농학박사 학위논문

기주식물 침입 시 벼 도열병균의 선택적 스플라이싱 다양성 분석

Genome-wide analysis of alternative splicing complexity in *Magnaporthe oryzae* during infection

2021년 2월

서울대학교 대학원

협동과정 농생명유전체학 전공

전 중 범

Genome-wide analysis of alternative splicing complexity in *Magnaporthe oryzae* during infection

A dissertation submitted in partial
fulfillment of the requirement for
the degree of

DOCTOR OF PHILOSOPHY

to the Faculty of
Interdisciplinary Program in Agricultural Genomics
at

SEOUL NATIONAL UNIVERSITY

by

Jongbum Jeon

FEBRUARY 2021

농학박사 학위논문

기주식물 침입 시 벼 도열병균의 선택적 스플라이싱 다양성 분석

지도교수 이 용 환

이 논문을 농학박사 학위논문으로 제출함

2020년 12월

서울대학교 대학원
협동과정 농생명유전체학 전공

전 종 범

전종범의 박사 학위논문을 인준함

2020년 12월

위 원 장 최 도일 (인)

부 위 원 장 이 용환 (인)

위 원 최 진희 (인)

위 원 고 순옥 (인)

위 원 최 재혁 (인)

A THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Genome-wide analysis of alternative splicing complexity in *Magnaporthe oryzae* during infection

UNDER THE DIRECTION OF DR. YONG-HWAN LEE

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

BY

JONGBUM JEON

INTERDISCIPLINARY PROGRAM IN AGRICULTURAL GENOMICS

DECEMBER 2020

APPROVED AS A QUALIFIED THESIS OF JONGBUM JEON FOR THE DEGREE OF DOCTOR OF PHILOSOPHY BY THE COMMITTEE MEMBERS

CHAIRMAN Doil CHOI

VICE CHAIRMAN Yong-Hwan Lee

MEMBER Jin Hoe Huh

MEMBER Soonok Kim

MEMBER Taehyuk Choi

ABSTRACT

Genome-wide analysis of alternative splicing complexity in *Magnaporthe oryzae* during infection

Jongbum Jeon
Interdisciplinary Program in Agricultural Genomics
The Graduate School
Seoul National University

The recent advancement of RNA sequencing technology and the accumulation of transcriptome has enabled comprehensive transcriptome research. Transcriptomic analysis revealed that gene is controlled by post-transcriptional regulation. One of the post-transcriptional regulating mechanisms is alternative splicing (AS). This mechanism is known to modulate gene expression and protein function. In human and plant genomes, the AS mechanism was revealed to be regulated by cell differentiation, environmental adaptation, and assorted stress cues. The pathogen also has been understood as a cue to induce AS reprogramming of transcriptome networks in the plant genome. However, previous study of AS diversification has focused on plant transcriptome, and the AS mechanism of the pathogen during microbe-plant interaction was poorly understood. Therefore, profiling of AS

diversification in the fungal genome is necessary to decipher how fungal pathogens counteract or overcome the innated plant defense system.

This study profiled the AS repertoire of *Magnaporthe oryzae* during rice-infection to understand the response of the fungal system. Especially the identification of AS repertoire from diverse development or infection stages provides the multifaceted transcriptome to understand the temporal infection process. These AS repertoires increased during the whole infection compared to vegetative growth stages. Moreover, the specific AS isoform in the early infection stages increased after the fungus penetrated a host cell, which continued up to the biotrophic stages and decreased in necrotrophic stages. This compositional pattern suggests that the AS repertoire correlated to the fungal lifestyle during the infection process.

To decipher the roles of AS, understanding the post-transcriptional process of isoforms needs to follow. This study considered the two major AS roles, the nonsense-mediated decay (NMD) and protein diversification. In the roles of NMD, we found that majority of intron retention could regulate transcript decay. Subsequently, we identified proteins with domain structure transitions and alteration of secretion structure at potential translated isoforms. These protein variations in a functional region provide the clues of neo-functionalization in proteome during infection.

This study provides the AS repertoire throughout the infection process and will provide a foundation for molecular studies to unveil pathogenicity and adaptation mechanisms to environmental cues in *M. oryzae*. Moreover, the *ab initio* proteome analysis of the AS repertoire proposed a pleiotropic concept of the fungal genome.

Collectively, these AS profiles will provide insights into fungal AS transcriptome complexity during host-pathogen interaction.

Keywords: alternative splicing, pathogen-plant interaction, *Magnaporthe oryzae*, transcriptome diversification, pathogenicity

Student number: 2013-31043

CONTENTS

LIST OF TABLES

CHAPTER I

CHAPTER II

LIST OF FIGURES

CHAPTER I

CHAPTER II

CHAPTER I

Alternative splicing in fungi

ABSTRACT

Fungi are prevalent components of most ecosystems. These organisms inhabit a variety of ecological niches and so are effective models as models of responses to deleterious environmental factors. Alternative splicing (AS) is one such stress-response mechanism that is widespread in eukaryotes. Although RNA sequencing technology enables the genome-wide profiling of AS, fewer AS studies have focused on fungi than on plants and animals. In this chapter, we provide collective information on AS profiling of fungi. We summarize AS profiling studies using RNA sequencing and classify the condition-specific AS repertoire in fungi. Further, we summarize the role of AS in controlling gene functions under different environmental conditions.

INTRODUCTION

Alternative splicing (AS) is a key regulatory mechanism of gene expression via the production of at least two different transcripts from a pre-mRNA by different splicing processes (Chen and Manley, 2009). AS occurs in eukaryotes of the plant, animal, and fungal kingdoms. Assorted transcriptional elements in the transcriptome, including AS isoforms, modulate genomic systems to promote adaptation to a changing environment (Jabre et al., 2019). In eukaryotes, AS modulates RNA stability, protein localization, structure, and function as well as translational efficiency (Park et al., 2018). Thus, AS expands the functional repertoire of the eukaryotic genome.

In higher eukaryotes, AS modulates cell differentiation (Fiszbein and Kornblihtt, 2017) and controls responses to abiotic and biotic stresses (Ling et al., 2017; Laloum et al., 2018), and regulates the circadian rhythm (Torres et al., 2018). As a result, the production of alternatively spliced transcripts is subject to condition-specific regulation and evidence obtained under multiple conditions is needed to investigate AS. RNA-Seq enables the profiling of AS based on the transcript patterns in a single RNA library (Stark et al., 2019). Up to 95% and 70% of human and plant genes, respectively, are subjected to AS (Pan et al., 2008; Chamala et al., 2015). Despite the increasing number of fungal RNA-Seq-based transcriptomes under different environmental conditions, available, limited information is characterized and profiled on the mechanism and patterns of AS in fungi.

We first review the advances in AS methods and fungal AS profiling (Section I).

Moreover, we summarize genome-scale profiles of AS in multiple fungal species—both saprotrophs and animal and plant pathogens. We also overview AS in fungi in response to abiotic and biotic environmental cues (Figure 1) (Section II). Finally, we review the functions of AS in fungi as discovered by global transcriptomics and other molecular analyses (Section III).

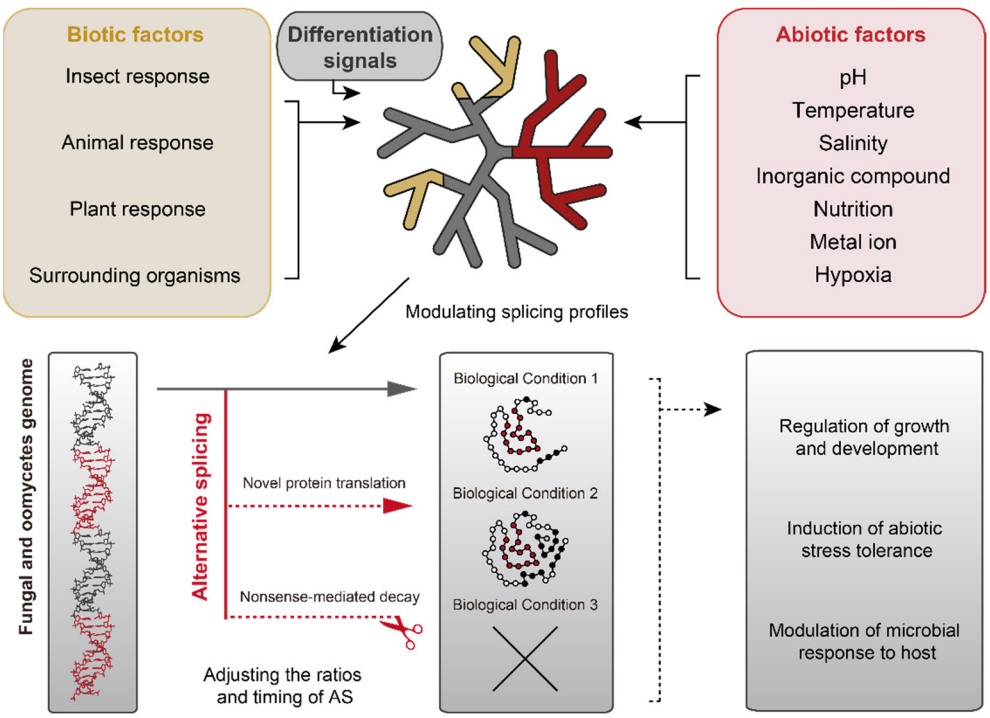


Figure 1. Schematic diagram of the effect of AS on cellular mechanisms. Environmental cues induce splicing complexity and modulate the expression and translation of stress-responsive genes. Splicing profiles in fungi are fine-tuned by the modulation of transcript ratios and modulating time points, increasing transcriptome and proteome composition diversity and facilitating adaptation to deleterious environments (dashed lines indicate less well-understood mechanisms).

I. Advances in genome-wide AS profiling of fungi

Genome-wide AS studies aim to identify gene models and evaluate transcriptome diversified conditions. Before the advent of RNA-Seq, only a small fraction of AS events could be identified by the available transcript assays. Reverse transcription-polymerase chain reaction (RT-PCR) (Davis et al., 2000) and Northern blotting (Langford and Gallwitz, 1983; Engebrecht et al., 1991) is the most simple and powerful approach to identify the specific AS transcripts. RT-PCR enables the identification of single alternatively spliced transcripts. However, these molecular approaches have limitations for evaluating the genome-wide AS pattern. The earliest attempts to profile AS were conducted in yeast using expressed sequence tags (ESTs) (Kupfer et al., 2004) and DNA microarrays (Clark et al., 2002). However, these methods detected only a narrow alternatively spliced region, and so could not identify the patterns of alternatively spliced isoforms.

RNA-Seq enables the identification and quantification of transcripts by high-throughput sequencing (HTS), improving the identification efficiency of alternatively spliced transcripts (Wang et al., 2019). Following the first profiling of AS by RNA-Seq—that of *Fusarium graminearum* (Zhao et al., 2013)—multiple AS profiling studies of fungi have been conducted (Table 1). Moreover, a new RNA-Seq technology based on long-read sequences has improved the identification of AS (Steijger et al., 2013; Kuang et al., 2017). Also, long-read sequencing, including Iso-Seq, allows the identification and characterization of AS by generating full-length transcript sequences (Kuang et al., 2017).

II. Environment-specified AS profiling in fungi

Fungi inhabit almost all ecological niches, including extreme environments (Cantrell et al., 2011). To adapt to environmental transitions, they have developed mechanisms to respond to environmental cues (Braunsdorf et al., 2016), one of which is AS (Jabre et al., 2019). Fungi have relatively small genomes (Mohanta and Bae, 2015) and genes with small introns (Neuveglise et al., 2011) compared to plants or animals. For these reasons, AS mechanisms in fungi have not been investigated extensively. However, the advent of RNA-Seq has accelerated research on AS. Below, we summarize the genome-wide diversity of AS across fungal lifestyle (Table 1).

Saprotrophic fungi: abiotic stress and differentiation models

Fungi decompose and obtain nutrients from non-living organic material or detritus. Saprotrophic fungi reproduce at a lower rate than pathogenic fungi (Grutzmann et al., 2014). RNA-Seq-based AS studies enable characterization of the AS repertoires of saprotrophic fungi. In plants, AS mediates rapid adjustment of the expression and function of factors that respond to environmental cues (Laloum et al., 2018). The multiple saprotrophic studies also discovered the different AS transcripts production in surrounding abiotic stress. Abiotic stresses—including salinity, temperature, and nutrient deficiencies—modulate the production of alternatively spliced transcripts. Transcriptome analysis of the halophilic fungus *Wallemia ichthyophaga* showed that the NaCl concentration affected the alternatively spliced transcript profile (Zajc et al., 2013). Similarly, temperature altered alternatively

spliced transcript production in the psychrophilic fungus *Mrakia psychrophile* (Su et al., 2016). Moreover, nitrogen starvation altered the list of alternatively spliced transcripts in *Aspergillus nidulans* (Sibthorp et al., 2013), as did nutrient starvation in *Trichoderma longibrachiatum* (Xie et al., 2015). AS modulates morphogenesis in animals (Baralle and Giudice, 2017). In the mushroom-forming fungi *Armillaria ostoyae*, *Rickenella mellea*, *Lentinus tigrinus*, *Coprinopsis cinerea*, *Phanerochaete chrysosporium*, and *Schizophyllum commune*, the list of alternatively spliced transcripts changes during vegetative growth to fruitbody development (Gehrmann et al., 2016; Krizsan et al., 2019).

