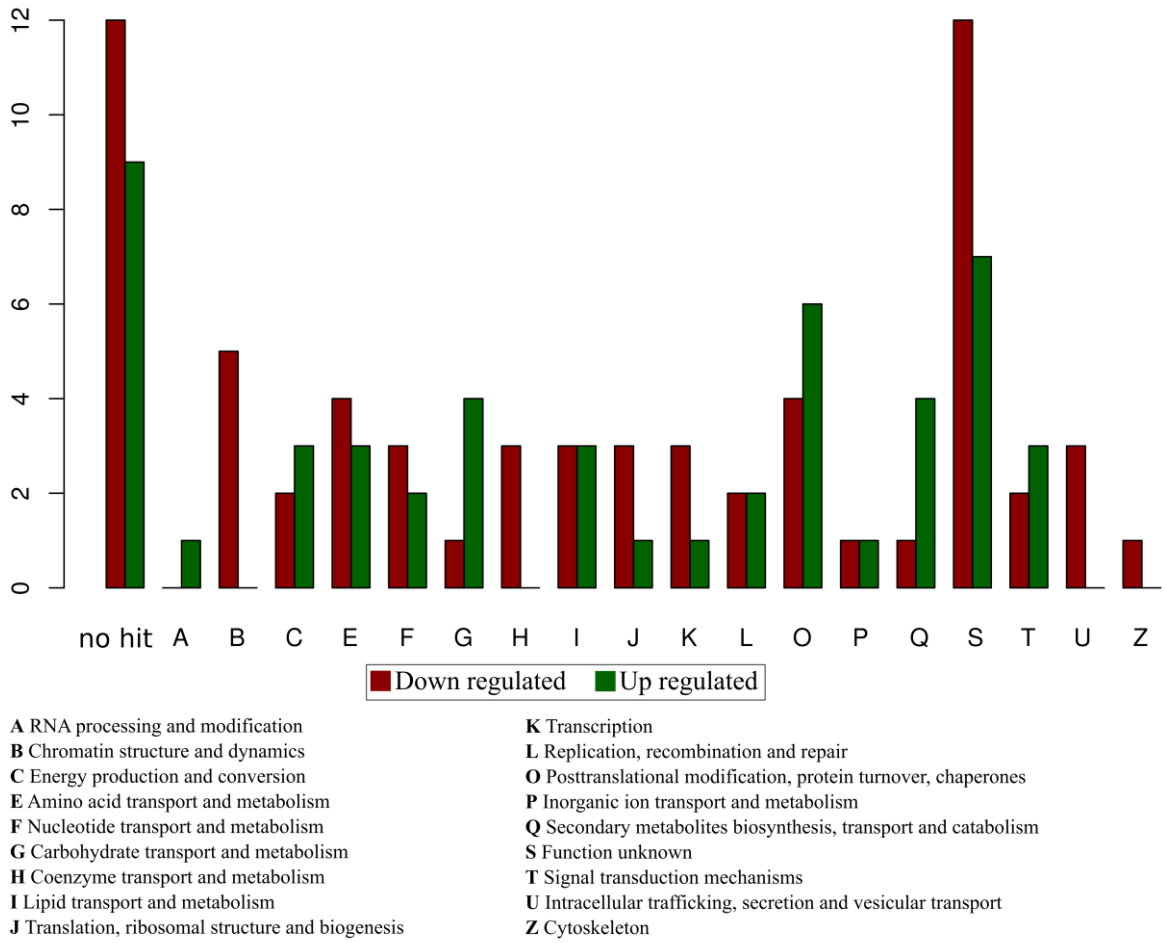


Figure 2 – KOG categories for the *S. scitamineum* differentially expressed genes in both genotypes (SP80-3280 (resistant) – green and IAC66-6 (susceptible) - red). The "no hit" category indicates that no KOG class were assigned to the gene. The Y axis represents the frequency of genes in each category. The X axis represents each of the KOG category.

3.4 DEGs related with the plant-pathogen interaction

Some of the DEGs were suggestive to show relevant differences of *S. scitamineum* gene expression when the fungus was infecting susceptible and resistant plants. For instance, when infecting susceptible plants, the fungus induced the expression of genes potentially involved with chromatin structure and dynamics as mentioned before, besides also inducing other five genes encoding homologous of proteins related to the ubiquitin-dependent pathway for protein recycling (g3992_chr10_Ss, protein mlo2; g5105_chr15_Ss, related to SIZ1-E3-like factor in the SUMO pathway; g1892_chr04_Ss, proteasome-interacting domain protein; g5570_chr17_Ss, ubiquitin-conjugating enzyme variant MMS2; and g2167_chr04_Ss, related to cop9 complex subunit 3), and five other genes relevant to growth and development (g2062_chr04_Ss, probable cyclophilin b; g2546_chr05_Ss, protein APG2-required for

sporulation; g75_chr01_Ss, HMF1-heat-shock inducible inhibitor of cell growth; g2998_chr07_Ss, quinone reductase; and g1019_chr02_Ss, HMG box-containing protein).

On the other hand, the fungus infecting resistant plants induced beside catalases katG and katE mentioned above, the pepsin g74_chr01_Ss and the protease g5644_chr17_Ss, related to PRC1-carboxypeptidase y involved in the degradation of small peptides (P00729); and a hydrolase g1926_chr04_Ss probably of GH23 family. The fungus also seems to induce carbon partitioning changes up-regulation FBP1-fructose-1,6-biphosphatase (g2833_chr16_Ss), glucose oxidase (g5302_chr16_Ss), and phosphoenolpyruvate carboxykinase (g3602_chr09_Ss). The analysis also revealed two kinases induced, one uncharacterized Ste20-like kinase Don3 and the serine/threonine-protein kinase gad8, potentially involved in a signaling module for sexual development and cell growth under stress conditions (Q9P7J8).

3.5 Genes highly expressed *in planta*

We used a second approach to investigate which genes were highly expressed in the fungus when infecting plant tissues compared to the fungus growing in axenic culture. We considered highly expressed those genes whose expression was a fold change greater than 10, considering the three replicates' FPKM average. We reasoned that if a fungal gene was highly expressed during the interaction when our experiment produced a low total number of genes expressed, it must be relevant to the plants contact. The results showed 521 genes in these conditions, of which 197 were shared among treatments, considering fungal colonization of the two sugarcane genotypes (Figure 3, Supplementary Figure 1).

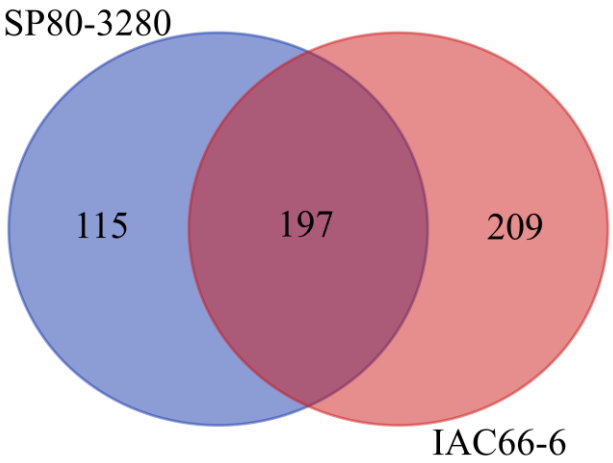


Figure 3 – Venn diagram comparing the highly expressed genes (fold change > 10, in average, compared with the expression of genes when the fungi cells were growing in axenic culture) in both resistant and susceptible genotypes.

To further enrich our understanding of the top genes most relevant to the fungus when colonizing the plants 48 hai, we defined the genes only expressed *in planta* plotting altogether in a Venn diagram the expressed genes in each treatment (at least one CPM in the three biological replicates). We captured 21 genes expressed only *in planta*, 16 exclusively expressed during the smut-resistant genotype infection, one in the -susceptible, and four shared in both genotypes (Figure 4). From the 16 genes expressed exclusively when *S. scitamineum* was infecting the resistant genotype, most of them encoded hypothetical or uncharacterized proteins (12); and the other four translated homologous of a subtilisin-like serine protease (g3042_chr07_Ss), a probable aldehyde dehydrogenase (g1282_chr03_Ss), a DNA binding protein containing HGM-box Hmg3 (g1282_chr03_Ss), and a protein related to a POL polyprotein of retrovirus (g4048_chr10_Ss). The only gene expressed solely in smut-susceptible plants encoded a sugar transporter of the Major Facilitator Superfamily (MFS) (g6215_chr20_Ss). From the four of the remaining genes expressed in both genotypes, two encoded transmembrane proteins related to *U. maydis* virulence factor Dik6 (g4409_chr12_Ss and g4169_chr11_Ss); one a C_2H_2-type zinc fingers transcription factor (g3838_chr10_Ss); and the last one encoded a glycosyltransferase of the group 2 family (g6010_chr19_Ss).

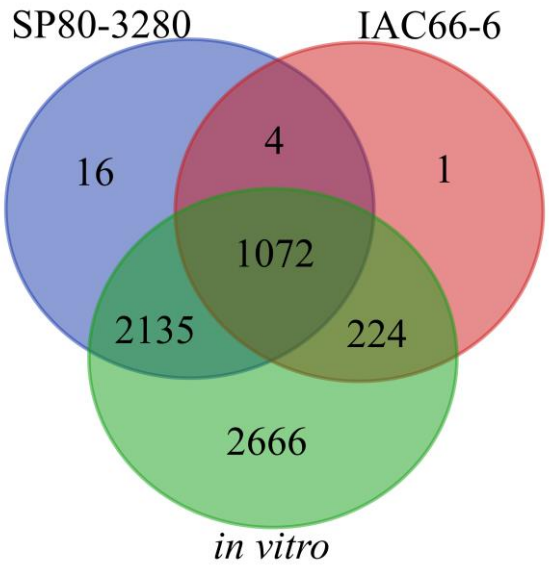


Figure 4 – Venn diagram comparing *S. scitamineum* expressed genes during infection of smut-resistant (SP80-3280) and -susceptible (IAC66-6) plants compared to expression in cells growing in axenic culture (*in vitro*).

At the end of these three strategies, 611 genes (Figure 5A) composed the set used to define molecular events most relevant to the fungus when infecting susceptible or resistant plants, different responses, and commonalities to be further explored as targets to prevent fungal colonization. Among the selected genes, 28 were solely present in the genome of *S. scitamineum*

and potentially only relevant to the interaction with sugarcane; another 309 were conserved among various fungi but with uncharacterized functions, and the remaining 274 had an annotated homolog in a public database (Table 3, Supplementary Figures 2 and 3).

Because secreted proteins are usually relevant to infection's initial phases, we analyzed them separately (Figure 5B, Supplementary Figure 4). Forty-one genes composed this set of encoded-proteins, showing a specific gene expression profile for each treatment, among them were candidate effectors (effectors of the Eff and Mig families and Pep1); proteases (subtilisin, aspartic protease), hydrolases (lysozyme, pectin lyase, glucosidases), and cell wall modification (chitin deacetylation, chitin-binding protein).

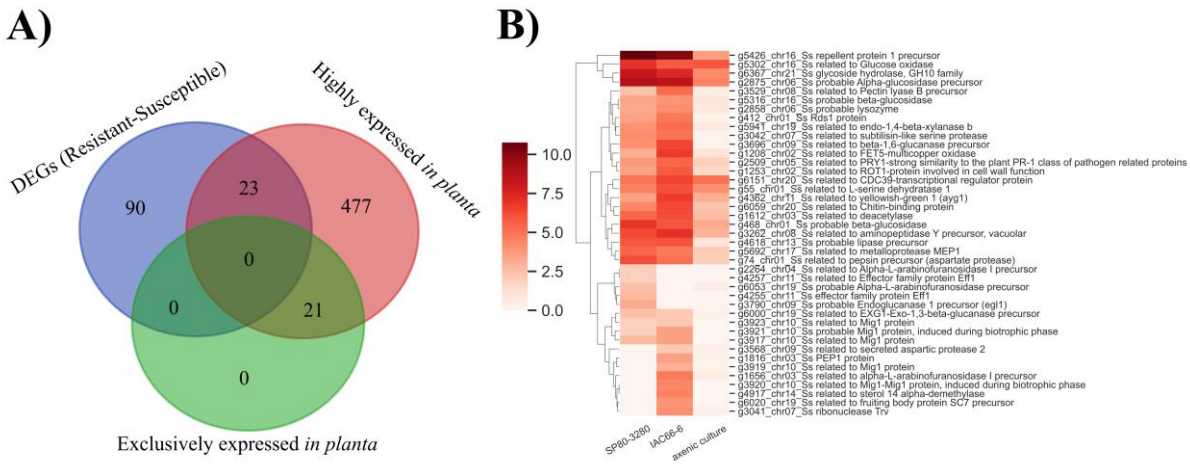


Figure 5 – Overview of the three strategies used to select genes relevant to plant tissues' fungal colonization. A) Venn diagram comparing the three strategies analyzed. B) Heatmap using log2 from FPKM values (average of the biological triplicate), for the annotate secreted protein-encoding genes.

Table 3 - Secretome prediction, and annotation status of the genes selected for further studies.