Pathogenic and symbiotic fungi: models of biotic stresses

Fungi engage in symbiotic and pathogenic interactions with their hosts (Remmele et al., 2015). Fungi are causal agents of plant and animal diseases, in which recognition of host cues induces adaptive transcriptional reprogramming to, for instance, evade host immunity (Hube, 2009; van der Does and Rep, 2017). Human pathogenic fungi have a higher tendency to produce alternatively spliced transcripts than non-pathogens (Grutzmann et al., 2014). Studies of animal pathogenic fungi have investigated their contribution to stress adaptation in host cells. Differential AS is implicated in infection by five human pathogenic fungi (Candida albicans, *Candida parapsilosis*, *Histoplasma capsulatum*, *Cryptococcus neoformans*, and *Lichtheimia corymbifera*). However, the host response is less affected than by other abiotic stresses (Sieber et al., 2018). Fungi are the causal agents of numerous plant diseases. Three smut fungi (two *Ustilago* species and *Sporisorium reilianum*),

Colletotrichum graminicola, and *Pseudoperonospora cubensis* generate alternatively spliced transcripts to adapt to the host (Schliebner et al., 2014; Burkhardt et al., 2015; Donaldson et al., 2017). The AS repertoires during infection by *Rhizoctonia solani* and *Sclerotinia sclerotiorum* are specific to the host (Xia et al., 2017; Ibrahim et al., 2020). The AS profiling of *Magnaporthe oryzae* showed that the transcriptome is reprogrammed by AS in a manner dependent on lifestyle (Jeon et al., this thesis). Although two studies of symbionts were conducted, the symbiotic conditions were not considered.

Table 1. Genome-wide identification of fungal alternative splicing

Species Name	Life style*	Host	Condition	Reference
Saccharomyces cerevisiae	NP	-	Vegetative growth	Schreiber et al., 2015
Armillaria ostoyae	NP	-	Developmental stages	Krizsan et al., 2019
Coprinopsis cinerea	NP	-	Developmental stages	Krizsan et al., 2019
Phanerochaete chrysosporium	NP	-	Developmental stages	Krizsan et al., 2019
Rickenella mellea	NP	-	Developmental stages	Krizsan et al., 2019
Schizophyllum commune	NP	-	Developmental stages	Gehrmann et al., 2016, Krizsan et al., 2019
Lentinus tigrinus	NP	-	Developmental stages	Krizsan et al., 2019
Wallemia ichthyophaga	NP	-	Salinity (30% NaCl), vegetative growth	Zajc et al., 2013
Mrakia psychrophila	NP	-	Temperature (4℃ Cold shock), vegetative growth	Su et al., 2016
Schizosaccharomyces pombe	NP	-	Meiosis	Kuang et al., 2017
Trichoderma longibrachiatum	NP	-	Nutrition scarse, mycelia	Xie et al., 2015
Aspergillus nidulans	NP	-	Nitrogen starvation, ammonia, mycelia	Sibthorp et al., 2013
Candida albicans	AP	Human	Infection, hydroperoxide, organic acid, vegetative growth	Sieber et al., 2018, Muzafar et al., 2020
Candida parapsilosis	AP	Human	Infection, hypoxia, vegetative growth	Sieber et al., 2018
Candida glabrata	AP	Human	Infection, DMSO, pH(pH4, pH8), oxidative stress, vegetative growth	Linde et al., 2015, Sieber et al., 2018
Aspergillus fumigatus	AP	Human	Oxidative stress, antifungal drug (caspofungin),	Sieber et al., 2018
Histoplasma capsulatum	AP	Human	Infection, hyphal growth	Sieber et al., 2018
Cryptococcus neoformans	AP	Human	Temperature, nutrition, capsule induced conditions	Sieber et al., 2018
Lichtheimia corymbifera	AP	Human	Infection, hypoxia, metal ion (iron depletion)	Gonzalez-Hilarion et al., 2016, Sieber et al., 2018

Beauveria bassiana	AP	Insect	Infection, mycelia	Dong et al., 2017
Cordyceps militaris	AP	Insect	Infection, mycelia	Chen et al., 2019
Fusarium graminearum	PP	Wheat	mycelia	Zhao et al., 2013
Rhizoctonia solani	PP	Rice, wheat, maize	Infection (3 different hosts), mycelia	Xia et al., 2017
Pseudoperonospora cubensis	PP	Cucumber	Infection, sporangia	Burkhardt et al., 2015
Sclerotinia sclerotiorum	PP	Dicots	Infection (6 different hosts), mycelia	Ibrahim et al., 2020
Colletotrichum graminicola	PP	Maize	Infection	Schliebner et al., 2014
Verticillium dahliae	PP	Arabidopsis	Mycelia	Jin et al., 2017
Ustilago maydis	PP	Maize	Haploid mycelia, infection (dikaryon)	Donaldson et al., 2017
Ustilago hordei	PP	Barley, oat	Haploid mycelia, dikaryon mycelia	Donaldson et al., 2017
Sporisorium reilianum	PP	Maize,sorghum	Haploid mycelia, dikaryon mycelia	Donaldson et al., 2017
Magnaporthe oryzae **	PP	Rice	Infection (5 time points), mycelia	Jeon et al., this thesis
Trichophyton rubrum	AP,Sym	Human	Antifungal drug (undecanoic acid), mycelia	Mendes et al., 2018
Shiraia bambusicola	Sym	Bamboo	Chemical (Triton X-100), UV, mycelia	Liu et al., 2020

* NP : Non-pathogenic fungus, AP : Animal pathogen, PP : Plant pathogen, Sym : Symbiont
** The profiling reported in this Ph.D. thesis

III. Major biological functions of the alternatively spliced transcripts in fungi

AS allows expansion of the proteome (Nilsen and Graveley, 2010) and modulates gene expression by nonsense-mediated decay (NMD) (Drechsel et al., 2013; Tress et al., 2017) (Figure 1). The biological functions are being regulated through AS mechanisms is important for deciphering how fungi respond to stress, but roles of AS repertoires in fungi are rarely studied. However, molecular studies (Table 2) and in silico prediction (Table 1) have provided evidence on the functions of AS in fungi (Fang et al., 2020). In the following sections, we summarize the roles of AS.

Fungal growth and differentiation

AS modulates the outcome of pluripotent cell differentiation and somatic reprogramming (Fiszbein and Kornblihtt, 2017). The lifecycle of fungi encompasses vegetative, asexual, and sexual phases. Mushrooms are used as models to investigate the link between AS and cell differentiation. The AS repertoire has been investigated during mushroom development (Gehrmann et al., 2016; Krizsan et al., 2019). During the development of *S. commune*, AS affects carbohydrate-active enzymes, metabolic genes, transcription factors, secreted proteins, and cytochrome P450. The expression of some alternatively spliced transcripts differed across the developmental stages (Gehrmann et al., 2016).

To date, several genes producing alternatively spliced isoforms related to fungal growth have been characterized. In vegetative *S. cerevisiae*, *SRC1* is modulated by AS (Grund et al., 2008). AS of *SRC1* generates alternatively spliced isoforms with

different membrane-spanning sequences (Grund et al., 2008). Moreover, a study of *hub1* showed that upstream spliceosome regulators can disrupt the fine-tuning of AS (Mishra et al., 2011). In addition, *Mer2* and *Mer3* in *S. cerevisiae* produce alternatively spliced isoforms during meiosis. AS also altered the localization of YlMDH1 (Kabran et al., 2012). Collectively, AS could affect the subcellular localization and binding motifs of proteins during development.

Responses to abiotic stresses

Fungi inhabit diverse niches, in which they are exposed to multiple environmental cues. These include biotic (described below section) and abiotic (temperature, pH, salinity, metal ion) cues. In M. psychrophile, AS affects transporters, the ribonucleoprotein complex, metal-ion-binding proteins, and hydrolase activity at low temperatures (Su et al., 2016). Moreover, temperature modulates AS in *S. cerevisiae* (*Fes1, SUS1*) and *Lachancea kluyveri* (*SKI7/HBS1*) (Hossain et al., 2011; Marshall et al., 2013; Gowda et al., 2016). AS of *Fes1* in response to temperature controls protein localization by means of the nuclear localization signal (NLS) (Gowda et al., 2016). SUS1 induces NMD under heat stress (Hossain et al., 2011). In *L. kluyveri*, AS of *SKI7/HBS1* results in the generation of two functionally distinct proteins, HBS1 and SKI7, from a single gene region. Moreover, the choice of AS and duplication of these two genes provide evidence how AS replaces duplication mechanism in fungal species (Marshall et al., 2013).

Responses to temperature, pH, and oxygen levels involve the AS of genes encoding DNA binding proteins and transcription factors (Sieber et al., 2018). AS

isoforms play roles in the responses to abiotic stresses (e.g., pH) and drug resistance. *Neurospora crassa hex1* and *A. nidulans palB* generate alternatively spliced transcripts depending on the phosphate concentration and pH and modulate *pacC* expression (Leal et al., 2009; Trevisan et al., 2011). AS also modulates the drug resistance mediated by *YlHAC1* (dithiothreitol and tunicamycin), *SOD3* (amphotericin B and menadione), and *sbPKS* (nourseothricin) using protein domain or structure transition (Oh et al., 2010; Liu et al., 2020; Muzafar et al., 2020). Therefore, AS regulates the expression of genes involved in responses to a variety of abiotic stresses; also, fine-tuning of AS modulates such responses.

Defenses against and adaptation to the host environment

Recent studies have been revealed that AS is one of key mechanisms in microbe-host interactions (Rigo et al., 2019; Chauhan et al., 2019), but few studies have focused on its role in the interactions of eukaryotic pathogens and symbionts (Table 2). A comparative study of six human pathogens showed that membrane regulation genes were subjected to AS during invasion (Sieber et al., 2018). The AS profile of the entomopathogenic fungus *Beauveria bassiana* was enriched in the category protein synthesis, cell differentiation, binding protein, energy, regulatory mechanisms, cellular communication, and cell fate. AS of *BbATG8* generates two isoforms linked to conidia/blastospore formation and pathogenesis (Dong et al., 2017).

In the causal agent of cucurbit downy mildew, *P. cubensis*, AS modulates protein secretion (Burkhardt et al., 2015). Moreover, AS of *P. cubensis PscRXLR1* generates

isoforms with signal peptide variation. AS is implicated in the control of rapid necrosis in *Nicotiana benthamiana* (Savory et al., 2012). In *S. sclerotiorum*, AS affects genes encoding secreted proteins, oxidoreductases, and genes linked to carbohydrate metabolism (Ibrahim et al., 2020). Moreover, two serine and arginine-rich (SR) proteins of *FgSrp1* and *FgSrp2* are regulated by AS; in addition to controlling the splicing mechanism, FgSrp1, but not FgSrp2, regulates pathogenesis (Zhang et al., 2017; Zhang et al., 2020). These SR proteins are fine-tuned by a specific kinase during infection (Wang et al., 2018). Studies of *M. oryzae MoGrp1* and *MoHMT1* showed that several genes regulate AS during infection (Gao et al., 2019; Li et al., 2020). MoGrp1 regulates genes linked to infection-related morphology (Gao et al., 2019) and MoHMT1 regulates genes related to autophagosome formation (Li et al., 2020). Therefore, fungal AS studies provide evidence that AS is essential component of the response to the host.

Table 2. Molecular characterization of fungal alternative splicing related genes

Species	Gene	AS inducing cues	Reference
Aspergillus nidulans	*palB*	pH response	Trevisan et al., 2011
Lachancea kluyveri	*SKI7/HBS1*	Temperature	Marshall et al., 2013
Monascus pilosus	*MpLaeA*	Secondary metabolism	Zhang and Miyake, 2009
Neurospora crassa	*hex1*	Chemical (inorganic phosphate), pH	Leal et al., 2009
	RRP44	Assembly of the spliceosome	Zhang et al., 2015
	SRC1	Growth	Grund et al., 2008
	Hub1	Growth	Mishra et al., 2011
Saccharomyces cerevisiae	*MER2*	Meiosis	Engebrecht et al., 1991
	MER3	Meiosis	Nakagawa and Ogawa, 1999
	SUS1	Temperature	Hossain et al., 2011
	Fes1	Temperature	Gowda et al., 2016
	PTC7	-	Juneau et al., 2009
Yarrowia lipolytica	*YlMDH2*	Growth	Kabran et al., 2012
	YlHAC1	Drug resistance (dithiothreitol and tunicamycin)	Oh et al., 2010
	YlUPF1, YlUPF2	-	Mekouar et al., 2010
Xanthophyllomyces dendrorhous	*crtI, crtYB*	Response to environment	Lodato et al., 2003
Candida albicans	*SOD3*	Drug resistance	Muzafar et al., 2020
	RPL30, SPR28	-	Mitrovich et al., 2007
Beauveria bassiana	*BbATG8*	Autophagy	Dong et al., 2017
Epichloë festucae	*Rose City salicylate hydroxylase*	Enzymatic activity	Ambrose et al., 2015
Fusarium graminearum	*Srk1*	Vegetative growth, conidiogenesis, sexual reproduction, pathogenesis	Wang et al., 2018
	FgSrp1	Asexual reproduction, pathogenesis	Zhang et al., 2017
	FgSrp2	Vegetative growth, perithecia melanization	Zhang et al., 2020
Magnaporthe oryzae	*MoGrp1*	Cell wall intetgrity, conidiation, pathogenesis	Gao et al., 2019
	MoHMT1	Growth, virulence	Li et al., 2020
Pseudoperonospora cubensis	*PscRXLR1*	Pathogenesis	Savory et al., 2012
Ustilago maydis	*Gapdh*	Virulence	Freitag et al., 2012
	UmRrm75	Cell morphology, pathogenesis	Rodriguez-Kessler et al., 2012
Shiraia bambusicola	*sbPKS*	Abiotic stress (fungicide, oxidative)	Liu et al., 2020

PERSPECTIVE

Fungi are some of the most diversified organisms and cope with environments ranging from the mild to the extreme. These organisms control the global cycling of nutrients and carbon sequestration in ecosystems. Moreover, some fungi cause physiological disorders in their host or are closely connected to other species and microbiome for benefits. These fungi have the ability to survive in a wide range of habitats and adapt to their environment. Investigation of the underlying mechanisms may provide insight into environmental adaptation by other organisms.

HTS has promoted AS profiling of fungi and the data from assorted conditions provide evidence to understand the correlation between AS and environmental cues. Studies to date have focused on AS properties and patterns. To unveil the roles of AS, functional characterization of transcripts and proteome prediction are needed. Recent technical advances such as Oxford Nanopore and PacBio Iso-Seq promote the detection of AS by the identification of full-length novel isoforms (Wang et al., 2019). Although long-read techniques provide much comprehensive data for plants and animals, they have been applied in few fungal studies. The application of these technological advances will enable us to more accurately predict the function of each alternatively spliced isoform in fungi.

Investigations of these roles of AS will enhance our understanding of the mechanisms of fungal adaptation to environmental cues. In addition, given the importance of fungi in industry, medicine, and agriculture, the mechanisms of AS that are active under a variety of conditions must be elucidated. Furthermore, the

accumulation of isoform repertoires will suggest novel targets for controlling fungal pathogens in medicine and agriculture.

LITERATURE CITED

Ambrose, K.V., Tian, Z.P., Wang, Y.F., Smith, J., Zylstra, G., Huang, B.R., and Belanger, F.C. (2015). Functional characterization of salicylate hydroxylase from the fungal endophyte *Epichloe festucae*. *Sci. Rep.* 5: 10939

Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18, 437-451.

Braunsdorf, C., Mailander-Sanchez, D., and Schaller, M. (2016). Fungal sensing of host environment. *Cell Microbiol.* 18, 1188-1200.

Burkhardt, A., Buchanan, A., Cumbie, J.S., Savory, E.A., Chang, J.H., and Day, B. (2015). Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. *Mol. Plant-Microbe Interact.* 28, 298-309.

Cantrell, S.A., Dianese, J.C., Fell, J., Gunde-Cimerman, N., and Zalar, P. (2011). Unusual fungal niches. *Mycologia* 103, 1161-1174.

Chamala, S., Feng, G., Chavarro, C., and Barbazuk, W.B. (2015). Genome-wide identification of evolutionarily conserved alternative splicing events in flowering plants. *Front. Bioeng. Biotechnol.* 3, 33.

Chauhan, K., Kalam, H., Dutt, R., and Kumar, D. (2019). RNA splicing: a new paradigm in host-pathogen interactions. *J. Mol. Biol.* 431, 1565-1575.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10, 741-754.

Chen, Y.J., Wu, Y.Q., Liu, L., Feng, J.H., Zhang, T.C., Qin, S., Zhao, X.Y., Wang,

C.X., Li, D.M., Han, W., Shao, M.H., Zhao, P., Xue, J.F., Liu, X.M., Li, H.J., Zhao, E.W., Zhao, W., Guo, X.J., Jin, Y.F., Cao, Y.M., Cui, L.W., Zhou, Z.Q., Xia, Q.Y., Rao, Z.H., and Zhang, Y.Z. (2019). Study of the whole genome, methylome and transcriptome of *Cordyceps militaris*. *Sci. Rep.* 9: 898

Clark, T.A., Sugnet, C.W., and Ares, M. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science* 296, 907-910.

Davis, C.A., Grate, L., Spingola, M., and Ares, M., Jr. (2000). Test of intron predictions reveals novel splice sites, alternatively spliced mRNAs and new introns in meiotically regulated genes of yeast. *Nucleic Acids Res.* 28, 1700-1706.

Donaldson, M.E., Ostrowski, L.A., Goulet, K.M., and Saville, B.J. (2017). Transcriptome analysis of smut fungi reveals widespread intergenic transcription and conserved antisense transcript expression. *BMC Genomics* 18, 340.

Dong, W.X., Ding, J.L., Gao, Y., Peng, Y.J., Feng, M.G., and Ying, S.H. (2017). Transcriptomic insights into the alternative splicing-mediated adaptation of the entomopathogenic fungus *Beauveria bassiana* to host niches: autophagy-related gene 8 as an example. *Environ. Microbiol.* 19, 4126-4139.

Drechsel, G., Kahles, A., Kesarwani, A.K., Stauffer, E., Behr, J., Drewe, P., Ratsch, G., and Wachtera, A. (2013). Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the *Arabidopsis* steady state transcriptome. *Plant Cell* 25, 3726-3742.

Engebrecht, J.A., Voelkel-Meiman, K., and Roeder, G.S. (1991). Meiosis-specific RNA splicing in yeast. *Cell* 66, 1257-1268.

Fang, S.M., Hou, X., Qiu, K.H., He, R., Feng, X.S., and Liang, X.L. (2020). The occurrence and function of alternative splicing in fungi. *Fungal Biol. Rev.* 34, 178-188.

Fiszbein, A., and Kornblihtt, A.R. (2017). Alternative splicing switches: important players in cell differentiation. *Bioessays* 39: 6

Freitag, J., Ast, J., and Bolker, M. (2012). Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature* 485, 522-U135.

Gao, X., Yin, C., Liu, X., Peng, J., Chen, D., He, D., Shi, W., Zhao, W., Yang, J., Peng, Y.-L. (2019). A glycine-rich protein MoGrp1 functions as a novel splicing factor to regulate fungal virulence and growth in *Magnaporthe oryzae*. *Phytopathol. Res*. 1, 1-15

Gehrmann, T., Pelkmans, J.F., Lugones, L.G., Wosten, H.a.B., Abeel, T., and Reinders, M.J.T. (2016). *Schizophyllum commune* has an extensive and functional alternative splicing repertoire. *Sci. Rep.* 6: 33640

Gonzalez-Hilarion, S., Paulet, D., Lee, K.T., Hon, C.C., Lechat, P., Mogensen, E., Moyrand, F., Proux, C., Barboux, R., Bussotti, G., Hwang, J., Coppee, J.Y., Bahn, Y.S., and Janbon, G. (2016). Intron retention-dependent gene regulation in *Cryptococcus neoformans*. *Sci. Rep*. 6: 32252

Gowda, N.K.C., Kaimal, J.M., Masser, A.E., Kang, W.J., Friedlander, M.R., and Andreasson, C. (2016). Cytosolic splice isoform of *Hsp70* nucleotide

exchange factor *Fes1* is required for the degradation of misfolded proteins in yeast. *Mol. Biol. Cell* 27, 1210-1219.

Grund, S.E., Fischer, T., Cabal, G.G., Antunez, O., Perez-Ortin, J.E., and Hurt, E. (2008). The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. *J. of Cell Biol*. 182, 897-910.

Grutzmann, K., Szafranski, K., Pohl, M., Voigt, K., Petzold, A., and Schuster, S. (2014). Fungal alternative splicing is associated with multicellular complexity and virulence: a genome-wide multi-species study. *DNA Res*. 21, 27-39.

Hossain, M.A., Rodriguez, C.M., and Johnson, T.L. (2011). Key features of the two-intron *Saccharomyces cerevisiae* gene *SUS1* contribute to its alternative splicing. *Nucleic Acids Res*. 39, 8612-8627.

Hube, B. (2009). Fungal adaptation to the host environment. *Curr. Opin. Microbiol*. 12, 347-349.

Ibrahim, H.M.M., Kusch, S., Didelon, M., and Raffaele, S. (2020). Genome-wide alternative splicing profiling in the fungal plant pathogen *Sclerotinia sclerotiorum* during the colonization of diverse host families. *Mol. Plant Pathol*. mpp.13006

Jabre, I., Reddy, A.S.N., Kalyna, M., Chaudhary, S., Khokhar, W., Byrne, L.J., Wilson, C.M., and Syed, N.H. (2019). Does co-transcriptional regulation of alternative splicing mediate plant stress responses? *Nucleic Acids Res*. 47, 2716-2726.

Jin, L.R., Li, G.L., Yu, D.Z., Huang, W., Cheng, C., Liao, S.J., Wu, Q.J., and Zhang, Y. (2017). Transcriptome analysis reveals the complexity of alternative splicing regulation in the fungus *Verticillium dahliae*. *BMC Genomics* 18.

Juneau, K., Nislow, C., and Davis, R.W. (2009). Alternative Splicing of PTC7 in *Saccharomyces cerevisiae* Determines Protein Localization. *Genetics* 183, 185-194.

Kabran, P., Rossignol, T., Gaillardin, C., Nicaud, J.M., and Neuveglise, C. (2012). Alternative splicing regulates targeting of malate dehydrogenase in *Yarrowia lipolytica*. *DNA Res*. 19, 231-244.

Krizsan, K., Almasi, E., Merenyi, Z., Sahu, N., Viragh, M., Koszo, T., Mondo, S., Kiss, B., Balint, B., Kues, U., Barry, K., Cseklye, J., Hegedus, B., Henrissat, B., Johnson, J., Lipzen, A., Ohm, R.A., Nagy, I., Pangilinan, J., Yan, J.Y., Xiong, Y., Grigoriev, I.V., Hibbett, D.S., and Nagy, L.G. (2019). Transcriptomic atlas of mushroom development reveals conserved genes behind complex multicellularity in fungi. *Proc. Natl. Acad. Sci. U.S.A*. 116, 7409-7418.

Kuang, Z., Boeke, J.D., and Canzar, S. (2017). The dynamic landscape of fission yeast meiosis alternative-splice isoforms. *Genome Res*. 27, 145-156.

Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A., and Murphy, J.W. (2004). Introns and splicing elements of five diverse fungi. *Eukaryot. Cell* 3, 1088-1100.

Laloum, T., Martin, G., and Duque, P. (2018). Alternative splicing control of abiotic stress responses. *Trends Plant Sci.* 23, 140-150.

Langford, C.J., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. *Cell* 33, 519-527.

Leal, J., Squina, F.M., Freitas, J.S., Silva, E.M., Ono, C.J., Martinez-Rossi, N.M., and Rossi, A. (2009). A splice variant of the *Neurospora crassa hex-1* transcript, which encodes the major protein of the Woronin body, is modulated by extracellular phosphate and pH changes. *FEBS Lett*. 583, 180-184.

Li, Z.Q., Wu, L.Y., Wu, H., Zhang, X.X., Mei, J., Zhou, X.P., Wang, G.L., and Liu, W.D. (2020). Arginine methylation is required for remodelling pre-mRNA splicing and induction of autophagy in rice blast fungus. *New Phytol*. 225, 413-429.

Linde, J., Duggan, S., Weber, M., Horn, F., Sieber, P., Hellwig, D., Riege, K., Marz, M., Martin, R., Guthke, R., and Kurzai, O. (2015). Defining the transcriptomic landscape of *Candida glabrata* by RNA-Seq. *Nucleic Acids Res*. 43, 1392-1406.

Ling, Y., Alshareef, S., Butt, H., Lozano-Juste, J., Li, L., Galal, A.A., Moustafa, A., Momin, A.A., Tashkandi, M., Richardson, D.N., Fujii, H., Arold, S., Rodriguez, P.L., Duque, P., and Mahfouz, M.M. (2017). Pre-mRNA splicing repression triggers abiotic stress signaling in plants. *Plant J*. 89, 291-309.

Liu, X.Y., Fan, L., Gao, J., Shen, X.Y., and Hou, C.L. (2020). Global identification of alternative splicing in *Shiraia bambusicola* and analysis of its regulation in hypocrellin biosynthesis. *Appl. Microbiol. Biotechnol.* 104, 211-223.

Lodato, P., Alcaino, J., Barahona, S., Retamales, P., and Cifuentes, V. (2003). Alternative splicing of transcripts from *crtI* and *crtYB* genes of *Xanthophyllomyces dendrorhous*. *Appl. Environ. Microbiol*. 69, 4676-4682.