Secretome	Annotation	Number of genes
Non-secreted	Hypothetical	23
	Conserved Hypothetical	252
	Annotated	233
Secreted	Hypothetical	5
	Conserved Hypothetical	57
	Annotated	41
	TOTAL	**611**

3.6 Functional categorization of the genes set using the three strategies (Enriched GO and KOG analyses)

The set of genes selected using the three strategies, *i.e.*, DEGs from preferentially expressed genes; Highly expressed *in planta*; and Exclusively expressed *in planta*, comprised a set of 611 protein-encoding genes, and were submitted as previously to two functional categorizations of GO and KOG. Because we now included the data of the fungus' axenic culture growth, we revealed a new set of genes differentiating their expression profile in resistant and susceptible plants (Figure 6). Once again, most of the genes were either not assigned to any KOG category (no hit) or classified as an unknown function (category S). The analysis also revealed a larger number of genes modulated in the susceptible genotype (254) than in the resistant (216). The most striking differences were category G (Carbohydrate transport and metabolism) for infection in resistant plants and categories B (Chromatin structure and dynamics) and U (Intracellular trafficking, secretion, and vesicular transport) for infection in susceptible plants. The new analysis confirmed that carbon metabolism was most important to resistant plants' infection, increasing the number of genes unveiled from 4 to 24. Moreover, the new analysis revealed many genes involved in intracellular trafficking besides confirming the relevance of chromatin dynamics for the susceptible infection.

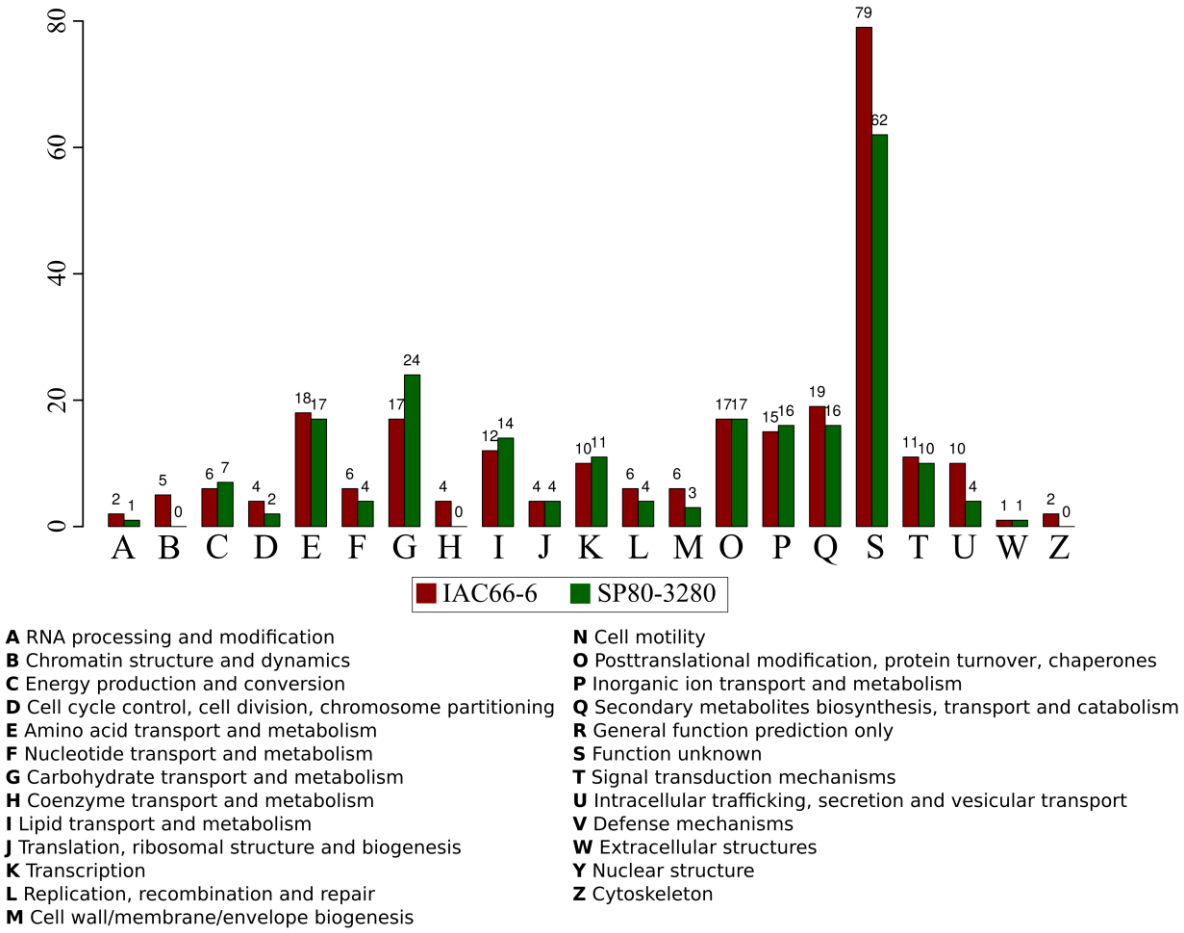


Figure 6 - KOG categories for the *S. scitamineum* expressed genes detected in the sugarcane genotypes (SP80-3280 (resistant) - green and IAC66-6 (susceptible) – red) summarizing three approaches (DEGs + Highly Expressed *in planta* + Exclusively Expressed *in planta*). The KOG classes that were not assigned to a gene (no hit) are not presented in this graph. The Y axis represents the frequency of genes in each category. The X axis represents each of the KOG category.

The GO enrichment conducted as before but applied for the 611 genes among all genes with an assigned GO term (4401) revealed twenty-three over-represented functional categories ($P < 0.05$) with the most expressive terms associated to the membrane (GO:0016020) 648/71 total genes/enriched genes; and an intrinsic component of membrane (GO:0031224) 580/66; transport (GO:0006810) 393/39; and transmembrane transport (GO:0055085) 281/39; and oxidation-reduction process (GO:0055114) 348/39 (Figure 7, Table 4).

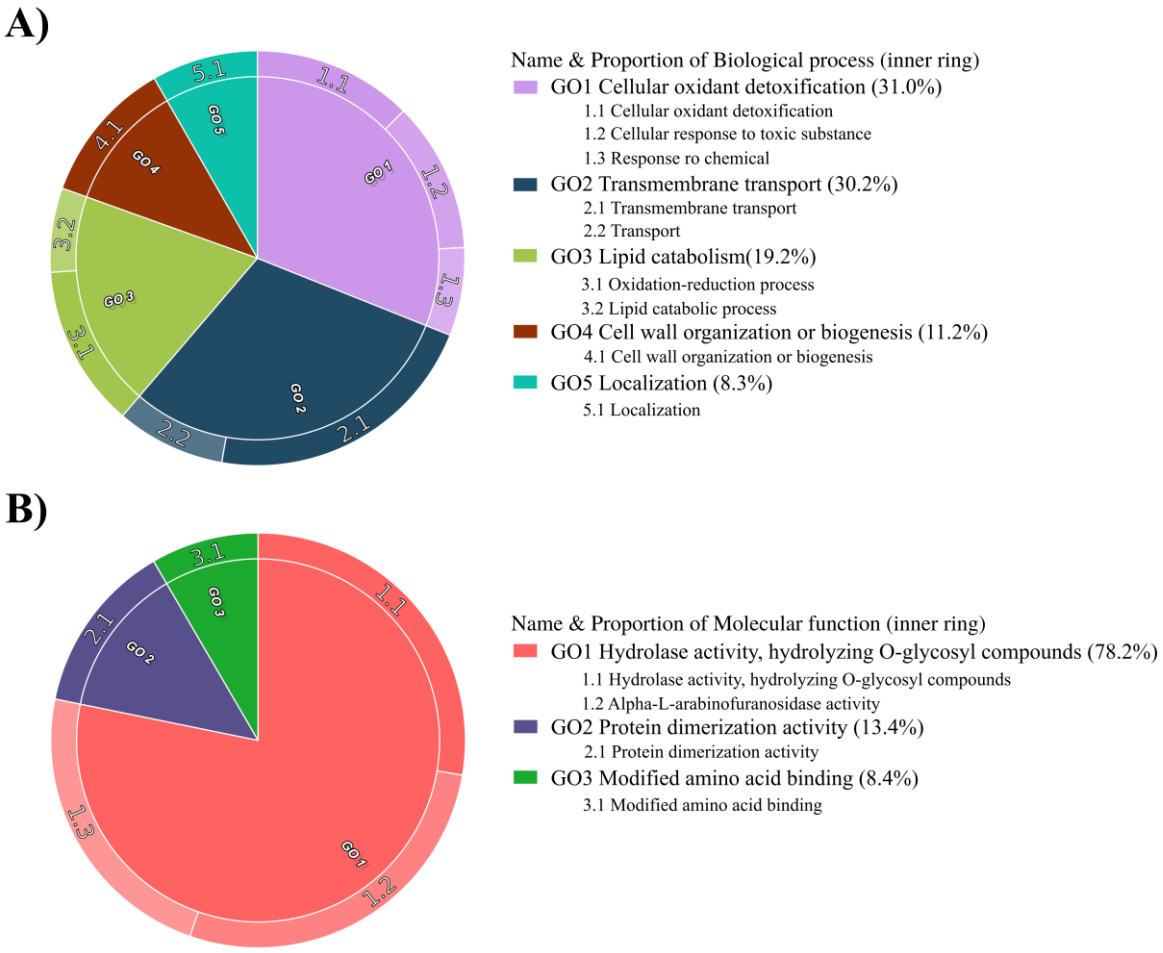


Figure 7 - Overview of the enrichment GO using three approaches (DEGs + Highly expressed *in planta* + Exclusively expressed *in planta*). A) Enrichment GO of Biological Process. B) Enrichment GO of Molecular Function. Figure generate using REVIGO (Supek *et al.*, 2011) and CirGO v1.0 (Kuznetsova *et al.*, 2019).

Table 4 – Enrichment GO summary. GO enrichment applied for all genes of the three approaches (307) related to all genes of *S. scitamineum* with an assigned GO term (4401).

GO term	GO ID	Total genes	Enriched genes	p-value
transmembrane transport	GO:0055085	281	39	0.000018
integral component of membrane	GO:0016021	579	66	0.000019
intrinsic component of membrane	GO:0031224	580	66	0.000020
hydrolase activity, hydrolyzing O-glycosyl compounds	GO:0004553	15	7	0.000030
alpha-L-arabinofuranosidase activity	GO:0046556	7	5	0.000030
membrane	GO:0016020	648	71	0.000034
membrane part	GO:0044425	612	67	0.000061
hydrolase activity, acting on glycosyl bonds	GO:0016798	19	7	0.000184
oxidation-reduction process	GO:0055114	348	39	0.001707
cellular oxidant detoxification	GO:0098869	14	5	0.001905
cellular response to toxic substance	GO:0097237	15	5	0.002696
cellular detoxification	GO:1990748	15	5	0.002696
cell wall organization or biogenesis	GO:0071554	10	4	0.003499
response to toxic substance	GO:0009636	16	5	0.003701
detoxification	GO:0098754	16	5	0.003701
protein dimerization activity	GO:0046983	41	8	0.006475
cellular response to chemical stimulus	GO:0070887	21	5	0.012919
transport	GO:0006810	393	39	0.013756
establishment of localization	GO:0051234	395	39	0.014878
localization	GO:0051179	408	40	0.015300
response to chemical	GO:0042221	26	5	0.031386
lipid catabolic process	GO:0016042	11	3	0.036578
modified amino acid binding	GO:0072341	5	2	0.042235

3.7 Expression profiles comparison: *in planta* vs. *in vitro*

The comparison between IAC66-6 (susceptible) and *S. scitamineum* growing in axenic culture (*in vitro*), according to the methods used, revealed 1,296 genes preferentially expressed, and from these, 466 were DEGs with *FDR* < 0.05, 242 down-regulated and 224 up-regulated (Figure 8A). The enrichment analysis of GO terms revealed thirty enriched categories (P < 0.05) and the most expressive related to biosynthetic process: oxidation-reduction (GO:0055114), 102/45 total genes/enriched genes; organic substance biosynthetic process (GO:1901576), 73/34; biosynthetic process (GO:0009058), 73/34; and cellular biosynthetic process (GO:0044249), 70/33 (Figure 9A).

On the other hand, the comparison between SP80-3280 (resistant) and *S. scitamineum* growing *in vitro* revealed 3,207 genes as preferentially expressed, and 1,106 DEGs identified with an *FDR* < 0.05, 525 down-regulated and 581 up-regulated (Figure 8B). The analysis revealed thirty-two GO enriched categories ($P_{adjusted}$ < 0.05) with the most expressive terms similar to those identified in the susceptible infection: oxidation-reduction process (GO:0055114), 198/94 total genes/enriched genes, transmembrane transport (GO:0055085), 154/70, biosynthetic process (GO:0009058), 141/64, organic substance biosynthetic process (GO:1901576), 136/64 and cellular biosynthetic process (GO:0044249), 136/63 (Figure 9B).