Marshall, A.N., Montealegre, M.C., Jimenez-Lopez, C., Lorenz, M.C., and Van Hoof, A. (2013). Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. *PLoS Genet*. 9: e1003376

Mekouar, M., Blanc-Lenfle, I., Ozanne, C., Da Silva, C., Cruaud, C., Wincker, P., Gaillardin, C., and Neuveglise, C. (2010). Detection and analysis of alternative splicing in *Yarrowia lipolytica* reveal structural constraints facilitating nonsense-mediated decay of intron-retaining transcripts. *Genome Biol*. 11: R65

Mendes, N.S., Bitencourt, T.A., Sanches, P.R., Silva-Rocha, R., Martinez-Rossi, N.M., and Rossi, A. (2018). Transcriptome-wide survey of gene expression changes and alternative splicing in *Trichophyton rubrum* in response to undecanoic acid. *Sci. Rep*. 8: 2520

Mishra, S.K., Ammon, T., Popowicz, G.M., Krajewski, M., Nagel, R.J., Ares, M., Holak, T.A., and Jentsch, S. (2011). Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature* 474, 173-U205.

Mitrovich, Q.M., Tuch, B.B., Guthrie, C., and Johnson, A.D. (2007). Computational and experimental approaches double the number of known introns in the pathogenic yeast *Candida albicans*. *Genome Res.* 17, 492-502.

Mohanta, T.K., and Bae, H. (2015). The diversity of fungal genome. *Biol. Proced. Online* 17, 8.

Muzafar, S., Sharma, R.D., Shah, A.H., Gaur, N.A., Dasgupta, U., Chauhan, N., and Prasad, R. (2020). Identification of genomewide alternative splicing events in sequential, isogenic clinical isolates of *candida albicans* reveals a novel mechanism of drug resistance and tolerance to cellular stresses. *Msphere* 5: 4

Nakagawa, T., and Ogawa, H. (1999). The *Saccharomyces cerevisiae MER3* gene, encoding a novel helicase-like protein, is required for crossover control in meiosis. *EMBO J.* 18, 5714-5723.

Neuveglise, C., Marck, C., and Gaillardin, C. (2011). The intronome of budding yeasts. *C. R. Biol.* 334, 662-670.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457-463.

Oh, M.H., Cheon, S.A., Kang, H.A., and Kim, J.Y. (2010). Functional characterization of the unconventional splicing of *Yarrowia lipolytica HAC1* mRNA induced by unfolded protein response. *Yeast* 27, 443-452.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet*. 40, 1413-1415.

Park, E., Pan, Z., Zhang, Z., Lin, L., and Xing, Y. (2018). The expanding landscape of alternative splicing variation in human populations. *Am. J. Hum. Genet*. 102, 11-26.

Remmele, C.W., Luther, C.H., Balkenhol, J., Dandekar, T., Muller, T., and Dittrich, M.T. (2015). Integrated inference and evaluation of host-fungi interaction

networks. *Front. Microbiol*. 6, 764.

Rigo, R., Bazin, J.R.M., Crespi, M., and Charon, C.L. (2019). Alternative splicing in the regulation of plant-microbe interactions. *Plant Cell Physiol*. 60, 1906-1916.

Rodriguez-Kessler, M., Baeza-Montanez, L., Garcia-Pedrajas, M.D., Tapia-Moreno, A., Gold, S., Jimenez-Bremont, J.F., and Ruiz-Herrera, J. (2012). Isolation of *UmRrm75*, a gene involved in dimorphism and virulence of *Ustilago maydis*. *Microbiol. Res*. 167, 270-282.

Savory, E.A., Zou, C., Adhikari, B.N., Hamilton, J.P., Buell, C.R., Shiu, S.H., and Day, B. (2012). Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits a rapid cell death. *PLoS One* 7: e0034701

Schliebner, I., Becher, R., Hempel, M., Deising, H.B., and Horbach, R. (2014). New gene models and alternative splicing in the maize pathogen *Colletotrichum graminicola* revealed by RNA-Seq analysis. *BMC Genomics* 15, 842.

Schreiber, K., Csaba, G., Haslbeck, M., and Zimmer, R. (2015). Alternative splicing in next generation sequencing data of *Saccharomyces cerevisiae*. *PLoS One* 10: e0140487

Sibthorp, C., Wu, H.H., Cowley, G., Wong, P.W.H., Palaima, P., Morozov, I.Y., Weedall, G.D., and Caddick, M.X. (2013). Transcriptome analysis of the filamentous fungus *Aspergillus nidulans* directed to the global identification of promoters. *BMC Genomics* 14: 847

Sieber, P., Voigt, K., Kammer, P., Brunke, S., Schuster, S., and Linde, J. (2018).

Comparative study on alternative splicing in human fungal pathogens suggests its involvement during host invasion. *Front. Microbiol*. 9.

Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat. Rev. Genet*. 20, 631-656.

Steijger, T., Abril, J.F., Engstrom, P.G., Kokocinski, F., Consortium, R., Hubbard, T.J., Guigo, R., Harrow, J., and Bertone, P. (2013). Assessment of transcript reconstruction methods for RNA-seq. *Nat. Methods* 10, 1177-1184.

Su, Y., Jiang, X.Z., Wu, W.P., Wang, M.M., Hamid, M.I., Xiang, M.C., and Liu, X.Z. (2016). Genomic, transcriptomic, and proteomic analysis provide insights into the cold adaptation mechanism of the obligate psychrophilic fungus *Mrakia psychrophila*. *G3* 6, 3603-3613.

Torres, M., Becquet, D., Franc, J.L., and Francois-Bellan, A.M. (2018). Circadian processes in the RNA life cycle. *Wiley Interdiscip. Rev. RNA* 9, e1467.

Tress, M.L., Abascal, F., and Valencia, A. (2017). Alternative splicing may not be the key to proteome complexity. *Trends Biochem. Sci*. 42, 98-110.

Trevisan, G.L., Oliveira, E.H.D., Peres, N.T.A., Cruz, A.H.S., Martinez-Rossi, N.M., and Rossi, A. (2011). Transcription of *Aspergillus nidulans pacC* is modulated by alternative RNA splicing of *palB*. *FEBS Lett*. 585, 3442-3445.

Van Der Does, H.C., and Rep, M. (2017). Adaptation to the host environment by plant-pathogenic fungi. *Annu. Rev. Phytopathol*. 55, 427-450.

Wang, B., Kumar, V., Olson, A., and Ware, D. (2019). Reviving the transcriptome studies: An insight into the emergence of single-molecule transcriptome sequencing. *Front. Genet*. 10: 3389

Wang, G.H., Sun, P., Gong, Z.W., Gu, L.F., Lou, Y., Fang, W.Q., Zhang, L.H., Su, L., Yang, T., Wang, B.H., Zhou, J., Xu, J.R., Wang, Z.H., and Zheng, W.H. (2018). *Srk1* kinase, a SR protein-specific kinase, is important for sexual reproduction, plant infection and pre-mRNA processing in *Fusarium graminearum*. *Environ. Microbiol*. 20, 3261-3277.

Xia, Y., Fei, B.H., He, J.Y., Zhou, M.L., Zhang, D.H., Pan, L.X., Li, S.C., Liang, Y.Y., Wang, L.X., Zhu, J.Q., Li, P., and Zheng, A.P. (2017). Transcriptome analysis reveals the host selection fitness mechanisms of the *Rhizoctonia solani* AG1IA pathogen. *Sci. Rep*. 7: 10120

Xie, B.B., Li, D., Shi, W.L., Qin, Q.L., Wang, X.W., Rong, J.C., Sun, C.Y., Huang, F., Zhang, X.Y., Dong, X.W., Chen, X.L., Zhou, B.C., Zhang, Y.Z., and Song, X.Y. (2015). Deep RNA sequencing reveals a high frequency of alternative splicing events in the fungus *Trichoderma longibrachiatum*. *BMC Genomics* 16: 54

Zajc, J., Liu, Y.F., Dai, W.K., Yang, Z.Y., Hu, J.Z., Gostincar, C., and Gunde-Cimerman, N. (2013). Genome and transcriptome sequencing of the halophilic fungus *Wallemia ichthyophaga*: haloadaptations present and absent. *BMC Genomics* 14: 617

Zhang, L., Wan, Y.F., Huang, G.B., Wang, D.N., Yu, X.Y., Huang, G.C., and Guo, J.H. (2015). The exosome controls alternative splicing by mediating the gene expression and assembly of the spliceosome complex. *Sci. Rep*. 5: 13403

Zhang, M.Y., and Miyake, T. (2009). Development and media regulate alternative splicing of a methyltransferase pre-mRNA in *Monascus pilosus*. *J. Agric.*

and Food Chem. 57, 4162-4167.

Zhang, Y., Dai, Y., Huang, Y., Wang, K., Lu, P., Xu, H., Xu, J.R., and Liu, H. (2020). The SR-protein FgSrp2 regulates vegetative growth, sexual reproduction and pre-mRNA processing by interacting with FgSrp1 in *Fusarium graminearum*. *Curr. Genet*. 66, 607-619.

Zhang, Y.M., Gao, X.L., Sun, M.L., Liu, H.Q., and Xu, J.R. (2017). The *FgSRP1* SR-protein gene is important for plant infection and pre-mRNA processing in *Fusarium graminearum*. *Environ. Microbiol*. 19, 4065-4079.

Zhao, C.Z., Waalwijk, C., De Wit, P.J.G.M., Tang, D.Z., and Van Der Lee, T. (2013). RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen *Fusarium graminearum*. *BMC Genomics* 14: 21

CHAPTER II

Alternative splicing diversifies the transcriptome and proteome of the rice blast fungus during infection

ABSTRACT

Alternative splicing (AS) contributes to diversifying and regulating cellular responses to environmental cues by differentially producing multiple mRNA and protein variants from a single gene. Previous studies on AS in pathogenic fungi have focused on profiling AS isoforms under a limited number of conditions to check if AS occurs during host infection. To understand how AS participates in pathogenesis, comprehensive analyses of AS patterns at multiple developmental and infection stages are needed. Here, we performed AS profiling in the rice blast fungus *Magnaporthe oryzae*, a worldwide threat to rice production, using high-quality transcriptome data representing its vegetative growth and different host infection stages. We identified 4,270 AS isoforms derived from 2,413 genes, including 499 genes regulated by infection-specific AS. AS appears to increase during infection, with 32.7% of the AS isoforms being detected only during infection. Analysis of the isoforms observed at each stage showed that 636 genes produce isoforms dominantly during infection, especially after initial fungal penetration into host cell. Many such isoforms were predicted to encode regulatory proteins such as transcription factors and phospho-transferases. We also identified the genes encoding proteins with distinct domain structures via AS and confirmed the translation of some isoforms via proteomic analysis, suggesting potential AS-mediated neo-functionalization during infection. This systematic study revealed the nature and regulation of AS during plant-pathogen interactions and established a foundation data resource for validating the role of AS during pathogenesis.

INTRODUCTION

Alternative splicing (AS) is one of the regulatory mechanisms for gene expression and produces multiple protein variants from a single gene by modulating the maturation of mRNA precursors in more than one way to generate heterogeneous transcripts (Shi, 2017). AS expands the proteome without creating or acquiring new genes (Wang et al., 2019) and also modulates protein expression via the generation of transcripts containing a premature termination codon (PTC). Aberrant transcripts with PTC are degraded by the nonsense-mediated decay (NMD) process (Yu et al., 2016; Tress et al., 2017). AS produces mRNA variants from pre-mRNAs by removing different combinations of introns, which makes some intron(s) become part of the coding sequence, or using alternative splice codes, causing the removal of some exon(s). his process is carried out utilizing the specific splicing complex, spliceosome. The spliceosome activity is modulated by cis- and trans-acting regulatory factors and cis-acting components include 5'- and 3'-splice sites, surrounding motif of branch point A, and polypyrimidine tracts (Kupfer et al., 2004). AS is also modulated by the trans-acting regulators, which are members of the SR (Ser-Arg) and heterogeneous nuclear RNP (hnRNP) protein families that function to recognize the splicing codes (Chen and Manley, 2009; Busch and Hertel, 2012; Jeong, 2017). Resulting mRNA isoforms may exhibit distinct stability and translational efficiency and produce proteins with different cellular localization, structure, and function. Accumulating evidence supports that AS plays crucial tasks in development, tissue differentiation, and response to abiotic and biotic stresses in

most eukaryotes (Staiger and Brown, 2013; Scotti and Swanson, 2016; Wang et al., 2019).

Almost 95% of the human genes appear to be regulated by AS (Pan et al., 2008; Sultan et al., 2008). In *Arabidopsis thaliana*, an RNA-seq analysis showed that 61% of the genes are subjected to AS (Syed et al., 2012). Similarly, high-throughput transcriptome analyses revealed that 33% to 61% of the genes undergo AS in soybean, cotton, maize, and rice (Filichkin et al., 2010; Marquez et al., 2012; Wang et al., 2016; Dong et al., 2018; Feng et al., 2019). However, analyses of genome-wide AS repertoires have been performed in only a few fungal species (Wang et al., 2010; Sibthorp et al., 2013; Gehrmann et al., 2016; Kuang et al., 2017). Human fungal pathogens appeared to increase the extent of AS during infection (Grutzmann et al., 2014; Sieber et al., 2018). AS repertoire during plant infection has been reported in *Pseudoperonospora cubensis, Rhizoctonia solani,* and *Sclerotinia sclerotiorum*, but samples collected at only one time point were used (Burkhardt et al., 2015; Jin et al., 2017; Xia et al., 2017; Ibrahim et al., 2020). Since AS is likely subjected to condition-specific regulation, comparative analyses AS patterns under multiple conditions are needed to understand its role and regulation during pathogenesis.

Here, we characterized AS repertoires in the rice blast fungus *Magnaporthe oryzae* during different stages of rice infection and vegetative growth. This fungus is one of the most scientifically and economically important plant pathogens (Dean et al., 2012) and causes up to 30% of rice yield loss every year (Talbot, 2003). The fungus is hemibiotrophic, switching its lifestyle from biotrophy to necrotrophy

during infection. Genome sequences of many field isolates, transcriptomes under various conditions, and molecular research tools are available (Kim et al., 2019), making *M. oryzae* an excellent model for studying fungal pathogenesis mechanisms. The presence of several genes known to be involved in splicing, including *RBP35*, *MoGrp1*, and *MoHMT1*, suggested that AS-mediated regulation of gene expression likely plays roles in pathogenicity (Franceschetti et al., 2011; Li et al., 2020). However, genome-wide AS profiling using RNA-seq data has not yet been performed in *M. oryzae*. A previous study (Ebbole et al., 2004) only considered the presence and absence of AS in *M. oryzae* due to the low sequencing depth of transcripts.

We previously reported high-quality RNA-seq data from both *M. oryzae* and rice at multiple infection stages (Jeon et al., 2020). Here, we reanalyzed this data set to identify which genes are subjected to AS and the type, abundance, and stage-specific pattern of resulting mRNA isoforms as a foundational resource for evaluating the role of AS in pathogenesis. We identified AS isoforms from the transcriptomes of vegetative tissue (mycelia) and during five infection stages, including pre-penetration (18 hpi), biotrophy (27 hpi and 36 hpi), and necrotrophy (45 hpi and 72 hpi) (Figure 1). We applied the concept of isoform abundance to evaluate whether transcript functionality likely changes under different conditions. Neo-functionalization caused by AS-mediated changes in protein domains seems to occur in some mRNAs during infection, suggesting the potential involvement of AS-mediated transcriptome remodeling during pathogenesis.

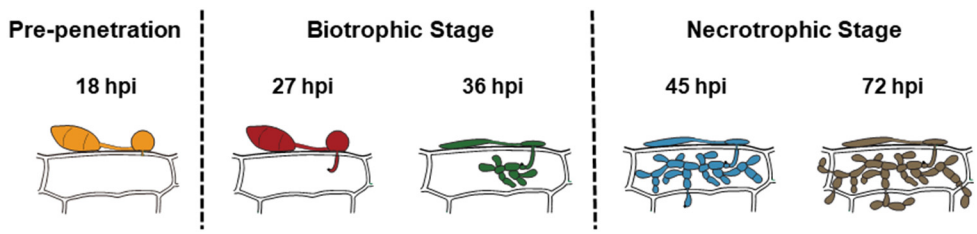


Figure 1. The schematic infection process of the rice blast fungus *M. oryzae*

M. oryzae is a hemibiotrophic pathogen that switches its lifestyle throughout the host cell infection process. Here, the development of fungus is depicted at pre-penetration (18 hpi), biotrophic stage (27-36 hpi), and necrotrophic stage (45-72 hpi).

MATERIALS AND METHODS

I. Genome sequencing and annotation of *M. oryzae* strain KJ201

Genomic DNA of KJ201 was purified using the blood and cell culture DNA midi kit (Qiagen, Germany). Sequencing was performed using a whole-genome shotgun strategy with Illumina Genome Analyzer IIx (Beijing Genome Institute, Shenzhen, China). The genome was assembled using SOAPdenovo 1.05 (Luo et al., 2012) and deposited to NCBI (Accession No. ANSL00000000). The assembled genome was annotated using the Maker 2.31.8 pipeline (Holt and Yandell, 2011). We used the built-in PROMER package of MUMMER v3.23 (Kurtz et al., 2004) to align the 70-15 and KJ201 genome sequences.

II. Transcriptome data analysis

We downloaded published transcriptome datasets generated using KJ201 (Jeon et al., 2020) from NCBI (Accession Nos. SRR8259727 to SRR8259732). The adapter sequences were removed via Cutadapt-1.8.1 using TruSeq universal sequences (Martin, 2011). Truncated reads with low quality at the 3' end and the length shorter than 20 bp were removed. Only the pair mapped sequences to the genome were used to eliminate those with false-positive splicing junctions (single pair and discordantly mapped reads were removed, using the option of --no-mixed --no-discordant). The reads were aligned to the KJ201 genome using HISAT2-2.1.0

(Pertea et al., 2016).

III. Identification of the transcripts generated via AS

A two-step process was used for their identification. Initially, consensus sequences of individual isoforms were identified using reference annotation-based transcript assembly processed by StringTie 1.3.5 (the minimum isoform abundance of 0.01) (Pertea et al., 2015). Information for isoforms, including the type of isoforms, was recorded as General Feature Format (GFF) using ASTALAVISTA 4.0 (Foissac and Sammeth, 2015) and eventGenerator embedded in SUPPA2 (Trincado et al., 2018). Subsequently, the expression value of all genes was calculated using the default settings of Cufflinks 2.2.1 (Trapnell et al., 2010). The transcripts with FPKM value less than one and fused to neighboring genes were eliminated.

IV. Identification of translated isoforms and their predicted functions

The transcript sequences of whole annotated forms and isoforms were retrieved using gffread v0.9.8c from constructed GFF (Pertea and Pertea, 2020). The transcripts with low coding potential were filtered out using CPC 2.0 (Kang et al., 2017). The isoforms with coding potential were classified into NMD candidate transcript (PTC+) and potential protein encoding transcript (PTC-). The PTC position, which is defined as the stop codon within 50 nt upstream of the last exon junction or in the intronic sequence, was calculated to eliminate potential NMD candidates. After removing the NMD candidates, the remaining isoforms were

translated using getORF program (translation from start to stop codon) in EMBOSS v 6.0.0 (Rice et al., 2000). There are several potential translation models which did not start at the start codon of annotated form (Brown et al., 2015). Because we could not determine the authentic translation start site from RNA-seq data and the possibility of leaky scanning translation (Carpenter et al., 2014), we predicted proteins from both the annotated start codon and the novel start codon from the longest ORF. First, we collected the proteins which started from the same start codon of annotated gene. Second, we used the longest ORF in transcripts without non-overlapping upstream ORF to eliminate the isoforms covering multiple genes. These two predicted proteomes were merged for further analysis. Putative functional domains of the predicted proteins were identified using PFAM information of InterProScan v68 (Mitchell et al., 2019). Functional category of each gene product was predicted using GO term information of InterProScan v68 (Mitchell et al., 2019). The enrichment of GO term analysis was performed using the web-based tool WEGO 2.0 (http://wego. genomics.org.cn) (Ye et al., 2018). The protein sequences derived from predicted isoforms were used to query the proteins archived in PHI-base using BLASTP-2.2.26.

V. Gene family analysis

Several splicing regulatory factors (SR proteins and hnRNP) were identified. We first identified the gene products containing RNA recognition motif (RRM) using InterProScan v68 (Mitchell et al., 2019). Among these proteins, we collected those carrying RS domain (SR protein) and G rich motif (hnRNP) using the following rules:

N-terminal RRMs and downstream RS domain of at least 50 amino acids with >40% R[S/D] sequence and G-rich motif characterized by consecutive repeats (Barta et al., 2010; Busch and Hertel, 2012). Putative transcription factors (TFs) were identified using the Fungal Transcription Factor Database (FTFD) pipeline (Park et al., 2008) (Table 1). Putative kinases were identified using Hidden Markov Models of the protein sequences from Superfamily (v1.75), Kinomer, and Microbial Kinome (Kannan et al., 2007; Miranda-Saavedra and Barton, 2007; Wilson et al., 2009) (Table 2). Secreted proteins were identified using the Fungal Secretome Database (Choi et al., 2010). We collected all signal peptide-containing proteins and assessed their likelihood of secretion using SignalP 5.0b (Almagro Armenteros et al., 2019) and TMHMM-2.0c (Moller et al., 2001). The longest ORF sequences without upstream uORF were used for this analysis. Identification of putative effectors, including those with the size of ≤300 amino acids and excluding putative enzymes, was performed as previously described (Kim et al., 2016).

VI. Validation of the production of AS transcripts and their translation

We validated the production of highly expressed isoforms in infected rice plants and also checked that they are translated to produce proteins. Rice plants were grown in a growth chamber set at 28°C and 80% humidity with 16 h-light/8 h-dark photoperiod. Four-week-old rice cultivar Nakdong was inoculated with KJ201 conidial suspension (5×104 conidia/mL in 250 ppm tween20) using a sprayer. The inoculated plants were incubated for 3 and 6 days, including 12 hours in dew

chamber. Total RNAs were isolated from infected leaves using the Easy-spin total RNA extraction kit (iNtRON Biotechnology, Korea) according to the manufacturer's instructions. First-strand cDNA was synthesized from 2 µg of total RNAs using the ImProm-II Reverse Transcription System (Promega) with oligo (dT) primers. Real-time PCR (RT-PCR) reactions were performed using i-star-max II premix (iNtRON Biotechnology, Korea) and primers designed to amplify tested isoforms (Table 3). Each reaction consisted of PCR master mix, 25 ng of cDNA, and 15 pmol of each primer. The thermal cycling conditions were 10 min at 94°C followed by 30 cycles of 15 s at 94°C and 1 min at 60°C. Poly-Acrylamide Gel Electrophoresis (PAGE) was performed to analyze RT-PCR products.

We analyzed the proteomes extracted from infected rice leaves (7 dpi) and mock sample (0 hpi). Protein digestion was conducted using a filter-aided sample preparation (FASP) approach as previously reported (Wisniewski et al., 2009). The collected peptides were dissolved in solvent-A (2% acetonitrile (ACN) in water (v/v) with 0.1% formic acid). The solution separation conducted by reversed-phase chromatography using a UHPLC Dionex UltiMate ® 3000 (Thermo Fisher Scientific, MA, USA). The LC analytical gradient was run 90 min at 2% to 35% solvent B (100% ACN and 0.1% formic acid). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed with an electrospray ionization using QExactive™ Hybrid Quadrupole-Orbitrap High-Resolution Mass Spectrometer (Thermo Fisher Scientific, MA, USA). The peptides were electro-sprayed through a coated silica tip (Scientific Instrument Service, NJ, USA) at an ion spray voltage of 2000 eV, and MS spectra were collected at a resolution of 70,000 (200 m/z) in a mass range of 350-

1650 m/z. The MaxQuant (version 1.5.3.30) program was used for database searching (Tyanova et al., 2016; Gupta et al., 2018). All three technical replicates were merged to profile the proteins.

Table 1. Classification of transcriptional factors identified in this study

No of genes	Family
131	Zn2Cys6
118	Zinc finger, CCHC-type
79	C2H2 zinc finger
44	Nucleic acid-binding, OB-fold
43	HMG
31	Homeodomain-like
23	Winged helix repressor DNA-binding
18	bZIP
17	Myb
8	bHLH
8	Transcription factor jumonji
7	Heteromeric CCAAT factors
5	Zinc finger, DHHC-type
4	p53-like transcription factor
4	Lambda repressor-like, DNA-binding
3	APSES
3	NDT80/PhoG like DNA-binding
3	Zinc finger, MIZ-type
3	AT-rich interaction region
3	Helix-turn-helix, AraC type
3	Zinc finger, Rad18-type putative
3	Bromodomain transcription factor
2	Centromere protein B, DNA-binding region
2	ssDNA-binding transcriptional regulator
2	Zinc finger, GRF-type
2	Transcription factor TFIIS
1	Grainyhead/CP2
1	YL1 nuclear protein
1	Negative transcriptional regulator
1	Tubby transcription factors
1	SGT1
1	SART1
1	RFX DNA-binding domain
1	Zinc finger, NF-X1-type
1	Zinc finger, BED-type predicted
1	CCR4-Not complex component, Not1
1	DDT
1	Helix-turn-helix type 3
1	Iron dependent repressor

Table 2. Classification of kinases identified in this study

No of genes	Subfam	Family	Kinomer
51	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	TKL
34	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	CAMK
25	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	CMGC
19	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	Alpha
18	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	STE
16	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	AGC
11	P-loop containing nucleoside triphosphate hydrolases	Nucleotide and nucleoside kinases	Alpha
7	Protein kinase-like (PK-like)	APH phosphotransferases	TKL
7	Homodimeric domain of signal transducing histidine kinase	Signal transducing histidine kinase	Alpha
6	P-loop containing nucleoside triphosphate hydrolases	Nucleotide and nucleoside kinases	TKL
4	Protein kinase-like (PK-like)	Phoshoinositide 3-kinase (PI3K), catalytic domain	PIKK
4	P-loop containing nucleoside triphosphate hydrolases	Nucleotide and nucleoside kinases	CAMK, CMGC, RGC, TK
4	P-loop containing nucleoside triphosphate hydrolases	Phosphoribulokinase/pantothenate kinase	Alpha, TKL
4	Cysteine-rich domain	Protein kinase cysteine-rich (cys2, phorbol-binding)	AGC, Alpha, TKL
3	Homodimeric domain of signal transducing histidine kinase	Signal transducing histidine kinase	TKL
3	P-loop containing nucleoside triphosphate hydrolases	Gluconate kinase	TKL
3	P-loop containing nucleoside triphosphate hydrolases	Shikimate kinase (AroK)	TKL
3	ATPase domain of HSP90 chaperone/DNA topoisomerase II/histidine kinase	alpha-ketoacid dehydrogenase kinase, C-terminal domain	PDHK
3	P-loop containing nucleoside triphosphate hydrolases	Phosphoribulokinase/pantothenate kinase	CAMK, CMGC, TK
3	P-loop containing nucleoside triphosphate hydrolases	6-phosphofructo-2- /fructose-2,6-bisphosphatase, kinase	Alpha, RIO, TKL
3	Ribokinase-like	Ribokinase-like	Alpha, RIO, TK