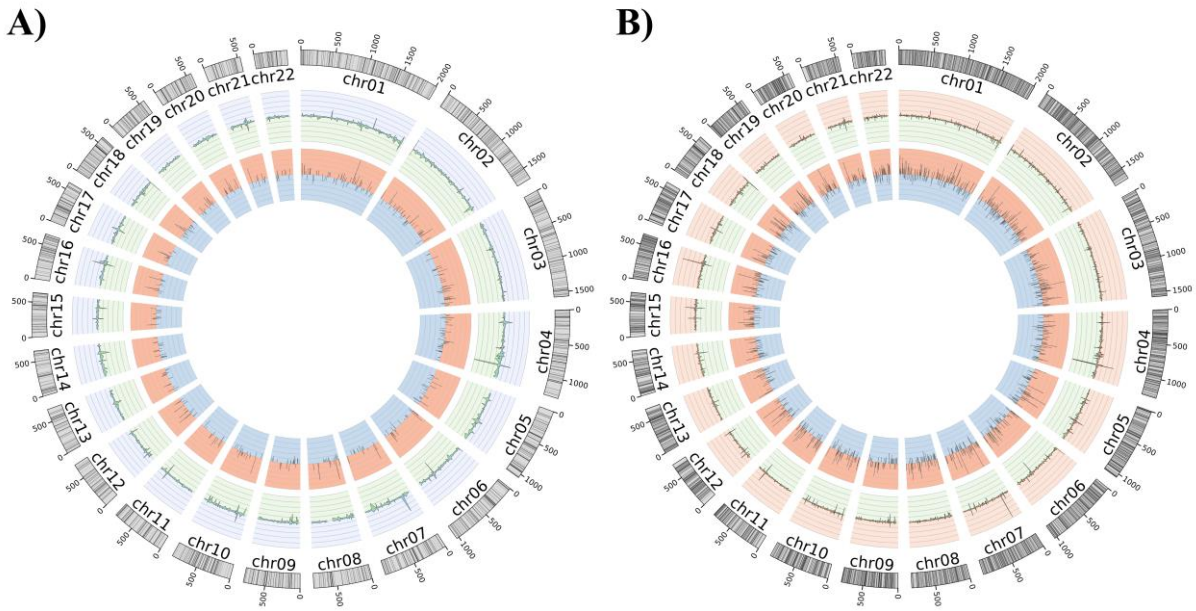


Figure 8 – Chromosomal overview of the expression data. The first outer circle represents 22 of 26 *S. scitamineum* chromosomes, indicating the length in kbp. The next two circles show the expression, FPKM normalized, of *S. scitamineum* infecting (A) IAC66-6 (light blue) and growing in axenic culture (green); and (B) SP80-3280 (light red) and growing in axenic culture (green). The last circle shows the log2FC for up-regulated (red) and down-regulated (blue) DEGs in (A) log2FC varying from -11 to 11, and (B) log2FC varying from -8 to 8.

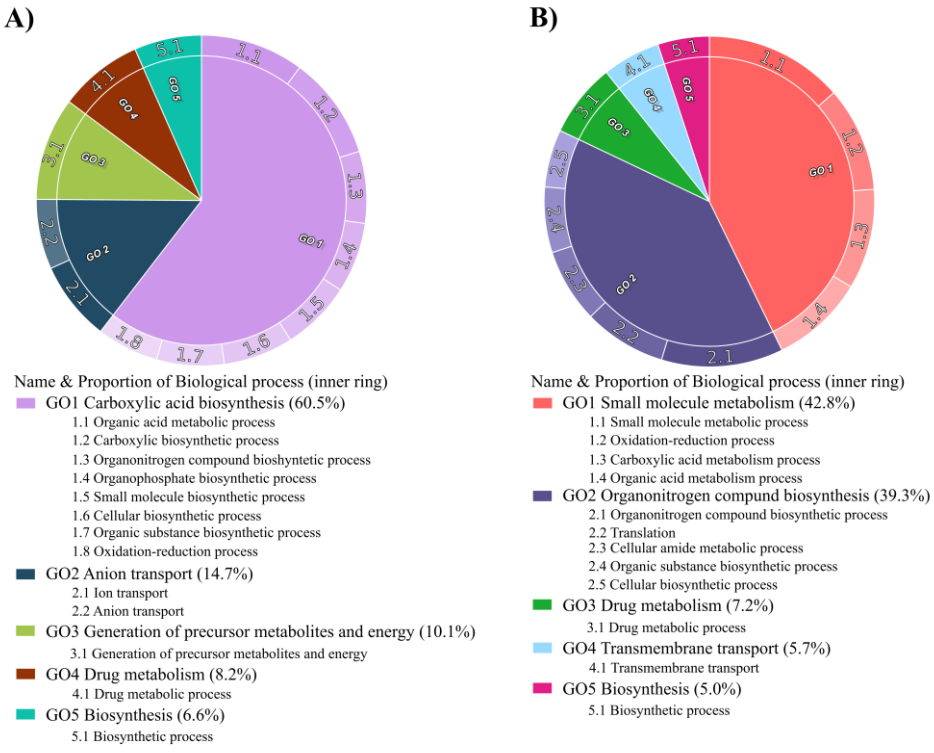


Figure 9 – Overview of the enrichment GO on the comparisons *in planta* vs. *in vitro*. A) IAC66-6 vs. *in vitro* growth. B) SP80-3280 vs. *in vitro* growth. Figure generate using REVIGO (Supek *et al.*, 2011) and CirGO v1.0 (Kuznetsova *et al.*, 2019).

The KOG analysis presented some categories enriched of genes up-regulated *in vitro*, such as energy production and conversion (C); amino acid transport and metabolism (E); and translation, ribosomal structure, and biogenesis (J) (Figure 10). Examining the comparison of plants infected of each genotype (resistant and susceptible) and axenic culture growth, we detected similar behavior for some and high discrepancies for others. For instance, some categories had more genes up-regulated in both genotypes (Figure 10), whereas categories of carbohydrate transport and metabolism (G); and intracellular trafficking, secretion, and vesicular transport (U) were contrasting between genotypes, having more down-regulated genes in the IAC66-6 (susceptible) and more up-regulated genes in SP80-3280 when compared to axenic growth (Figure 10). These same categories were detected before but unveiling an increased number of genes.

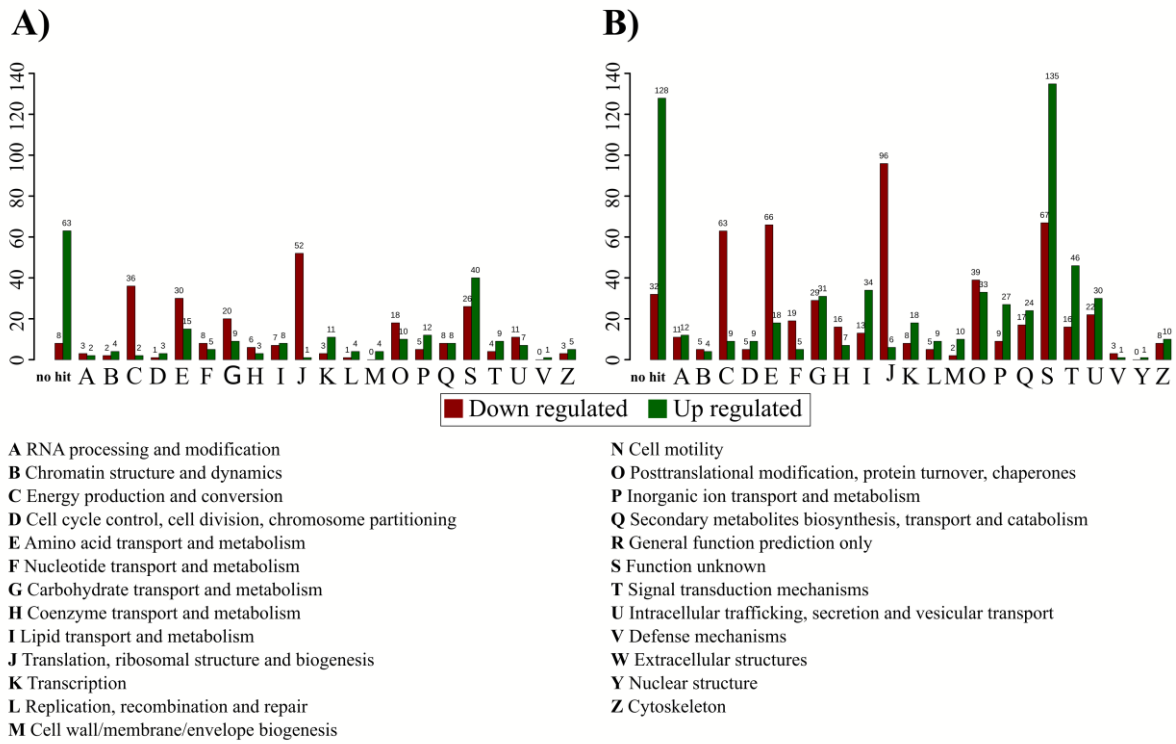


Figure 10 – KOG categories for the differentially expressed genes in both genotypes comparing with growing in axenic culture. A) – IAC66-6 (susceptible); B) – SP80-3280 (resistant). The "no hit" category indicates that no KOG class were assigned to the gene. The Y axis represents the frequency of genes in each category. The X axis represents each of the KOG categories.

3.8 Secretome and candidate effector genes prediction

We used the same approach as Taniguti *et al.* (2015) to define the secretome, obtaining 305 genes encoding proteins with a signal peptide and without transmembrane domains and

GPI-anchors. To predict the effector candidates, we performed an EffectorP v2.0 (Sperschneider *et al.*, 2018) analysis using the predicted secretome as input, obtaining 55 effectors candidates. Additionally, the effector candidate's dataset included orthologs genes of functionally characterized effectors in the smut model *U. maydis* (Table 1).

To explore the candidate effectors in the RNA-Seq data, we generated a heatmap with the log2 from FPKM values (average of biological triplicate) of the three treatments and added DEGs' information (Figure 11). We observed a set of candidate effectors genes more expressed when *S. scitamineum* infected the susceptible sugarcane genotype (central portion of the heatmap – Figure 11). Although these genes were not DEGs in any of the comparisons analyzed, because of the lack of reads mapped to all biological replicates (low-depth), they may have a potential role during the colonization of susceptible plants. Another set of genes identified only in the resistant genotype reasonably had the same explanation. Despite this, these are good candidates for further investigation.

We detected an up-regulated gene in all the three comparisons by exploring DEGs in the candidate effectors set. The gene g1450_chr03_Ss translate for a conserved hypothetical protein present among the cereal smuts, and it is another suitable candidate for further functional studies (Figure 11).

Among candidate effector DEGs identified in the infection of resistant plants, the fungus up-regulated the expression of five genes (g6307_chr21_Ss, Cmu1; g2067_chr04_Ss, conserved hypothetical protein; g1612_chr03_Ss, related to deacetylase; g6535_chr22_Ss, related to extracellular elastinolytic metalloproteinase precursor; and g706_chr10_Ss, conserved hypothetical protein), and down-regulated the expression of g5910_chr19_Ss – related to FPR2-FK506/rapamycin-binding protein of the endoplasmatic reticulum. We identified two DEGs, one up-regulated (g2412_chr05_Ss, conserved hypothetical protein), and one down-regulated (g2127_chr04_Ss, conserved hypothetical protein) analyzing the expression of genes in susceptible plants. Also intriguing is the repression of the candidate effector g3626_chr09_Ss (conserved hypothetical protein), when the fungus colonized the host plant independently if they were resistant or susceptible.

When focusing on the genes functionally characterized in other smut fungi, mainly in the smut model *U. maydis* (Table 1), we found two DEGs comparing resistant vs. axenic growth, the Cmu1 (g6307_chr21_Ss) and Fly1 (g6535_chr22_Ss). Other genes seemingly more expressed in infected susceptible genotype (Tin2, Stp1, ApB73, and Pep1), had reads mapped only to one biological replicate and will need further validation. The same consideration is valid for Pit2 and Rsp3 genes, apparently more expressed *in planta*. The effector ortholog of UhAvr1

did not present an expressive transcript level at 48 hai in any of the two sugarcane genotypes nor did in the axenic culture growth.

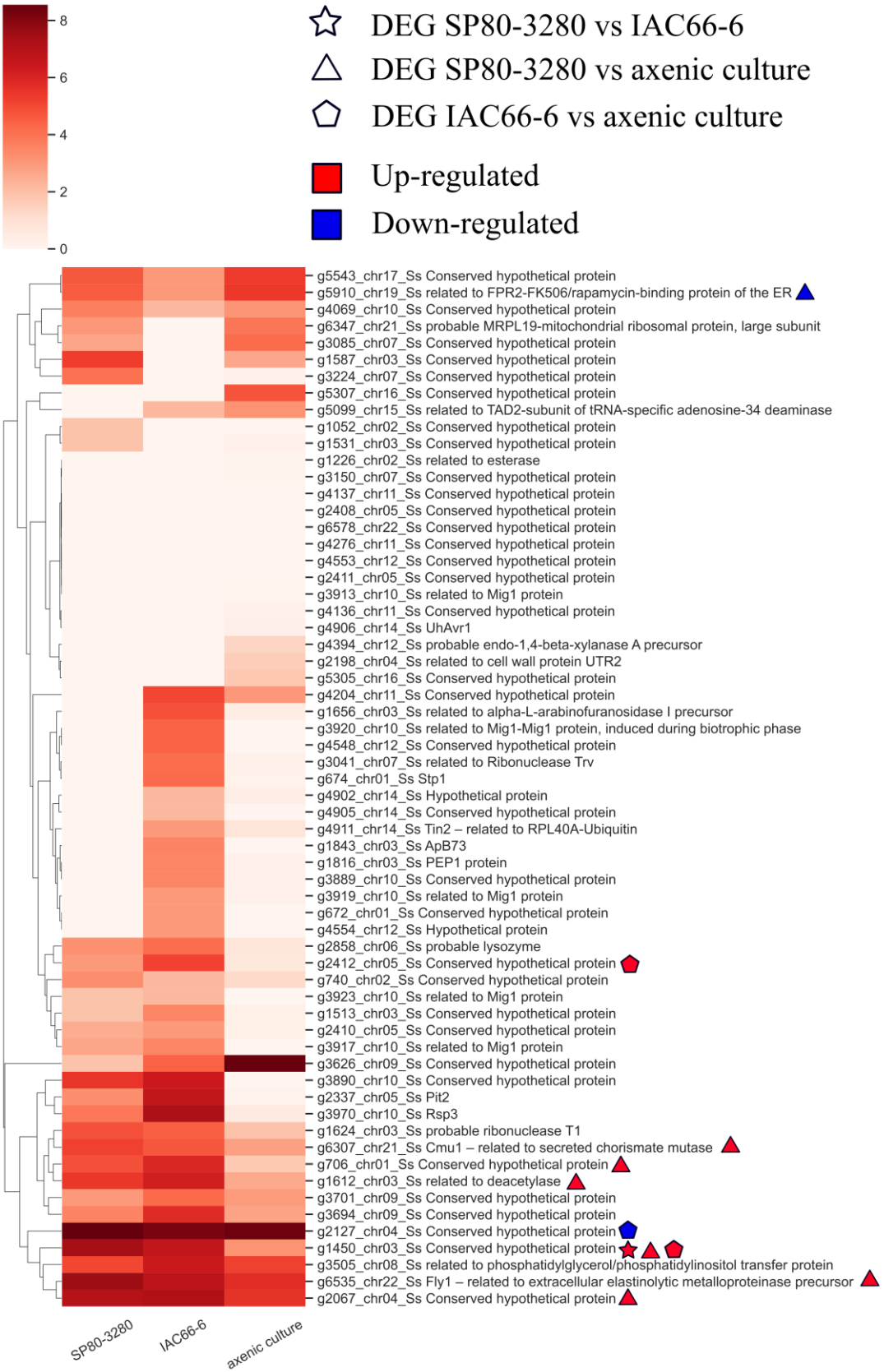


Figure 11 – Heatmap of candidate effectors genes presenting log2 of FPKM values (average of biological triplicate) for the three treatments used in this work. Symbols represent DEGs in a given comparison.