3	ARM repeat	Phoshoinositide 3-kinase (PI3K) helical domain	Alpha, PIKK, TKL
3	NAD kinase/diacylglycerol kinase-like	NAD kinase-like	Alpha, PDHK, TKL
3	Protein kinase-like (PK-like)	MHCK/EF2 kinase	Alpha
2	Actin-like ATPase domain	Glycerol kinase	TKL
2	DAK1/DegV-like	DAK1	TKL
2	Ribokinase-like	Ribokinase-like	TKL
2	Ribokinase-like	Thiamin biosynthesis kinases	TKL
2	Ribosomal protein S5 domain 2-like	GHMP Kinase, N-terminal domain	TKL
2	SAICAR synthase-like	Phosphatidylinositol phosphate kinase IIbeta	TKL
2	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	TK
2	P-loop containing nucleoside triphosphate hydrolases	Nucleotide and nucleoside kinases	STE
2	Casein kinase II beta subunit	Casein kinase II beta subunit	CK1,TK
2	Protein kinase-like (PK-like)	Protein kinases, catalytic subunit	CK1
2	Riboflavin kinase-like	ATP-dependent riboflavin kinase-like	CAMK, TKL
2	Carbamate kinase-like	PyrH-like	Alpha, TKL
2	Protein kinase-like (PK-like)	Choline kinase	Alpha, TKL
2	SAICAR synthase-like	Inositol polyphosphate kinase (IPK)	Alpha, TKL
2	P-loop containing nucleoside triphosphate hydrolases	Adenosine-5'phosphosulfate kinase (APS kinase)	Alpha, STE
2	GHMP Kinase, C-terminal domain	Mevalonate kinase	Alpha, RIO
2	Actin-like ATPase domain	Glycerol kinase	Alpha, PDHK
2	NAD kinase/diacylglycerol kinase-like	Diacylglycerol kinase-like	Alpha, CAMK
2	P-loop containing nucleoside triphosphate hydrolases	Shikimate kinase (AroK)	Alpha
2	Protein kinase-like (PK-like)	APH phosphotransferases	Alpha
1	Actin-like ATPase domain	Acetokinase-like	TKL
1	Nucleoside diphosphate kinase, NDK	Nucleoside diphosphate kinase, NDK	TKL
1	Phosphoglycerate kinase	Phosphoglycerate kinase	TKL

1	Ribokinase-like	PfkB-like kinase	TKL
1	Ribokinase-like	YjeF C-terminal domain-like	TKL
1	Protein kinase-like (PK-like)	APH phosphotransferases	TK
1	Protein kinase-like (PK-like)	RIO1-like kinases	RIO
1	Winged helix DNA-binding domain	Rio2 serine protein kinase N-terminal domain	RIO
1	KA1-like	Kinase associated domain 1, KA1	CAMK
1	P-loop containing nucleoside triphosphate hydrolases	Shikimate kinase (AroK)	CAMK
1	Carbamate kinase-like	N-acetyl-l-glutamate kinase	Alpha
1	Glycerate kinase I	Glycerate kinase I	Alpha
1	HAD-like	phosphatase domain of polynucleotide kinase	Alpha
1	PK beta-barrel domain-like	Pyruvate kinase beta-barrel domain	Alpha
1	P-loop containing nucleoside triphosphate hydrolases	Gluconate kinase	Alpha
1	Ribosomal protein S5 domain 2-like	GHMP Kinase, N-terminal domain	Alpha
1	SAICAR synthase-like	Phosphatidylinositol phosphate kinase IIbeta	Alpha
1	AMPKBI-like	AMPKBI-like	
1	AraD/HMP-PK domain-like	Phosphomethylpyrimidine kinase C-terminal domain	
1	Protein kinase-like (PK-like)	MHCK/EF2 kinase	

Table 3. The primers used in this study

Primer	Strand	Sequence
(MGG_08019T0) maker-scaffold000032-augustus-gene-22.85-mRNA-1	forward	GCCAGCTGCTTCTGACTATGGAAA
Total.5386.4	reverse	TTCTCCACAGTGCTCGTCCTG
maker-scaffold000094-snap-gene-8.57-mRNA-1	forward	TGGACAGGATAACACCCACGATAC
Total.10136.2	reverse	GATGCGGCAAACGTCGAATAAGAG
maker-scaffold000032-snap-gene-16.0-mRNA-1	forward	GGCCACGGACATGTTCCTCATCTA
Total.5261.3	reverse	CGTACCAGTAGGTCATGAGAAGGAG
(MGG_11426T0) maker-scaffold000002-augustus-gene-3.119-mRNA-1	forward	AAGCCAAATCCAAATGGCGAAGAG
Total.685.3	reverse	GCGATCGTTGCTGGTTCATTTTGG
maker-scaffold000062-snap-gene-6.62-mRNA-1	forward	CAACACGGTCGCCATCTCGTTAAA
Total.7087.8	reverse	CGTATCCTGTGCCTTGCGAGTAA
(MGG_07681T0) maker-scaffold000077-snap-gene-2.60-mRNA-1	forward	CAGCACTGGGGTAGCTGCTTA
Total.8570.2	reverse	CCGTGGTGATGTCCAGCGTAA
(MGG_01368T0) maker-scaffold000097-augustus-gene-8.41-mRNA-1	forward	GCCTCAGCCGATAAGCTCATCAAG
Total.12011.2	reverse	AGGTGCTGTCTGGGTGGTTC
maker-scaffold000096-augustus-gene-3.128-mRNA-1	forward	CTCCGAGGTACCTGTTGACGA
Total.11556.3	reverse	CGGTTGGAAGGTCGTTGTAGTCG
(MGG_11132T0) maker-scaffold000060-augustus-gene-5.73-mRNA-1	forward	GTAGCCTCGGGGAATCTGGTA
Total.6116.2	reverse	CGAGGTCTATGCTCTGCACAAGTA
(MGG_17060T0) maker-scaffold000061-snap-gene-12.79-mRNA-1	forward	CGATGAGCACAATCGCCATAAGCGA
Total.6717.5	reverse	GACGCAGTGGAGCTTAAGCCTA
B-tublin_F	forward	ACAACTTCGTCTTCGGTCAG
B-tublin_R	reverse	GTGATCTGGAAACCCTGGAG

MGG gene numbers represent ortholog gene in 70-15 genome v8

RESULTS

I. Identification of the *M. oryzae* genes containing multi-exons

We previously reported RNA-seq datasets derived from KJ201, an *M. oryzae* field isolate from infected rice, during infection (Jeon et al., 2020). Because we used the genome of 70-15, a laboratory strain created through a series of genetic crosses (Dean et al., 2005), as a reference for mapping transcript sequences of KJ201 in this study, we may have failed to identify some KJ201 transcripts due to genome differences between the two strains. To support a more accurate transcriptome analysis of KJ201, we sequenced the genome of KJ201. The assembled genome, represented by 123 scaffolds with a total length of 41.7 Mb and N50 values of 2.3 Mb, contains 13,306 genes (Table 4). Synteny comparison revealed that the genomes of KJ201 and 70-15 are mostly conserved with a few segmental differences (Figure 2A). However, only 9,951 KJ201 genes (78.6%) have orthologs in 70-15 when their predicted proteomes were compared with >90% sequence identity and coverage as the threshold. Moreover, only 62.5% of the genes showed the same number of intron(s) as their orthologs in 70-15 (Figure 2B). Based on the annotated gene model of KJ201, 10,178 (76.5%) genes are predicted to contain multi-exons. Among other *M. oryzae* isolates, the proportion of multi-exon genes ranged from 64.5 to 80.7% (Figure 3).

Table 4. Statistics of KJ201 genome

Genome statistics	KJ201
Genome size (Mb)	41.7
N50 (bp)	2,318,557
L50	14
No. of scaffolds	123
Number of protein coding genes	13,306

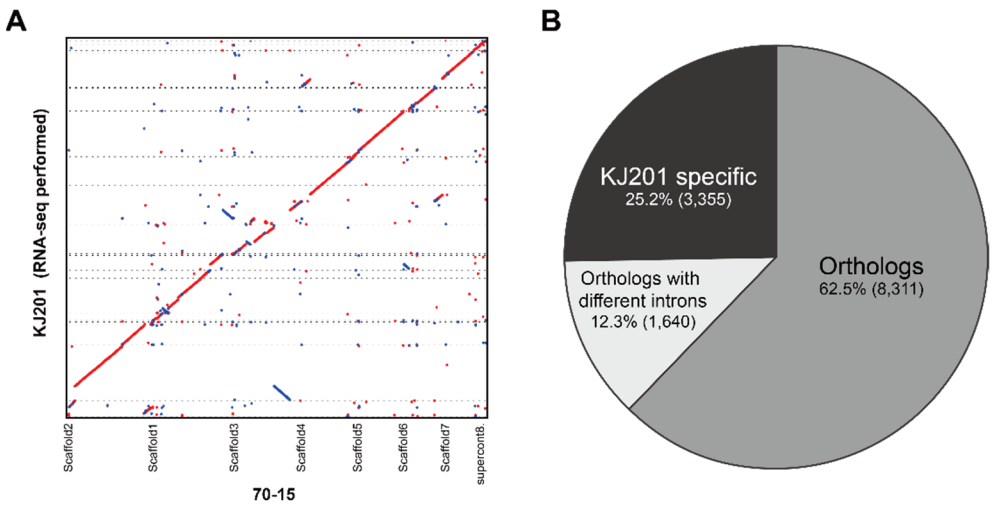


Figure 2. Genome synteny and ortholog between KJ201 and 70-15

(A) Synteny dot plot showing the aligned genomes of KJ201 and 70-15. The x-axis corresponds to the 70-15 genome, and the y-axis denotes the KJ201 genome. (B) A summary of results from bidirectional blast between the KJ201 and 70-15 proteomes. The pie graph shows three classes of 13,306 KJ201 annotations, including orthologs with identical intron(s) (Orthologs), orthologs with different intron numbers (Orthologs with different intron(s)), and proteins displaying lower than 90% identity (KJ201 specific).

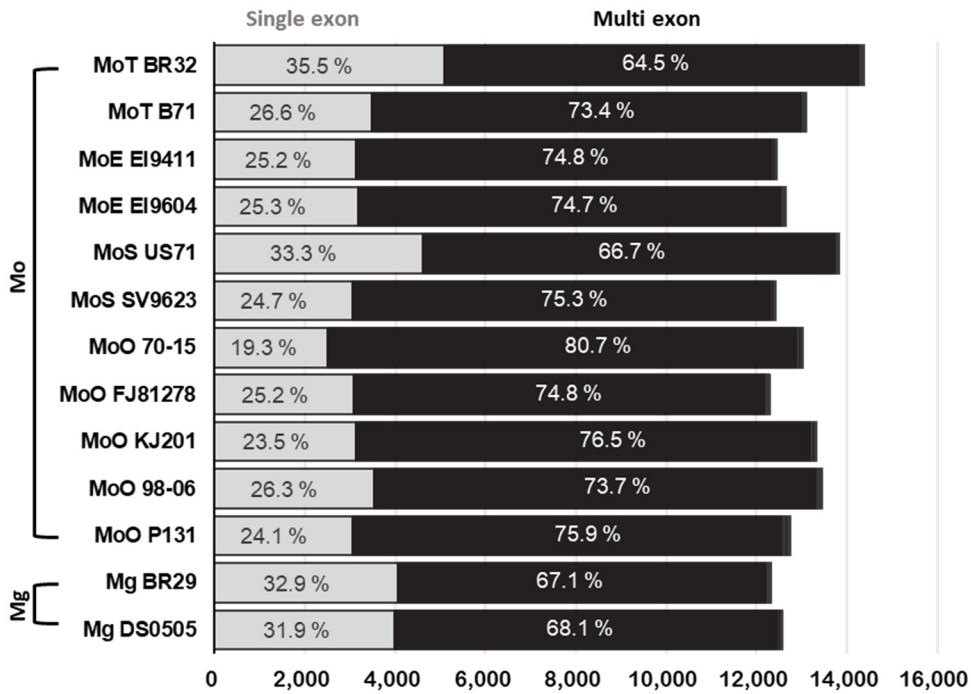


Figure 3. Single- and multi-exon gene distribution patterns among the genomes of diverse *Magnaporthe* isolates

The proportions of single- and multi-exon genes in the genomes of selected isolates of *M. oryzae* (Mo) and *M. grisea* (Mg) are shown. The grey and black boxes indicate single- and multi-exon genes, respectively.