3.9 Mating and filamentation growth-realted genes

We searched for genes related to mating and filamentous growth because of their relevance considering plant infection. *S. scitamineum* presents a yeast-like growth in culture, and the fusion of two cells forming filamentous dikaryotic hyphae is essential for infecting plant tissues. We identified 19 genes potentially involved in mating and multicellular growth based on the various work developed for *U. maydis* and *S. scitamineum* (Shaw *et al.*, 1991; Colman-Lerner *et al.*, 2001; Smith *et al.*, 2003; König *et al.*, 2009; Martínez-Soto *et al.*, 2015; Chang *et al.*, 2019; Zhu *et al.*, 2019; Martínez-Soto *et al.*, 2020) (Figure 12).

They are involved in various steps considering signal transduction and sensing the environment (g5272_chr16_Ss; g5334_chr16_SS; g5909_chr19_Ss; g4003_chr10_Ss; g2842_chr06_Ss; g5892_chr19_Ss; g3706_chr09_Ss); cell wall changes (g496_chr01_Ss, g1612_chr03_Ss, g2818_chr06_Ss); multicellular growth (g1183_chr02_Ss; g5426_chr16_Ss); and internal REDOX control (g4614_chr13_Ss; g1075_chr02_Ss; g3315_chr08_Ss) (Figure 12).

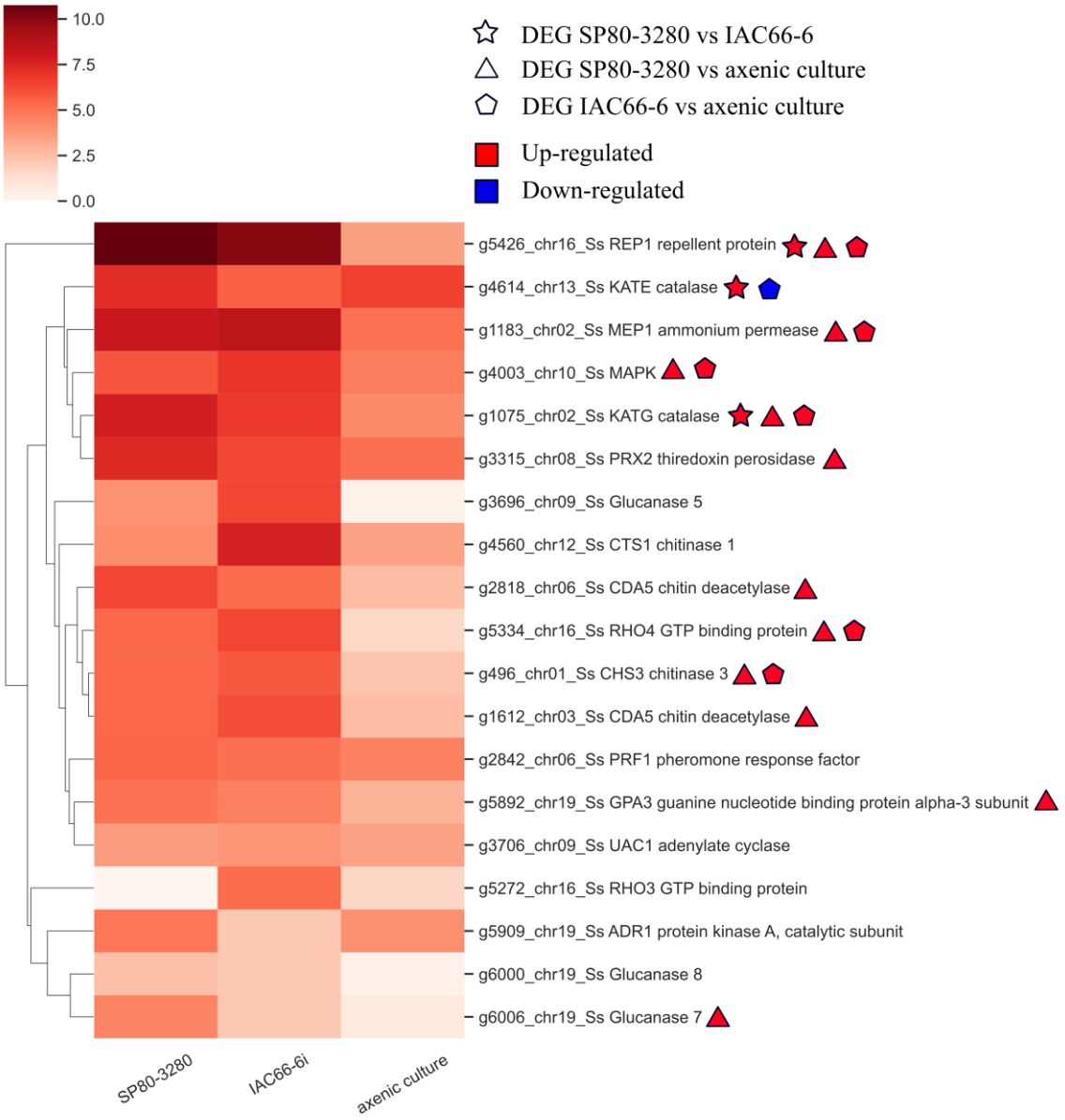


Figure 12 – Heatmap of genes related with mating/filamentation presenting log2 of the FPKM values (average of biological triplicate) for the three treatments used in this work.

4. Discussion

Defining *S. scitamineum* genes' expression profile when infecting sugarcane genotypes with contrasting smut-resistance levels can lead us to a more comprehensive knowledge of its infection mechanisms and help develop efficient control strategies. In this study, we applied high-throughput RNA-Seq technology to identify DEGs in the transcriptome of *S. scitamineum* infecting sugarcane comparing to the fungal axenic growth.

4.1 RNA-Seq reads mapping

In our analysis, 0.13% of all reads aligned with *S. scitamineum* reference genome at 48 hai. Less than 40% of all *S. scitamineum* genes were detected as expressed at this time. The expected low in-depth coverage of fungal reads in our RNA-Seq data is associated with the amount of DNA present in plants resistant and susceptible to smut 48 hai in infected tissues. Peters (2016) determined that approximately one ng per 100 ng of total DNA of *S. scitamineum* was present in tissues of infected susceptible plants. Whereas at the same time, only one-quarter of that was present in tissues of infected resistant plants. These results are not surprising as faced in other studies of dual transcriptome analysis. For instance, Kawahara *et al.* (2012), studying the interaction between rice (*Oryza sativa* L.) and blast fungus (*Magnaporthe oryzae*), reported that around 0.2% of reads aligned to the pathogen reference genome at 24 hai. As other authors extended the time before the isolation of infected tissues working with sugarcane-smut, the number of reads recovered also increased, detecting more expressed genes (Taniguti *et al.*, 2015; Schaker *et al.*, 2016). The total amounts detected were 2% and 18% for 5 and 200 dai, respectively. Despite that, we explored the data with multiple approaches and performed differential expression analysis obtaining results with statistical significance. The reference genome of *S. scitamineum* is a high-quality chromosome-scale assembly from telomere-to-telomere (Taniguti *et al.*, 2015), which is essential when working with low in-depth RNA-Seq data (Naidoo *et al.*, 2018).

4.2 Plant-pathogen interaction in the compatible vs. incompatible interactions

Peters *et al.* (2017) demonstrated that the sugarcane smut-pathogen induced an early modulation of reactive oxygen species (ROS) when infecting a resistant sugarcane genotype (SP80-3280). Specifically, 48 hai *S. scitamineum* proved to delay appressorium formation and colonization compared to an infection of a susceptible genotype (IAC66-6). The authors studied the modulation of the antioxidant system gene expression and antioxidant enzyme activities of sugarcane in addition to the quantification of hydrogen peroxide and reactive oxygen molecules in infected plants. Here, two out of three DEGs related to peroxidases up-regulated their expression in resistant plant's infection. These were the g1075_chr02_Ss encoding the catalase KatE (EC:1.11.1.6) and g4614_chr13_Ss, encoding catalase KatG (EC:1.11.1.21). They both catalyze hydrogen peroxide decomposition into water and oxygen, preventing ROS from being deadly to the cells and maintaining the cell's REDOX status (Peters *et al.*, 2020 in press).

Supposedly, because resistant plants detect the pathogen contact and promote the oxidative burst (Peters *et al.*, 2017), the smut fungi survive the immune system dismantling the plant host's ROS avoiding cell death. Noteworthy, smut pathogen penetrates and colonizes resistant plants, however with a lower number of cells than in susceptible plants (Carvalho *et al.*, 2016; Peters, 2016; Peters *et al.*, 2017).

Proteases act to inactivate host defense molecules and break down molecules to serve as nutrients to support growth in different lifestyles (Muszewska *et al.*, 2017). Under limited nitrogen and carbon sources, pathogens use ammonia, certain amino acids, proteins, and polyamines to allow host tissue colonization. We identified two secreted proteases among DEGs and induced in the resistant genotype (up-regulated), a homolog of the vacuolar serine-type carboxypeptidase y (g5644_chr17_Ss; P00729; MEROPS family S10), involved in the degradation of small peptides in conditions of nitrogen starvation (Parzych *et al.*, 2018); and a secreted aspartyl protease (g74_chr01_Ss; MEROPS family S), that in some cases were associated with host protein degradation (Mandujano-González *et al.*, 2016). Besides, two other highly expressed secreted proteases, one of the subtilisin-type serine protease (g3042_chr07_Ss; MEROPS family S8) (Figueiredo *et al.*, 2018), and the other, an aminopeptidase y-type (g3262_chr08_Ss; MEROPS Peptidase M28 family) (Richards *et al.*, 2012), were identified in resistant infected plants. Subtilisin-type serine proteases were associated with pathogen recognition and immune priming (Figueiredo *et al.*, 2014), providing a candidate for further investigation in *S. scitamineum* ability to infect resistant plants and its potential use for priming sugarcane immune system.

On the other hand, instead of proteases, hydrolases were mostly up-regulated in the susceptible genotype. Eighteen genes encoded hydrolases more expressed in plant tissues than *in vitro*. Among them, secreted (14) and non-secreted proteins (4) acting on cellulose, hemicellulose, pectin, chitin, arabinose, xylan, glucan, lipid mostly supported the filamentous growth and allowed colonization of the fungus in both resistant and susceptible infections (Zhao *et al.*, 2013; Ene *et al.*, 2014).

S. scitamineum also induced the expression of genes involved in chromatin structure modulation and proteasome-dependent protein degradation for susceptible plants' infection. The impact of chromatin modification on gene expression regulation during infection lacks a precise definition. However, in some cases, *U. maydis* has served as a model for fungal plant pathogenesis (Elías-Villalobos *et al.*, 2019).

The ubiquitin/proteasome system contributed to protein recycling under nutrient-limiting in *C. albicans* (Leach *et al.*, 2011; Ene *et al.*, 2014), and for some plant-pathogens

contributed to virulence (Liu and Xue, 2011). We suggested that when invading a susceptible genotype, *S. scitamineum* may repress genes encoding secreted enzymes for nutrient assimilation and instead recycle proteins to escape detection of the host immune system.

One of our investigation highlights was identifying a glucose oxidase (GOx) (g5302_chr16_Ss) up-regulated in the infection of resistant plants. GOx (EC:1.1.3.4) is an oxidoreductase that catalyzes glucose oxidation to hydrogen peroxide and D-glucono-δ-lactone, successfully used in bioelectronic devices and biosensors to detect free glucose (Kornecki *et al.*, 2020). The induction of GOx determined in the *Penicillium expansum* infecting apples was essential for the fungus' necrotrophic development and, consequently, for pathogenicity (Hadas *et al.*, 2007). On the other hand, transgenic rice, potato, tobacco, and canola plants expressing a fungal glucose oxidase gene conferred long-lasting resistance to various pathogens (Wu Gusui *et al.*, 1995; Kazan *et al.*, 1998; Felcher *et al.*, 2003; Kachroo *et al.*, 2003). GOx led to increased endogenous levels of H_2O_2, which activated the expression of several defense genes. GOx was also considered a control agent against *Botrytis cinerea* infecting strawberries (Li *et al.*, 2019). Indirectly, GOx controlled *Fusarium solani* using biocontrol strategies with *Aspergillus tubingensis* CTM 507 expressing glucose oxidase (Kriaa *et al.*, 2015). Experimental data proved that GOx is involved in lesion formation and expansion of plant cell walls (Govrin and Levine, 2000), and in the smut model *U. maydis*, glyoxal oxidases play a role in the filamentous growth and pathogenicity (Leuthner *et al.*, 2005). For *S. scitamineum* infecting sugarcane, as mentioned before, Peters *et al.* (2017) showed increased H_2O_2 in resistant plants not detected during the infection of susceptible plants. An assay of GOx using transgenic approaches may bring new light in the control of smut disease.

Furthermore, this work unveiled other representative genes during compatible and incompatible interactions differentiating infection in resistant and susceptible sugarcane plants. For instance, we identified genes encoding heat shock proteins related to stress conditions such as those involved in host infection and colonization, heat and dry conditions, among other factors (Hahn *et al.*, 2004). There are still many proteins of uncharacterized function due to the lack of experimental data. Functional studies are long due to confirm the correct sequence (start, end, presence of introns and regulatory regions) and the role of the many proteins identified in this work.

4.3 *In planta* vs *in vitro* comparison

The use of previously obtained data of the fungus growing in the axenic culture helped us uncover candidates involved in the filamentous growth (discussed later), responses to the variant environment of resistant and susceptible hosts, and, most relevant, the genes encoding effector candidates. In general, we observed the oxidation-reduction process GO term mostly represented in genes with high expression in infected plants. Deviations from the normal redox values within cells are usually associated with energy production, filamentous development and cell wall modification, responses to the plant immune system, and other cell maintenance processes (Breitenbach *et al.*, 2015). The category KOG of carbohydrate metabolism and transport harbored the larger number of genes induced in the plant, after no-hit, and unknown function (S). It seems reasonable to assume that the fungus will change carbon assimilation in the plant compared to the available carbon sources present in axenic growth.

We already addressed some of the differences detected in compatible and incompatible interactions. Studies performed in *U. maydis* revealed modification in carbon acquisition during pathogenic development in plants (Goulet and Saville, 2017). As discussed before, we identified a large set of hydrolases acting on both the fungal cell wall to allow the filamentous growth and host cell wall degradation. Among the enzymes associated with host cell wall degradation were g600_chr19_Ss; g3696_chr19_Ss; g5316_chr16_Ss; g1300_chr03_Ss; g3529_chr08_Ss, encoding orthologs of cellulases, hemicellulases, and pectinases.

When searching for the glucose oxidase g5302_chr16_Ss, predicted as up-regulated in the comparison with the fungus growing in the two contrasting genotypes, we detected the same pattern when comparing the *S. scitamineum* growth *in planta* with *in vitro*, this gene was considered DEG in both comparisons, being up-regulated in the SP80-3280 and down-regulated in the IAC66-6. Other two glucose oxidases were found in the *in vitro* comparisons: g5805_chr18_Ss, being down-regulated in both comparisons, and g4780_chr13_Ss, being up-regulated in the IAC66-6, this one was not detected as expressed in the SP80-3280 (*i.e.*, have at least one biological replicate with less than 1 CPM).

4.4 Candidate effector genes

Fungal pathogens avoid triggering the plant host's innate immune system using different effector's repertoire. Patterns of temporal expression are a hallmark of candidate effectors to ensure fungal infection successfully and host plant colonization (Toruño *et al.*, 2016; Tang *et*

al., 2018). They express according to the fungal lifestyle in waves to ensure, for instance, its biotrophic behavior undermining host defense responses. Our data support the statement and add information about variation considering the host plant's genetic background. Resistant and susceptible host induced the expression of different sets of effectors at 48 hai. Teixeira-Silva (2018) described similar results analyzing 12 candidate effector genes expressed in experiments using the same sugarcane genotypes and fungal isolate.

Considering functionally characterized effectors in other smut fungi, we detected two DEGs, Fly1 and Cmu1. The Fly1 found up-regulated when comparing the resistant genotype infection with the *in vitro* growth is a protective effector in *U. maydis* encoding a secreted lysin metalloprotease that cleaves maize chitinase-A (Ökmen *et al.*, 2018). The other candidate effector also up-regulated in the same comparison is the Cmu1 ortholog. The Cmu1 acts in reprogramming plant metabolic pathways when translocates into the plants host cell (Djamei *et al.*, 2011; Tanaka *et al.*, 2014). It prevents the Salicylic Acid production, which is the primary defense hormone, sequestring its biosynthesis precursor chorismate (Djamei *et al.*, 2011; Djamei and Kahmann, 2012).

The other two effectors identified in our work protecting smut pathogens are the Rsp3 and Pit2. They both induced in the plant were more expressive in the susceptible genotype. The Rsp3 protects *U. maydis* mycelium from the activity of the maize mannose-binding proteins AFP1 and AFP2 by binding and shielding the fungal cell wall (Ma *et al.*, 2018). The Pit2 inhibits maize papain-like cysteine proteases, suppressing the apoplastic host defense responses, crucial for infection and virulence of *U. maydis* (Doehlemann *et al.*, 2011; Mueller *et al.*, 2013). These two effectors also play essential roles at the early stages of the sugarcane-smut infection, expressed by the pathogen to protects itself from the defense barriers imposed by the plant host.

4.5 Mating and filamentation growth-related genes

The analysis of the mating/filamentation genes showed that although the low in-depth of our RNA-Seq data, we evidenced differential expression of genes necessary for the first events mating-related in plant infection. The sexual reproduction and the filament development of dikaryotic hyphae are essential in the successful infection and colonization of the host plant by *S. scitamineum*. Some genes characterized in *S. scitamineum* are of significant importance to control mating/filamentation. For instance, a knock-out mutant sskpp2Δ in the MAPK SsKpp2 significantly reduced mating/filamentation in *S. scitamineum* (Deng *et al.*, 2018). The

authors found that the SsKpp2 is required for mating/filamentation in *S. scitamineum* via regulation of the pheromone signal transduction involving cAMP, and tryptophol biosynthesis, which is a quorum-sensing molecule. Later, other knock-out mutants (ssgpa3Δ, ssuac1Δ, and ssadr1Δ) of the cAMP/PKA pathway proved to be defective in mating/filamentation (Chang *et al.*, 2019). The final results concluded that cAMP positively regulated *S. scitamineum* mating/filamentation, via the cAMP/PKA pathway upstream of the transcription factor SsPrf1 (Chang *et al.*, 2019). The authors also revealed that increased peroxide suppressed filamentation, considering that the ROS levels positively signal mating/filamentation in *S. scitamineum.* We compare these genes and others (Martínez-Soto *et al.*, 2020) involved in the mating/filamentation development and growth (Figure 12). We identified all the genes related to mating/filamentation, functionally characterized by Deng *et al.* (2018) and Chang *et al.* (2019), in all three treatments, except SsAdr1 in susceptible plants. The MAPK Kpp2 detected up-regulated in both genotypes indicates its relevance *in planta* colonization. The gene SsGpa3, of the signaling pathway cAMP/PKA, and encoding a G-protein associated with mating signals' reception, was up-regulated in the SP80-3280. The other genes not detected were mostly because of our experiment's low in-depth coverage, as pointed before.

5. Conclusions

The sugarcane smut pathogen (*S. scitamineum*) uses different strategies when infecting the two contrasting genotypes SP80-3280 (resistant) and IAC66-6 (susceptible). *S. scitamineum* colonizing resistant plants, which has an early response identified by an oxidative burst, induce antioxidant enzymes' expression to escape the ROS's toxic environment to colonize plant tissues. The host defense response is potentially associated with the perception of proteases produced by the fungus triggering the immune system. On the other hand, *S. scitamineum* infecting susceptible plants supposedly uses a set of effectors to inhibit plant defenses and colonizing the host tissues.

This work provided insights into the expression profile to differentiate colonization of resistant and susceptible genotypes by sugarcane smut pathogen – a guide of candidate genes for further functional characterization studies. Functional studies are essential to comprehend the pathosystem sugarcane-smut and the fungal mechanisms to successfully colonize both genotypes and only complete its cycle in susceptible plants. We propose studies related to the characterization of the gene g5302_chr16_Ss, which is related to glucose oxidase and may play a role in how the fungus can colonize the resistant genotype.

REFERENCES

Akalach, M., and Touil, B. (1996). Occurrence and Spread of Sugarcane Smut Caused by Ustilago scitaminea in Morocco. *Plant Dis.* 80, 1363–1366.

Ali, S., Laurie, J. D., Linning, R., Cervantes-Chávez, J. A., Gaudet, D., and Bakkeren, G. (2014). An Immunity-Triggering Effector from the Barley Smut Fungus Ustilago hordei Resides in an Ustilaginaceae-Specific Cluster Bearing Signs of Transposable Element-Assisted Evolution. *PLoS Pathog.* 10, e1004223. doi:10.1371/journal.ppat.1004223.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ballouz, S., Pavlidis, P., and Gillis, J. (2017). Using predictive specificity to determine when gene set analysis is biologically meaningful. *Nucleic Acids Res.* 45, 1–18. doi:10.1093/nar/gkw957.

Benevenuto, J., Teixeira-Silva, N. S., Kuramae, E. E., Croll, D., and Monteiro-Vitorello, C. B. (2018). Comparative Genomics of Smut Pathogens: Insights From Orphans and Positively Selected Genes Into Host Specialization. *Front. Microbiol.* 9, 1–17. doi:10.3389/fmicb.2018.00660.

Brefort, T., Tanaka, S., Neidig, N., Doehlemann, G., Vincon, V., and Kahmann, R. (2014). Characterization of the Largest Effector Gene Cluster of Ustilago maydis. *PLoS Pathog.* 10, e1003866. doi:10.1371/journal.ppat.1003866.

Breitenbach, M., Weber, M., Rinnerthaler, M., Karl, T., and Breitenbach-Koller, L. (2015). Oxidative stress in fungi: Its function in signal transduction, interaction with plant hosts, and lignocellulose degradation. *Biomolecules* 5, 318–342. doi:10.3390/biom5020318.

Carvalho, G., Quecine, M. C., Longatto, D. P., Peters, L. P., Almeida, J. R., Shyton, T. G., *et al.* (2016). *Sporisorium scitamineum* colonisation of sugarcane genotypes susceptible and resistant to smut revealed by GFP-tagged strains. *Ann. Appl. Biol.* 169, 329–341. doi:10.1111/aab.12304.

Chang, C., Cai, E., Deng, Y. Z., Mei, D., Qiu, S., Chen, B., *et al.* (2019). cAMP/PKA signalling pathway regulates redox homeostasis essential for *Sporisorium scitamineum* mating/filamentation and virulence. *Environ. Microbiol.* 21, 959–971. doi:10.1111/1462-2920.14496.

Colman-Lerner, A., Chin, T. E., and Brent, R. (2001). Yeast Cbk1 and Mob2 activate daughter-specific genetic programs to induce asymmetric cell fates. *Cell* 107, 739–750. doi:10.1016/S0092-8674(01)00596-7.

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi:10.1093/bioinformatics/bti610.

Croft, B. J., and Braithwaite, K. S. (2006). Management of an incursion of sugarcane smut in Australia. *Australas. Plant Pathol.* 35, 113–122. doi:10.1071/AP05104.

Deng, Y. Z., Zhang, B., Chang, C., Wang, Y., Lu, S., Sun, S., *et al.* (2018). The MAP Kinase SsKpp2 Is Required for Mating/Filamentation in Sporisorium scitamineum. *Front. Microbiol.* 9, 2555. doi:10.3389/fmicb.2018.02555.

Djamei, A., and Kahmann, R. (2012). Ustilago maydis: Dissecting the Molecular Interface between Pathogen and Plant. *PLoS Pathog.* 8, e1002955. doi:10.1371/journal.ppat.1002955.

Djamei, A., Schipper, K., Rabe, F., Ghosh, A., Vincon, V., Kahnt, J., *et al.* (2011). Metabolic priming by a secreted fungal effector. *Nature* 478, 395–398. doi:10.1038/nature10454.

Doehlemann, G., Reissmann, S., Aßmann, D., Fleckenstein, M., and Kahmann, R. (2011). Two linked genes encoding a secreted effector and a membrane protein are essential for Ustilago maydis-induced tumour formation. *Mol. Microbiol.* 81, 751–766. doi:10.1111/j.1365-2958.2011.07728.x.

Doehlemann, G., Van Der Linde, K., Aßmann, D., Schwammbach, D., Hof, A., Mohanty, A., *et al.* (2009). Pep1, a secreted effector protein of Ustilago maydis, is required for successful invasion of plant cells. *PLoS Pathog.* 5, e1000290. doi:10.1371/journal.ppat.1000290.

Elías-Villalobos, A., Barrales, R. R., and Ibeas, J. I. (2019). Chromatin modification factors in plant pathogenic fungi: Insights from Ustilago maydis. *Fungal Genet. Biol.* 129, 52–64. doi:10.1016/j.fgb.2019.04.006.

Ene, I. V., Brunke, S., Brown, A. J. P., and Hube, B. (2014). Metabolism in fungal pathogenesis. *Cold Spring Harb. Perspect. Med.* 4, a019695. doi:10.1101/cshperspect.a019695.