II. Analysis of AS patterns in KJ201 during vegetative growth and infection

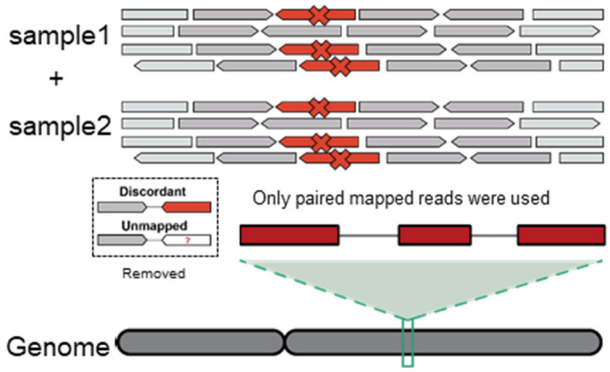
We analyzed AS profiles in KJ201 using our published RNA-seq data (Jeon et al., 2020) with the KJ201 genome as the reference (Table 5). The transcripts potentially generated via AS were identified using a two-step pipeline (Figure 4) (Angelini et al., 2014). We initially identified 2,772 candidate genes with their transcripts undergoing AS under one or more conditions. After filtering out those expressed at a very low level and chimeric transcripts, we found 4,270 novel mRNA isoforms from 2,413 annotated genes (18.1%) (Figure 5A). AS patterns in mycelia are distinct from those in infected rice at all expression thresholds (Figure 5B; Figure 6). In total, 1,914 genes produced AS isoforms in mycelia. AS was observed in transcripts of 2,135 genes at the stage of appressorium formation (18 hpi), 2,153 and 2,127 genes at two time points during the biotrophic stage (27 and 36 hpi, respectively), and 2,118 and 2,127 genes at 45 and 72 hpi, respectively (necrotrophic stage). A Principal Component Analysis (PCA) of the annotated transcripts and their isoforms (Figure 6) resulted in three distinct clusters.

Table 5. Statistics of KJ201 transcripts analyzed

	Mycelia	18 hpi	27 hpi	36 hpi	45 hpi	72 hpi
Proportion of fungal reads (%)	93.33	6.83	4.08	4.41	7.43	25.65
No. of mapped Reads (x1,000)	40,608	15,763	8,378	9,441	16,565	53,444
Depth (X)	204.6	90.5	54.2	60.9	96.2	247.8

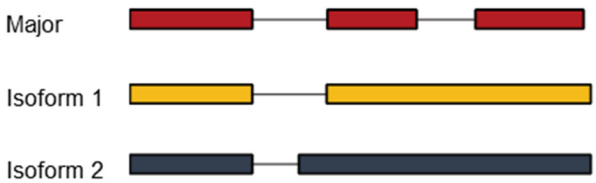
Collection of total reads for analysis

HISAT2



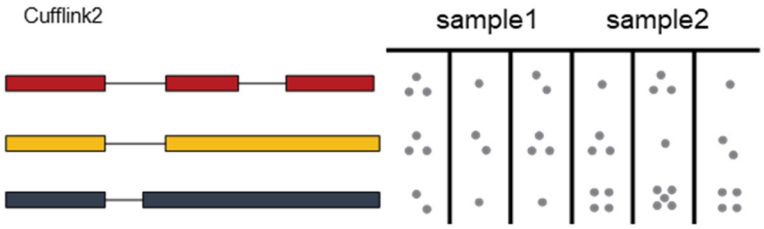
Identification of consensus isoforms

Stringtie, SUPPA2, ASTALAVISTA

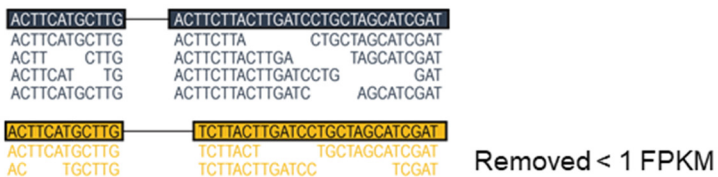


Re-estimation of isoform expression levels under different conditions

Cufflink2



Removal of the transcripts expressed at low levels



Comparison of expression patterns of individual genes under different conditions

Figure 4. Two-step pipeline used for identifying and analyzing transcripts generated via AS

In the first step, all RNA-seq reads from the samples representing different stages of *M. oryzae* life cycle are mapped to the genome of KJ201 to detect novel transcript variants using HISAT2. From the mapping data, consensus isoforms are identified using StringTie, and AS is detected using SUPPA2. In the second step, expression levels of individual isoforms as FPKM (Fragments Per Kilobase of transcripts per Million mapped reads) are calculated using Cufflinks2. After filtering out those expressed low levels (the ones with FPKM <1), the resulting data were subjected to comparative analyses.

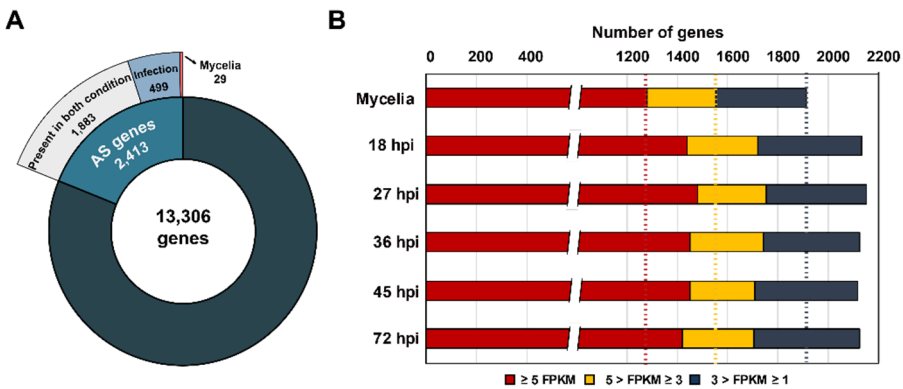


Figure 5. Characteristics of AS in *M. oryzae* KJ201 under different conditions

(A) The numbers and proportions of predicted AS genes expressed under different conditions are shown. The total number of annotated genes in KJ201 is 13,306. Among 2,413 AS genes, transcript variants from 29 and 499 genes were detected only in mycelia and during infection, respectively. The remaining 1,883 genes produced AS transcripts under both conditions. (B) The number of AS genes in each sample predicted using three FPKM values for cutoff.

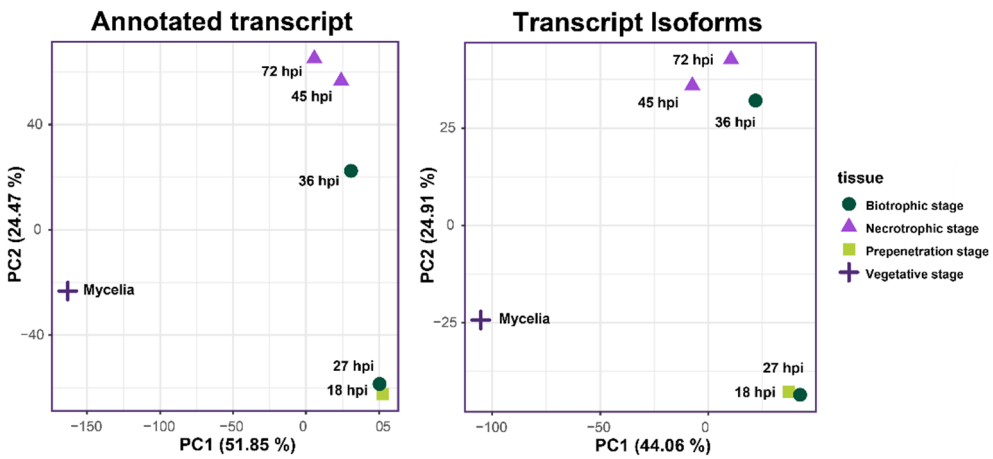


Figure 6. PCA analysis of the annotated transcripts and their isoforms

The log2 FPKM values of individual transcripts within each sample are used to calculate PCA distances. Each sample calculated using standard PCA function implemented in prcomp was used.

III. Vegetative- and infection-specific AS isoforms and the expression pattern of putative AS regulators

We compared the AS isoforms detected under different conditions to determine how the production of transcript variants is regulated at different growth and infection stages (Figure 7A). Among the 4,270 isoforms identified, 2,095 (49.1%) isoforms produced from 1,610 genes were present under all conditions. The total number of isoforms produced during infection was 1,368 (32.7%) from 955 genes, including 696 isoforms of 499 genes that were solely produced during infection. The number of vegetative stage-specific isoforms was 149, which were produced by 125 genes. The highest numbers of stage-specific isoforms were detected during the biotrophic stage (109 (104 genes) and 102 (98 genes) isoforms at 27 and 36 hpi, respectively). It was found that the proportion of infection specific isoforms was increased with the application of robust criteria (Table 6; Figure 8)

We also analyzed how the genes for putative AS regulators, including seven SR and 11 hnRNP proteins, are transcribed and processed under these conditions to explore how AS regulates their expression at each stage. Among these AS regulator genes, ten and six were up- and down-regulated at one or more infection stages, respectively, compared to the vegetative stage (Figure 7B). Five AS regulator genes produced isoforms, and isoforms of three genes displayed different expression patterns (Figure 9).

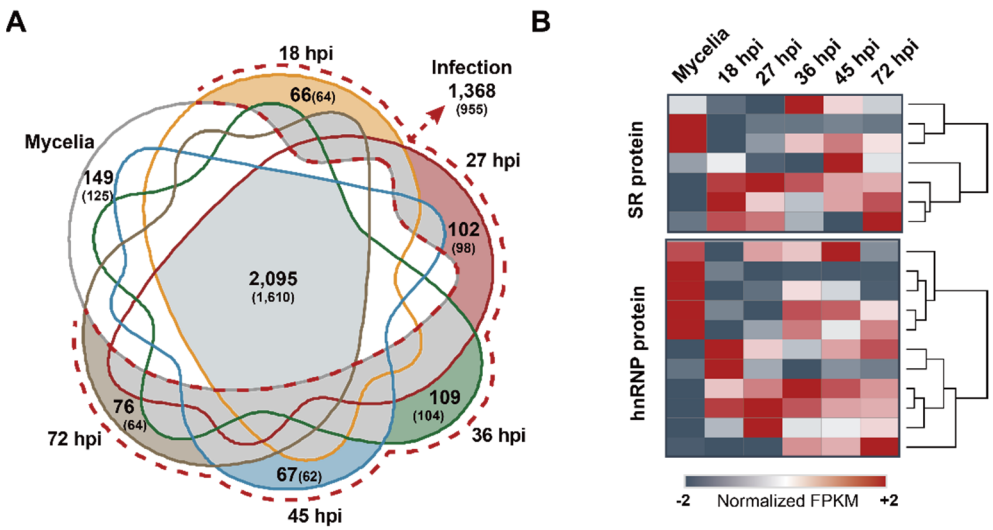


Figure 7. The mycelial growth/infection stage-specific AS repertoires and those produced under both conditions and expression patterns of putative AS regulator genes

(A) Venn diagram showing the number of AS isoforms and genes (in parenthesis) at each stage. The light grey region denotes the AS isoforms produced in mycelia and during infection (2,095). Each stage is color-coded: mycelia (grey), pre-penetration (yellow), biotrophic (red and green), and necrotrophic (blue and brown). The AS isoforms detected only during infection are shown in a dashed red line. (B) Heatmap showing the expression of SR and hnRNP protein-coding genes under different conditions.

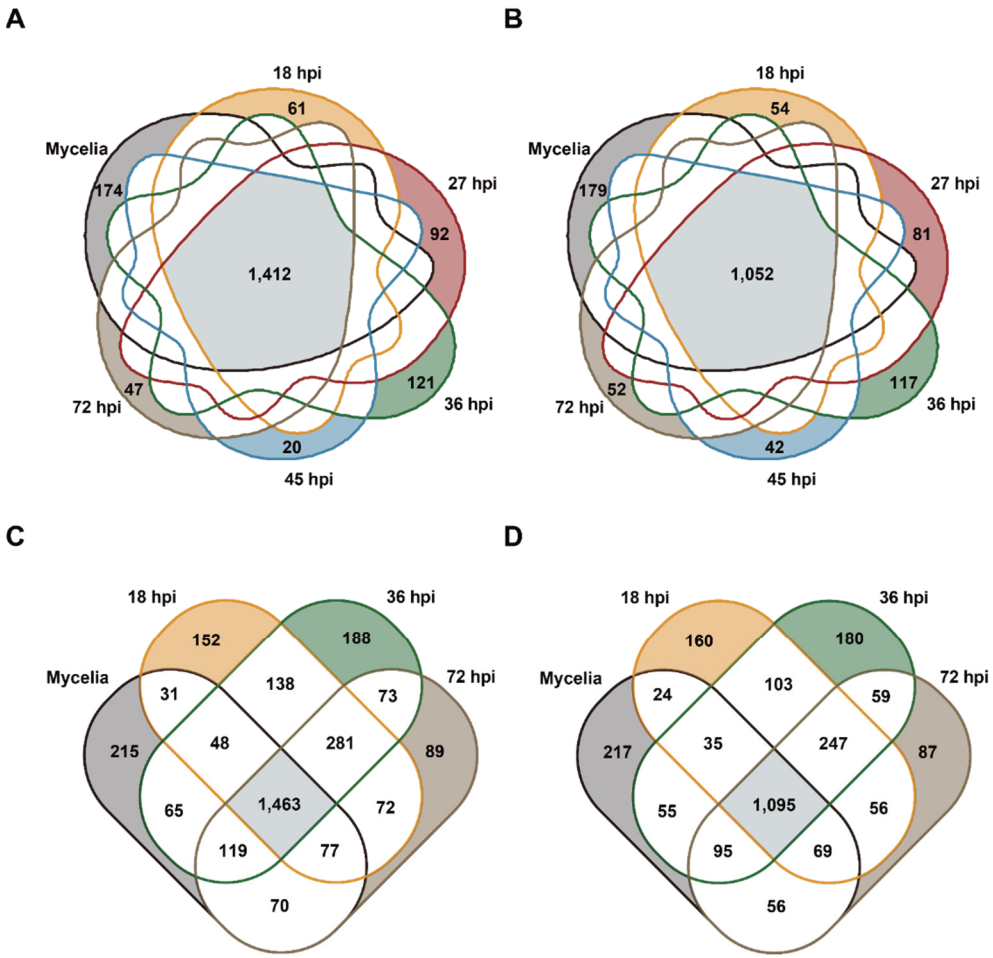


Figure 8. Distribution of AS repertoire by different criteria

Venn diagram showing the number of AS isoforms at each stage. The light grey region denotes the AS isoforms produced in mycelia and during infection. Each stage is color-coded: mycelia (grey), pre-penetration (yellow), biotrophic (red and green), and necrotrophic (blue and brown) (A) Isoforms over 3 FPKM and (B) over 5 FPKM. Venn diagram represents the comparison of vegetative and three representative infection stages (Mycelia; vegetative,18 hpi; pre-penetration, 36 hpi; biotroph, 72hpi; necrotroph) (C) Isoforms over 3 FPKM and (D) over 5 FPKM.