FAO. FAOSTAT. Available at: <http://www.fao.org/faostat/en/#data/QC>. Access date: 20/04/2020.

Felcher, K. J., Douches, D. S., Kirk, W. W., Hammerschmidt, R., and Li, W. (2003). Expression of a fungal glucose oxidase gene in three potato cultivars with different susceptibility to late blight (Phytophthora infestans Mont. deBary). *J. Am. Soc. Hortic. Sci.* 128, 238–245. doi:10.21273/jashs.128.2.0238.

Figueiredo, A., Monteiro, F., and Sebastiana, M. (2014). Subtilisin-like proteases in plant-pathogen recognition and immune priming: a perspective. *Front. Plant Sci.* 5, 739. doi:10.3389/fpls.2014.00739.

Figueiredo, J., Sousa Silva, M., and Figueiredo, A. (2018). Subtilisin-like proteases in plant defence: the past, the present and beyond. *Mol. Plant Pathol.* 19, 1017–1028. doi:10.1111/mpp.12567.

Goulet, K. M., and Saville, B. J. (2017). Carbon acquisition and metabolism changes during fungal biotrophic plant pathogenesis: insights from Ustilago maydis. *Can. J. Plant Pathol.* 39, 247–266. doi:10.1080/07060661.2017.1354330.

Govrin, E. M., and Levine, A. (2000). The hypersensitive response facilitates plant infection by the necrotrophic pathogen Botrytis cinerea. *Curr. Biol.* 10, 751–757. doi:10.1016/S0960-9822(00)00560-1.

Hadas, Y., Goldberg, I., Pines, O., and Prusky, D. (2007). Involvement of gluconic acid and glucose oxidase in the pathogenicity of Penicillium expansum in apples. *Phytopathology* 97, 384–390. doi:10.1094/PHYTO-97-3-0384.

Hahn, J.-S., Hu, Z., Thiele, D. J., and Iyer, V. R. (2004). Genome-wide analysis of the biology of stress responses through heat shock transcription factor. *Mol. Cell. Biol.* 24, 5249–56. doi:10.1128/MCB.24.12.5249-5256.2004.

Hemetsberger, C., Herrberger, C., Zechmann, B., Hillmer, M., and Doehlemann, G. (2012). The Ustilago maydis effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. *PLoS Pathog.* 8, e1002684. doi:10.1371/journal.ppat.1002684.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., *et al.* (2019). EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi:10.1093/nar/gky1085.

Kachroo, A., He, Z., Patkar, R., Zhu, Q., Zhong, J., Li, D., *et al.* (2003). Induction of H2O2 in transgenic rice leads to cell death and enhanced resistance to both bacterial and fungal pathogens. *Transgenic Res.* 12, 577–586. doi:10.1023/A:1025896513472.

Kämper, J., Kahmann, R., Bölker, M., Ma, L.-J., Brefort, T., Saville, B. J., *et al.* (2006). Insights from the genome of the biotrophic fungal plant pathogen Ustilago maydis. *Nature* 444, 97–101. doi:10.1038/nature05248.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45. doi:10.1093/nar/gkw1092.

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi:10.1093/nar/28.1.27.

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* 428, 726–731. doi:10.1016/J.JMB.2015.11.006.

Kawahara, Y., Oono, Y., Kanamori, H., Matsumoto, T., Itoh, T., and Minami, E. (2012). Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PLoS One* 7, e49423. doi:10.1371/journal.pone.0049423.

Kazan, K., Murray, F. R., Goulter, K. C., Llewellyn, D. J., and Manners, J. M. (1998). Induction of cell death in transgenic plants expressing a fungal glucose oxidase. *Mol. Plant-Microbe Interact.* 11, 555–562. doi:10.1094/MPMI.1998.11.6.555.

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317.

König, J., Baumann, S., Koepke, J., Pohlmann, T., Zarnack, K., and Feldbrügge, M. (2009). The fungal RNA-binding protein Rrm4 mediates long-distance transport of ubi1 and rho3 mRNAs. *EMBO J.* 28, 1855–1866. doi:10.1038/emboj.2009.145.

Kornecki, J. F., Carballares, D., Tardioli, P. W., Rodrigues, R. C., Berenguer-Murcia, Á., Alcántara, A. R., *et al.* (2020). Enzyme production of d -gluconic acid and glucose oxidase: successful tales of cascade reactions. *Catal. Sci. Technol.* 10, 5740–5771. doi:10.1039/d0cy00819b.

Kriaa, M., Hammami, I., Sahnoun, M., Azebou, M. C., Triki, M. A., and Kammoun, R. (2015). Biocontrol of tomato plant diseases caused by Fusarium solani using a new isolated Aspergillus tubingensis CTM 507 glucose oxidase. *Comptes Rendus - Biol.* 338, 666–677. doi:10.1016/j.crvi.2015.05.007.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi:10.1006/JMBI.2000.4315.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., *et al.* (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi:10.1101/gr.092759.109.

Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O., and Filipovska, A. (2019). CirGO: An alternative circular way of visualising gene ontology terms. *BMC Bioinformatics* 20, 84. doi:10.1186/s12859-019-2671-2.

Latiza, A. S., Ampusta, D. C., and Rivera, J. R. (1980). Reaction of sugarcane clones to strain B of Ustilago scitaminea Sydow. in *Proceeding ISSCT*, 1456–1462.

Leach, M. D., Stead, D. A., Argo, E., Maccallum, D. M., and Brown, A. J. P. (2011). Molecular and proteomic analyses highlight the importance of ubiquitination for the stress resistance, metabolic adaptation, morphogenetic regulation and virulence of Candida albicans. *Mol. Microbiol.* 79, 1574–1593. doi:10.1111/j.1365-2958.2011.07542.x.

Lee, H. K., Braynen, W., Keshav, K., and Pavlidis, P. (2005). ErmineJ: Tool for functional analysis of gene expression data sets. *BMC Bioinformatics* 6, 269. doi:10.1186/1471-2105-6-269.

Lemma, A., Hagos, H., and Zekarias, Y. Tekle, A. (2015). Study on the Reaction of Sugarcane Genotypes (CIRAD-2011) to Sugarcane Smut (Sporisorium scitamineum) in the Ethiopian Sugarcane Plantations. *Adv. Crop Sci. Technol.* 3, 1–9. doi:10.4172/2329-8863.1000181.

Leuthner, B., Aichinger, A. C., Oehmen, A. E., Koopmann, E., Muïler, A. O., Muüler, A. P., *et al.* (2005). A H 2 O 2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in Ustilago maydis. *Mol Gen Genomics* 272, 650. doi:10.1007/s00438-004-1085-6.

Li, X., Xie, X., Xing, F., Xu, L., Zhang, J., and Wang, Z. (2019). Glucose oxidase as a control agent against the fungal pathogen Botrytis cinerea in postharvest strawberry. *Food Control* 105, 277–284. doi:10.1016/j.foodcont.2019.05.037.

Liang, L. (2012). The role of Stp1, a secreted effector, in the biotrophic interaction of Ustilago maydis and its host plant maize.

Liao, Y., Smyth, G. K., and Shi, W. (2013). The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 41. doi:10.1093/nar/gkt214.

Liao, Y., Smyth, G. K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi:10.1093/bioinformatics/btt656.

Linning, R., Lin, D., Lee, N., Abdennadher, M., Gaudet, D., Thomas, P., *et al.* (2004). Marker-Based Cloning of the Region Containing the UhAvr1 Avirulence Gene. *Genetics* 166, 99–111.

Liu, T. B., and Xue, C. (2011). The ubiquitin-proteasome system and F-box proteins in pathogenic fungi. *Mycobiology* 39, 243–248. doi:10.5941/MYCO.2011.39.4.243.

Ma, L.-S. S., Wang, L., Trippel, C., Mendoza-Mendoza, A., Ullmann, S., Moretti, M., *et al.* (2018). The Ustilago maydis repetitive effector Rsp3 blocks the antifungal activity of mannose-binding maize proteins. *Nat. Commun.* 9, 1–15. doi:10.1038/s41467-018-04149-0.

Mandujano-González, V., Villa-Tanaca, L., Anducho-Reyes, M. A., and Mercado-Flores, Y. (2016). Aspartil-proteasas secretadas por hongos: revisión. *Rev. Iberoam. Micol.* 33, 76–82. doi:10.1016/j.riam.2015.10.003.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10. doi:10.14806/ej.17.1.200.

Martínez-Soto, D., González-Prieto, J. M., and Ruiz-Herrera, J. (2015). Transcriptomic analysis of the GCN5 gene reveals mechanisms of the epigenetic regulation of virulence and morphogenesis in Ustilago maydis. *FEMS Yeast Res.* 15, fov055. doi:10.1093/femsyr/fov055.

Martínez-Soto, D., Ortiz-Castellanos, L., Robledo-Briones, M., and León-Ramírez, C. G. (2020). Molecular Mechanisms Involved in the Multicellular Growth of Ustilaginomycetes. *Microorganisms* 8, 1072. doi:10.3390/microorganisms8071072.

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi:10.1093/nar/gks042.

Mueller, A. N., Ziemann, S., Treitschke, S., Aßmann, D., and Doehlemann, G. (2013). Compatibility in the Ustilago maydis-Maize Interaction Requires Inhibition of Host Cysteine Proteases by the Fungal Effector Pit2. *PLoS Pathog.* 9, e1003177. doi:10.1371/journal.ppat.1003177.

Muszewska, A., Stepniewska-Dziubinska, M. M., Steczkiewicz, K., Pawlowska, J., Dziedzic, A., and Ginalski, K. (2017). Fungal lifestyle reflected in serine protease repertoire. *Sci. Rep.* 7, 1–12. doi:10.1038/s41598-017-09644-w.

Naidoo, S., Visser, E. A., Zwart, L., Toit, Y. du, Bhadauria, V., and Shuey, L. S. (2018). Dual RNA-Sequencing to Elucidate the Plant-Pathogen Duel. *Curr. Issues Mol. Biol.* 27, 127–142. doi:10.21775/cimb.027.127.

Ökmen, B., Kemmerich, B., Hilbig, D., Wemhöner, R., Aschenbroich, J., Perrar, A., *et al.* (2018). Dual function of a secreted fungalysin metalloprotease in Ustilago maydis. *New Phytol.* 220, 249–261. doi:10.1111/nph.15265.

Parzych, K. R., Ariosa, A., Mari, M., and Klionsky, D. J. (2018). A newly characterized vacuolar serine carboxypeptidase, Atg42/Ybr139w, is required for normal vacuole function and the terminal steps of autophagy in the yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 29, 1089–1099. doi:10.1091/mbc.E17-08-0516.

Peters, L. P. (2016). A more detailed view of reactive oxygen species metabolism in the sugarcane and Sporisorium scitamineum interaction. Available at: https://www.teses.usp.br/teses/disponiveis/11/11137/tde-23012017-094942/en.php.

Peters, L. P., Carvalho, G., Vilhena, M. B., Creste, S., Azevedo, R. A., and Monteiro-Vitorello, C. B. (2017). Functional analysis of oxidative burst in sugarcane smut-resistant and -susceptible genotypes. *Planta* 245, 749–764. doi:10.1007/s00425-016-2642-z.

Peters, L.P.; Teixeira-Silva, N.S.; Bini, A.P.; Silva, M.M.L.; Crestana, G.S.; Azevedo, R.A.; Carvalho, G.; Monteiro-Vitorello, C.B. (2020) Differential responses of genes and enzymes associated with ROS protective responses in the sugarcane smut fungus. [Manuscript accepted for publication, in-press].

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi:10.1038/nmeth.1701.

Piepenbring, M., Stoll, M., and Oberwinkler, F. (2002). The generic position of Ustilago maydis, Ustilago scitaminea, and Ustilago esculenta (Ustilaginales). *Mycol. Prog.* 1, 71–80. doi:10.1007/s11557-006-0006-y.

Pierleoni, A., Martelli, P., and Casadio, R. (2008). PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* 9, 392. doi:10.1186/1471-2105-9-392.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richards, A., Gow, N. A. R., and Veses, V. (2012). Identification of vacuole defects in fungi. *J. Microbiol. Methods* 91, 155–163. doi:10.1016/j.mimet.2012.08.002.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp616.

Schaker, P. D. C., Palhares, A. C., Taniguti, L. M., Peters, L. P., Creste, S., Aitken, K. S., *et al.* (2016). RNAseq Transcriptional Profiling following Whip Development in Sugarcane Smut Disease. *PLoS One* 11, e0162237. doi:10.1371/journal.pone.0162237.

Schipper, K. (2009). Charakterisierung eines Ustilago maydis Genclusters, das für drei neuartige sekretierte Effektoren kodiert. Available at: https://archiv.ub.uni-marburg.de/diss/z2009/0723/.

Shaw, J. A., Mol, P. C., Bowers, B., Silverman, S. J., Valdivieso, M. H., Duran, A., *et al.* (1991). The function of chitin synthases 2 and 3 in the Saccharomyces cerevisiae cell cycle. *J. Cell Biol.* 114, 111–123. doi:10.1083/jcb.114.1.111.

Smith, D. G., Garcia-Pedrajas, M. D., Gold, S. E., and Perlin, M. H. (2003). Isolation and characterization from pathogenic fungi of genes encoding ammonium permeases and their roles in dimorphism. *Mol. Microbiol.* 50, 259–275. doi:10.1046/j.1365-2958.2003.03680.x.

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Singh, K. B., and Taylor, J. M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.* 19, 2094–2110. doi:10.1111/mpp.12682.

Stirnberg, A., and Djamei, A. (2016). Characterization of ApB73, a virulence factor important for colonization of Zea mays by the smut Ustilago maydis. *Mol. Plant Pathol.* 17, 1467–1479. doi:10.1111/mpp.12442.

Sundar, A. R., Barnabas, E. L., Malathi, P., and Viswanathan, R. (2012). A mini-review on smut disease of sugarcane caused by Sporisorium scitamineum. *InTech Bot.*, 107–128.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One* 6, e21800. doi:10.1371/journal.pone.0021800.

Tanaka, S., Brefort, T., Neidig, N., Djamei, A., Kahnt, J., Vermerris, W., *et al.* (2014). A secreted Ustilago maydis effector promotes virulence by targeting anthocyanin biosynthesis in maize. *Elife* 3, 1355. doi:10.7554/eLife.01355.001.

Tang, C., Xu, Q., Zhao, M., Wang, X., and Kang, Z. (2018). Understanding the lifestyles and pathogenicity mechanisms of obligate biotrophic fungi in wheat: The emerging genomics era. *Crop J.* 6, 60–67. doi:10.1016/j.cj.2017.11.003.

Taniguti, L. M., Schaker, P. D. C., Benevenuto, J., Peters, L. P., Carvalho, G., Palhares, A., *et al.* (2015). Complete Genome Sequence of Sporisorium scitamineum and Biotrophic Interaction Transcriptome with Sugarcane. *PLoS One* 10, e0129318. doi:10.1371/journal.pone.0129318.

Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., *et al.* (2003). The COG database: An updated vesion includes eukaryotes. *BMC Bioinformatics* 4. doi:10.1186/1471-2105-4-41.

Teixeira-Silva, N. S. (2018). Functional analysis of candidate effector proteins during Sporisorium scitamineum x sugarcane interaction.

Toruño, T. Y., Stergiopoulos, I., and Coaker, G. (2016). Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. *Annu. Rev. Phytopathol.* 54, 419–441. doi:10.1146/annurev-phyto-080615-100204.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., *et al.* (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi:10.1038/nbt.1621.

Wu Gusui, Shortt, B. J., Lawrence, E. B., Levine, E. B., Fitzsimmons, K. C., and Shah, D. M. (1995). Disease resistance conferred by expression of a gene encoding H2O2-generating glucose oxidase in transgenic potato plants. *Plant Cell* 7, 1357–1368. doi:10.1105/tpc.7.9.1357.

Zekarias, Y., Dejene, M., Tegegn, G., and Yirefu, F. (2010). Ethiopian journal of agricultural sciences EJAS. *Ethiop. J. Agric. Sci.* 20, 35–46.

Zhao, D., and Li, Y.-R. (2015). Climate Change and Sugarcane Production: Potential Impact and Mitigation Strategies. *Int. J. Agron.*, 1–10. doi:10.1155/2015/547386.

Zhao, Z., Liu, H., Wang, C., and Xu, J. R. (2013). Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 14, 1–15. doi:10.1186/1471-2164-14-274.

Zhu, G., Deng, Y., Cai, E., Yan, M., Cui, G., Wang, Z., *et al.* (2019). Identification and Functional Analysis of the Pheromone Response Factor Gene of Sporisorium scitamineum. *Front. Microbiol.* 10, 2115. doi:10.3389/fmicb.2019.02115.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Differentially expressed genes of *S. scitamineum* in the comparison resistant vs susceptible genotypes.

Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g520_chr01_Ss	2.68	3.78224908522736E-05	probable Hmp1-Mismatch base pair and cruciform DNA recognition protein	
g1713_chr03_Ss	2.63	0.000172049214709	probable Alcohol dehydrogenase	
g2447_chr05_Ss	2.26	0.001237164962945	probable DNA polymerase X-putative	
g5222_chr15_Ss	2.24	0.001800132158594	conserved hypothetical protein	
g6154_chr20_Ss	2.19	0.001621474050808	Ste20-like kinase Don3	
g1647_chr03_Ss	2.11	0.002397523226774	conserved hypothetical protein	X
g5200_chr15_Ss	2.09	0.000707496402191	related to sepB protein	
g1926_chr04_Ss	1.98	0.001385132405335	conserved hypothetical protein	X
g3602_chr09_Ss	1.95	0.002140492440631	probable phosphoenolpyruvate carboxykinase	
g4881_chr14_Ss	1.87	0.008028728641995	conserved hypothetical protein	
g1672_chr03_Ss	1.86	0.002774553988194	related to carnitine acetyl transferase FacC	
g2823_chr06_Ss	1.85	0.010464581174604	related to Endothelin-converting enzyme 1	
g5302_chr16_Ss	1.84	0.009319729856394	related to Glucose oxidase	X
g3823_chr10_Ss	1.77	0.003312749954399	probable heat shock protein 80	
g6546_chr22_Ss	1.76	0.012348234685065	probable IDP1-isocitrate dehydrogenase (NADP+). mitochondrial	

Continue of Supplementary Table 1…

Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g2050_chr04_Ss	1.75	0.02004307231915	conserved hypothetical protein	
g5951_chr19_Ss	1.69	0.011749276476472	conserved hypothetical protein	
g4535_chr12_Ss	1.68	0.011936006984778	related to Cholinesterase precursor	
g5932_chr19_Ss	1.66	0.015489386142915	related to TIF4631-mRNA cap-binding protein (eIF4F). 150K subunit (C-terminal fragment)	
g2833_chr06_Ss	1.65	0.016305781137879	probable FBP1-fructose-1.6-bisphosphatase	
g4339_chr11_Ss	1.61	0.024411755977046	probable STI1-Hsp90 cochaperone	
g3063_chr07_Ss	1.58	0.025520435047161	probable Aldo-keto reductase yakc	
g506_chr01_Ss	1.57	0.019445938090109	probable Serine/threonine-protein kinase gad8	
g5190_chr15_Ss	1.54	0.018560509629909	related to PIN4-protein involved in G2/M phase progression and response to DNA damage	
g779_chr02_Ss	1.53	0.012253358455521	related to DDR48-heat shock protein	
g2922_chr06_Ss	1.52	0.045425265791138	related to UV-damaged DNA-binding protein	
g94_chr01_Ss	1.52	0.025931069517799	conserved hypothetical protein	
g2859_chr06_Ss	1.52	0.02221853746736	probable YDJ1-mitochondrial and ER import protein	

Continue of Supplementary Table 1…

Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g3801_chr09_Ss	1.48	0.02702168604117	probable 2-methylcitrate dehydratase	
g4614_chr13_Ss	1.47	0.025970538376989	probable catalase 2	
g5788_chr18_Ss	1.46	0.025059578406517	probable enoyl-CoA hydratase precursor. mitochondrial	
g5644_chr17_Ss	1.45	0.037340388943937	related to PRC1-carboxypeptidase y. serine-type protease	
g388_chr01_Ss	1.45	0.030987747504783	conserved hypothetical protein	
g487_chr01_Ss	1.44	0.037054598780999	conserved hypothetical protein	
g6071_chr20_Ss	1.44	0.036839853306715	related to Long-chain-fatty-acid--CoA ligase 6	
g3209_chr07_Ss	1.43	0.016363844932045	fasciclin domain protein 3	
g74_chr01_Ss	1.42	0.043808533444278	related to pepsin precursor (aspartate protease)	X
g2907_chr06_Ss	1.41	0.033747353984028	probable nucleoside-diphosphate kinase	
g3013_chr07_Ss	1.40	0.045670830242125	related to YBT1-Vacuolar. ABC protein transporting bile acids	
g1075_chr02_Ss	1.40	0.026391484345633	peroxisomal catalase protein	
g1450_chr03_Ss	1.38	0.034148678695381	conserved hypothetical protein	X
g165_chr01_Ss	1.37	0.034905516564713	probable multifunctional beta-oxidation protein	
g2951_chr07_Ss	1.36	0.046289987311726	related to Aminoadipate-semialdehyde dehydrogenase	

Continue of Supplementary Table 1…

Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g5122_chr15_Ss	1.36	0.040384011424318	conserved hypothetical protein	
g2956_chr07_Ss	1.33	0.043135820513835	probable ribose-5-phosphate isomerase	
g6136_chr20_Ss	1.31	0.028770070808908	conserved hypothetical protein	X
g2632_chr06_Ss	1.26	0.034721909086741	indole-3-acetaldehyde dehydrogenase	
g1875_chr04_Ss	1.25	0.047462709179895	conserved hypothetical protein	
g5426_chr16_Ss	1.19	0.049632357000379	repellent protein 1 precursor	X
g5434_chr16_Ss	-1.14	0.044793872442972	histone H3	
g3464_chr08_Ss	-1.20	0.04833762712549	related to SEC59-Dolichol kinase	
g1523_chr03_Ss	-1.23	0.047040205219674	conserved hypothetical protein	
g4047_chr10_Ss	-1.25	0.036338832790609	conserved hypothetical protein	
g3992_chr10_Ss	-1.26	0.042391247678169	related to Protein mlo2	
g4454_chr12_Ss	-1.29	0.032435667664443	probable COP1-coatomer complex alpha chain of secretory pathway vesicles	
g4625_chr13_Ss	-1.30	0.032080003977392	conserved hypothetical protein	
g5093_chr15_Ss	-1.31	0.041470800491601	conserved hypothetical protein	
g4867_chr14_Ss	-1.31	0.041533231345816	related to succinate dehydrogenase [ubiquinone] cytochrome b small subunit. mitochondrial precursor	

Continue of Supplementary Table 1…

Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g75_chr01_Ss	-1.32	0.030610065061449	related to HMF1-Heat-shock induceable Inhibitor of cell Growth	
g5570_chr17_Ss	-1.33	0.02568534330685	probable MMS2-part of the error-free postreplication repair pathway	
g4685_chr13_Ss	-1.34	0.024911109333211	related to Histone H1	
g5332_chr16_Ss	-1.34	0.019468081362895	conserved hypothetical protein	
g6196_chr20_Ss	-1.35	0.044878571474764	related to hydroxyquinol-1.2-dioxygenase	
g403_chr01_Ss	-1.36	0.02679859351423	conserved hypothetical protein	
g2062_chr04_Ss	-1.40	0.018633929426712	probable cyclophilin b	
g411_chr01_Ss	-1.41	0.048175874039121	conserved hypothetical protein	
g55_chr01_Ss	-1.41	0.048166784491774	related to L-serine dehydratase 1	X
g5571_chr17_Ss	-1.42	0.018248993448677	putative dioxygenase Ssp1	
g93_chr01_Ss	-1.49	0.02385046533151	related to 5-oxoprolinase	
g5705_chr18_Ss	-1.49	0.014985525843682	conserved hypothetical protein	
g2845_chr06_Ss	-1.51	0.008404013298225	histone-fold-containing protein	
g5410_chr16_Ss	-1.53	0.014566853091292	conserved hypothetical protein	
g2310_chr05_Ss	-1.56	0.02493454042306	SEC27-coatomer complex beta subunit	
g5636_chr17_Ss	-1.59	0.008162009366951	RPL15A 60S large subunit ribosomal protein L15.e	

Continue of Supplementary Table 1…

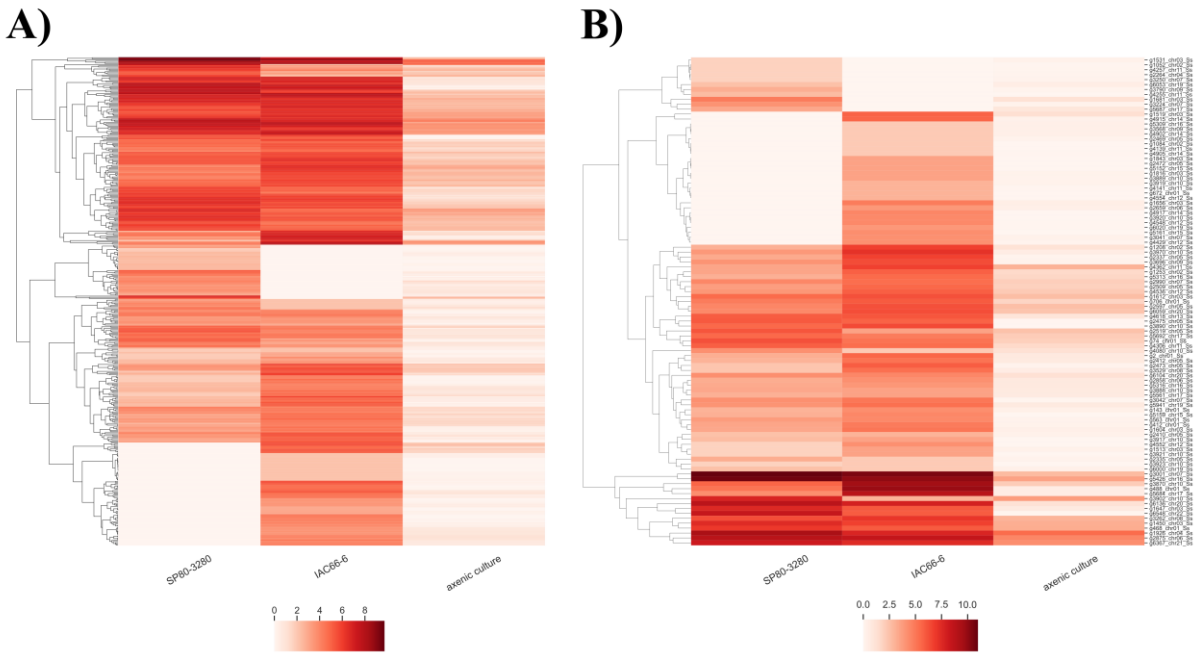
Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g5947_chr19_Ss	-1.61	0.012428699206305	probable POB3-protein that binds to DNA polymerase I	
g277_chr01_Ss	-1.62	0.009484875949404	conserved hypothetical protein	
g6100_chr20_Ss	-1.65	0.025555603394672	conserved hypothetical protein	
g5105_chr15_Ss	-1.65	0.025554186165774	related to SIZ1-E3-like factor in the SUMO pathway	
g1217_chr02_Ss	-1.65	0.025547893540859	probable TYS1-tyrosyl-tRNA synthetase	
g4365_chr11_Ss	-1.65	0.025546184976171	related to NMA2-nicotinate-nucleotide adenylyltransferase	
g2167_chr04_Ss	-1.65	0.025540456171605	related to cop9 complex subunit 3	
g6151_chr20_Ss	-1.72	0.005429514944931	related to CDC39-transcriptional regulator protein	X
g1434_chr03_Ss	-1.75	0.004898069962126	conserved hypothetical protein	
g564_chr01_Ss	-1.78	0.025409359884287	MON2-peripheral membrane protein	
g889_chr02_Ss	-1.82	0.006196323554888	related to GPI-transamidase subunit	
g2546_chr05_Ss	-1.82	0.006194903238711	related to APG2-required for sporulation	
g2721_chr06_Ss	-1.82	0.003708835012848	related to cleft lip and palate transmembrane protein 1 (CLPTM1)	
g1157_chr02_Ss	-1.94	0.01188578700381	probable DNA repair endonuclease rad2	

Continue of Supplementary Table 1…

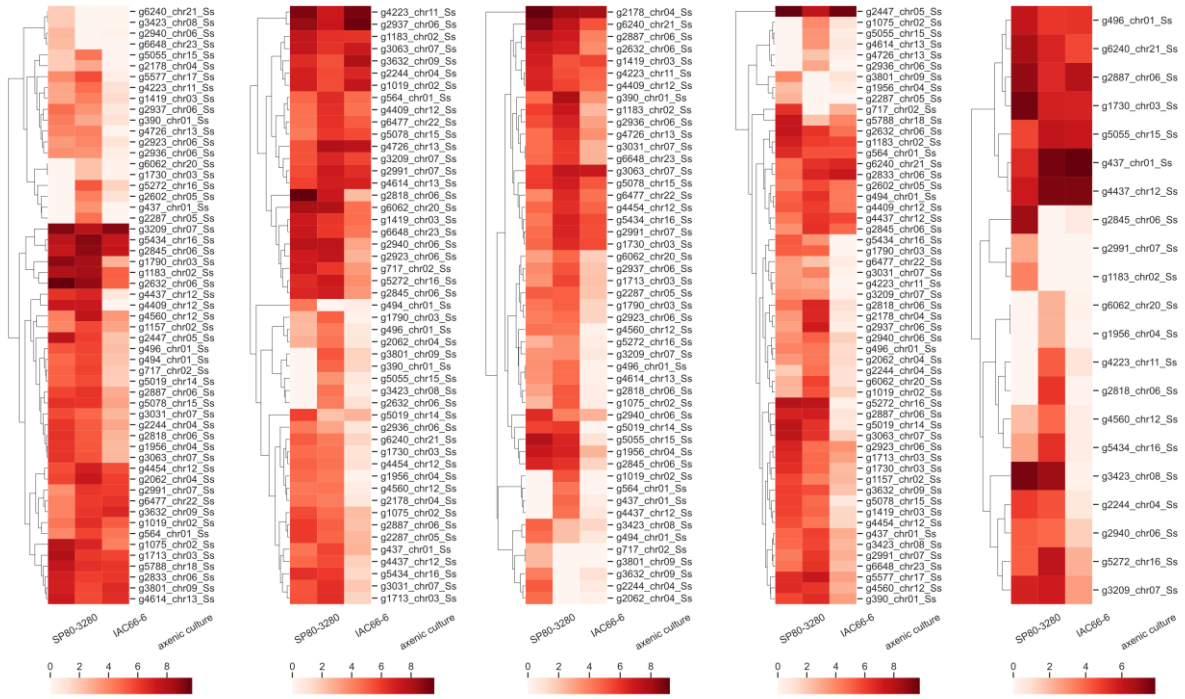
Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g5676_chr17_Ss	-1.94	0.011883748493177	conserved hypothetical protein	
g3632_chr09_Ss	-1.94	0.011882288806259	probable ADE6-phosphoribosylformyl glycinamidine synthetase	
g6201_chr20_Ss	-1.95	0.001308391558743	conserved hypothetical protein	
g3285_chr08_Ss	-1.95	0.000810021201589	conserved hypothetical protein	
g3591_chr09_Ss	-1.95	0.001376081786448	conserved hypothetical protein	X
g2998_chr07_Ss	-1.99	0.001390665886129	related to NADPH2:quinone reductase	
g2846_chr06_Ss	-2.01	0.001114122512299	probable HHT1-histone H3	
g1019_chr02_Ss	-2.06	0.002773400397267	HMG box-containing protein	
g6098_chr20_Ss	-2.11	0.001000992309217	conserved hypothetical protein	
g2305_chr05_Ss	-2.14	0.010652979165279	conserved hypothetical protein	
g2991_chr07_Ss	-2.14	0.010651762447358	ATP synthase subunit e, mitochondrial	
g4165_chr11_Ss	-2.19	0.000446128060246	conserved hypothetical protein	
g3349_chr08_Ss	-2.20	0.000521732087891	probable small subunit of ribonucleotide reductase	
g6477_chr22_Ss	-2.23	0.00243763005505	probable ADE4-amidophosphoribosyltransferase	
g3181_chr07_Ss	-2.31	0.004622526744423	related to ERG27-3-keto sterol reductase	

Continue of Supplementary Table 1…

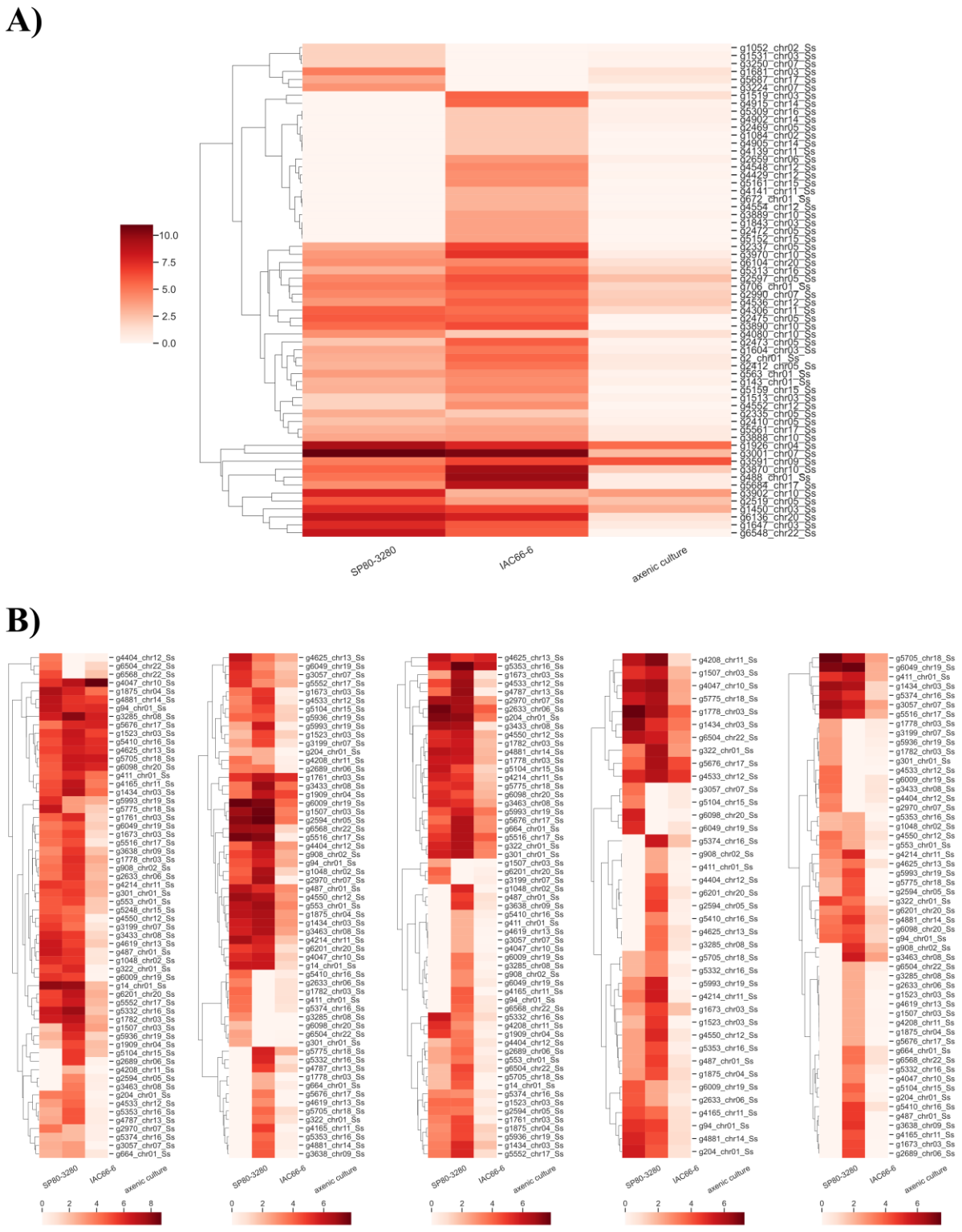
Gene id	logFC	PValue	Function (Blast2GO)	Secretome
g1709_chr03_Ss	-2.31	0.004621927510869	related to Oxidoreductase. short-chain dehydrogenase	
g1892_chr04_Ss	-2.37	0.000282463240102	related to Mig2 protein	
g4498_chr12_Ss	-2.46	0.001971361685424	conserved hypothetical protein	
g6107_chr20_Ss	-2.68	1.85803769076097E-05	conserved hypothetical protein	
g2648_chr06_Ss	-2.95	9.92451176053416E-06	conserved hypothetical protein	
g1102_chr02_Ss	-2.95	6.12560892105094E-06	probable Thiamin biosynthetic enzyme	
g3870_chr10_Ss	-4.25	0.008863268094894	conserved hypothetical protein	X
g488_chr01_Ss	-4.94	1.25046106807464E-06	conserved hypothetical protein	X



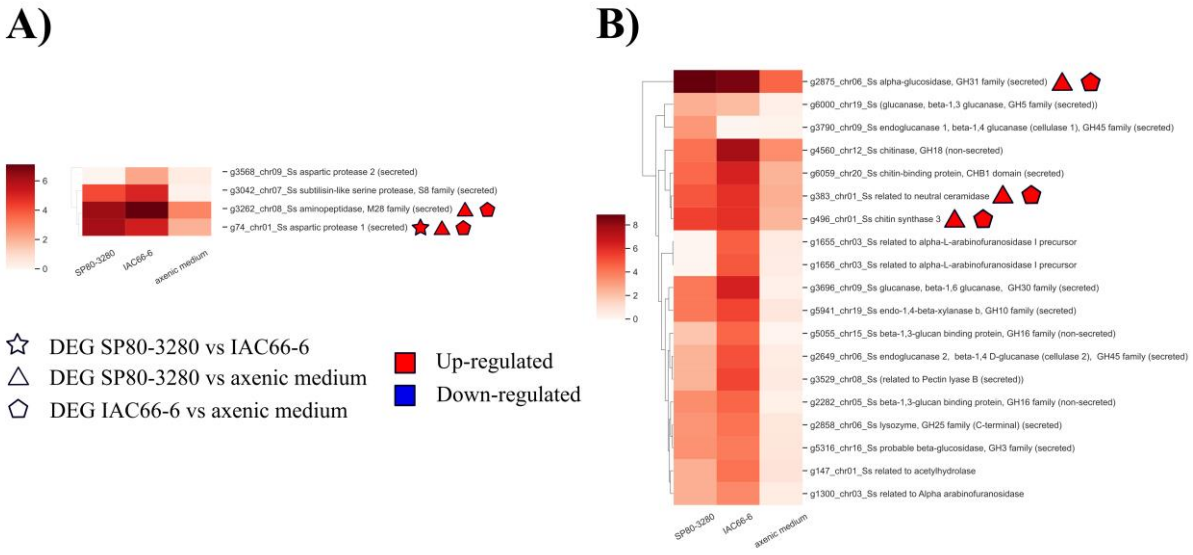
Supplementary Figure 1 – Heatmap for visualization of expression patterns (log2 of average FPKM of biological triplicate) of *S. scitamineum* genes highly expressed *in planta* in comparison with growth in axenic culture. A) Non-secreted genes. B) Secreted genes.



Supplementary Figure 2 – Heatmap for visualization of expression patterns (log2 of average FPKM of biological triplicate) of *S. scitamineum* genes non-secreted, with Blast2GO annotation (not shown) using the three approaches (DEGs + Highly expressed *in planta* + Exclusively expressed *in planta*).



Supplementary Figure 3 – Heatmap for visualization of expression patterns (log2 of average FPKM of biological triplicate) of *S. scitamineum* genes not annotated with Blast2GO (Hypothetical function) using the three approaches (DEGs + Highly expressed *in planta* + Exclusively expressed *in planta*). A) Secreted proteins-encoding genes. B) Non-secreted protein-encoding genes.



Supplementary Figure 4 – Heatmap for visualization of expression patterns (log of average FPKM of biological replicates) of *S. scitamineum* proteases (A) and hydrolases (B).