Table 6. Distribution of AS isoforms in different FPKM cutoff criteria

Filtering Criteria (FPKM)	The number of stage-specific AS isoforms						Core	Infection	Total	Core ratio (%)	Infection ratio (%)
	Mycelia	18 hpi	27 hpi	36 hpi	45 hpi	72 hpi					
1	149	66	102	109	67	76	2,095	1,368	4,270	49.06%	32.04%
2	173	66	98	102	50	62	1,673	1,268	3,637	46.00%	34.86%
3	174	61	92	121	47	47	1,412	1,142	3,230	43.72%	35.36%
4	183	61	82	119	41	53	1,208	1,074	2,928	41.26%	36.68%
5	179	54	81	117	42	52	1,052	1,021	2,667	39.45%	38.28%
6	196	49	80	86	34	51	913	954	2,460	37.11%	38.78%
7	188	54	74	73	30	50	816	892	2,275	35.87%	39.21%
8	178	42	67	79	31	48	749	839	2,118	35.36%	39.61%
9	163	43	57	74	33	49	685	812	1,979	34.61%	41.03%
10	158	39	43	67	29	51	621	761	1,858	33.42%	40.96%

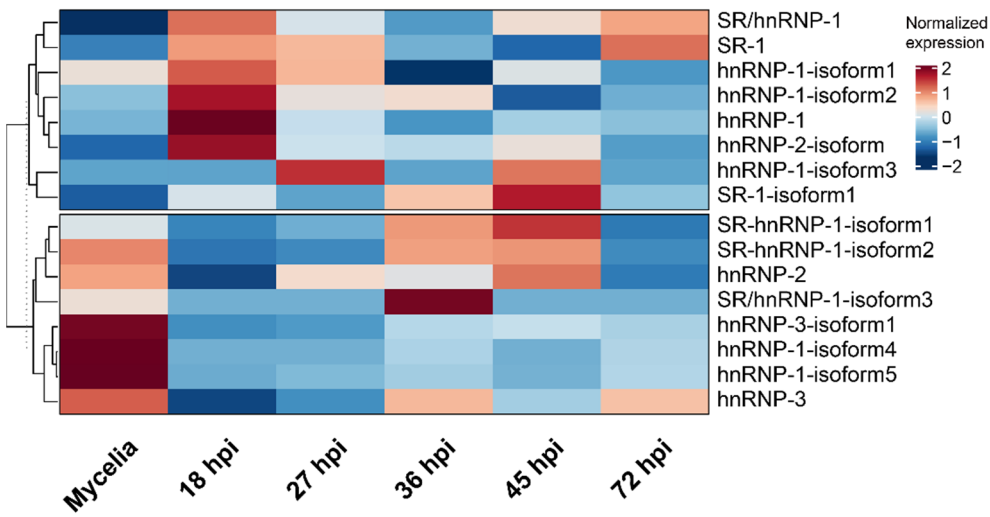


Figure 9. Expression patterns of alternatively spliced transcripts of SR and hnRNP genes

The heatmap shows the expression patterns of putative AS regulatory genes and their isoforms calculated using Z-score normalization across vegetative and infection stages.

Relative expression levels of AS transcripts at each stage were analyzed (Figure 10). We applied the concept of transcript usage, which describes the relative abundance of individual isoforms in a sample relative to one another, to predict which AS genes and isoforms are likely important for infection. The relative abundance of isoforms compared to corresponding annotated mRNAs was calculated using the formula (Fragment Per Kilobase of transcript per Million mapped reads (FPKM) of the isoform)/(FPKM of isoform + FPKM of the annotated form). We group them into three clusters based on their relative abundance pattern: (1) Constitutive AS (0.5 < relative abundance < 1), those that are highly abundant under all conditions, (2) Switching AS, those that fluctuate under these conditions, and (3) Low-frequency AS (relative abundance < 0.5), those that are not abundant under all conditions (Figure 11A). We found that 658 isoforms (559 genes) belong to Constitutive AS. Switching AS included 868 isoforms (636 genes), which were relatively higher at the infection stages than the vegetative stage. Low-frequency AS included 1,949 isoforms. For Switching AS, 126 isoforms (109 genes) were abundant throughout all infection stages. During infection, 200 isoforms (179 genes) at the biotrophic stage (27 – 36 hpi), 117 isoforms (108 genes) at the necrotrophic stage (45 – 72 hpi), and 66 isoforms (65 genes) at the pre-penetration stage (18 hpi) were identified (Figure 11B; Table 7). The most abundant isoforms during the biotrophic stage belonged to Switching AS. Furthermore, the result of Tau (τ) coefficient value to confirm time specificity at infection stage the value of AS isoforms showed little different value in all samples (Figure 12).

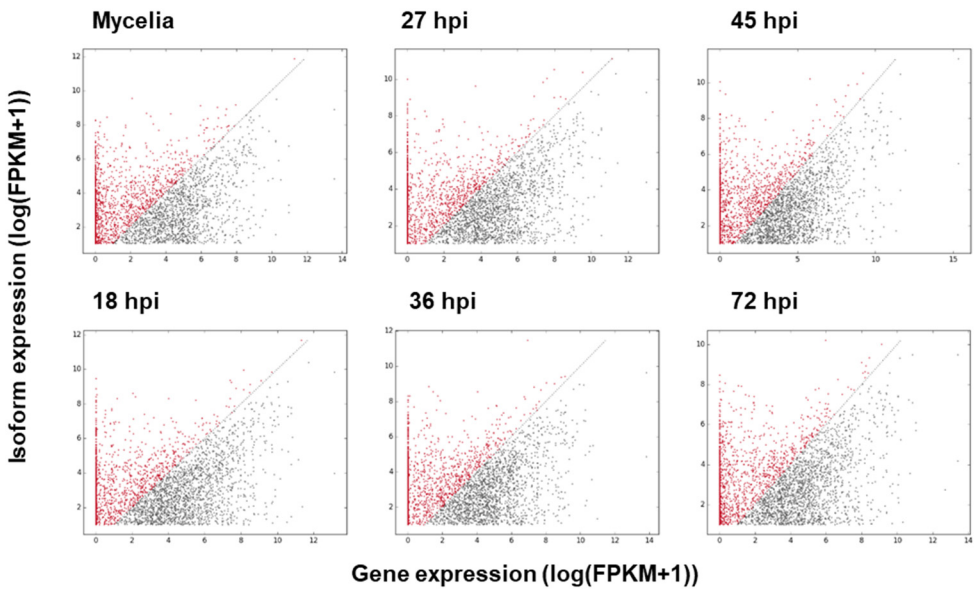
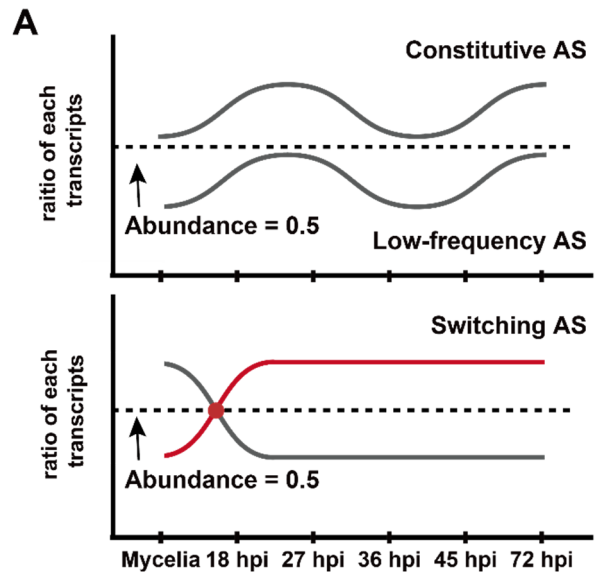
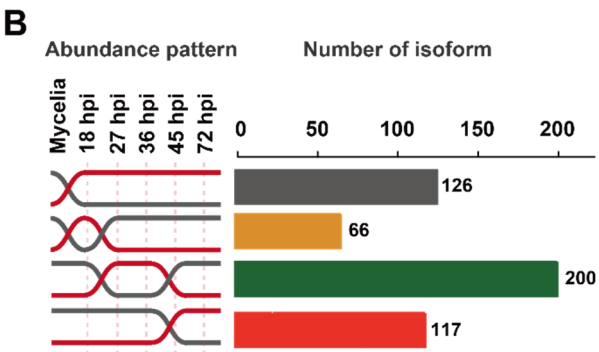


Figure 10. Relative expression levels of AS transcripts under each condition

The scatterplots show the patterns of expression between annotated transcripts and isoforms in each sample. The X-axis shows the values calculated using log(FPKM of annotated form +1), and the Y-axis shows the values calculated using log(FPKM of isoform +1). The red dots represent the isoforms with a ratio of >0.5.


A
Constitutive AS
raitio of each transcripts
Abundance = 0.5
Low-frequency AS
Switching AS
ratio of each transcripts
Abundance = 0.5
Mycelia 18 hpi 27 hpi 36 hpi 45 hpi 72 hpi


B
Abundance pattern
Number of isoform
Mycelia
18 hpi
27 hpi
36 hpi
45 hpi
72 hpi
0
50
100
150
200
126
66
200
117

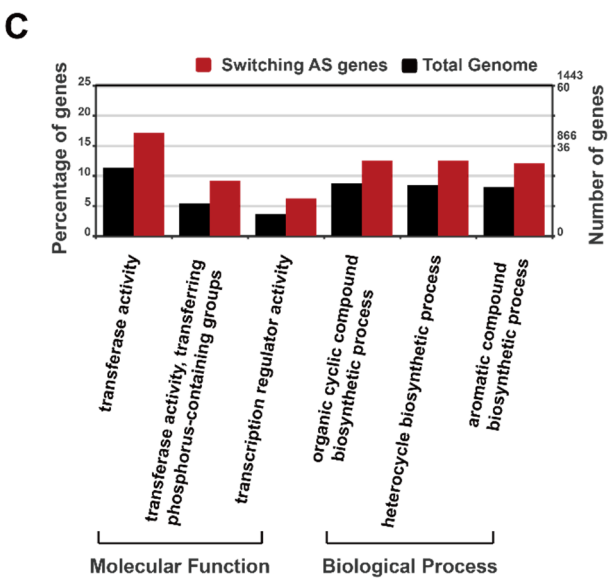

C
Switching AS genes
Total Genome
Percentage of genes
Number of genes
25
20
15
10
5
0
1443
60
866
36
0
transferase activity
transferase activity, transferring phosphorus-containing groups
transcription regulator activity
organic cyclic compound biosynthetic process
heterocycle biosynthetic process
aromatic compound biosynthetic process
Molecular Function
Biological Process

Figure 11. Expression types of AS isoforms and characteristics of Switching AS isoforms

(A) The diagrams show three types of AS transcripts categorized based on their relative abundance during the growth stages analyzed. The abundance indicates the ratio between isoforms and annotated forms. Constitutive AS and Low-frequency AS isoforms are classified based on relative abundances of > 0.5 and < 0.5, respectively. Switching AS isoforms include the isoforms with relative abundance > 0.5 and solely expressed during one or more infection stages. (B) Four groups of Switching AS isoforms and their expression patterns during infection. Grey box (the isoforms abundantly produced in all infection stages), yellow box (those produced higher than corresponding annotated forms only during the pre-penetration stage, 18 hpi)), green box (those produced higher than corresponding annotated forms only during the biotrophic stage, 25 hpi or 36 hpi), and red box (those produced higher than corresponding annotated forms only during the necrotrophic stage, 45 hpi or 72 hpi). (C) The GO terms enriched among the Switching AS genes (P-value < 0.05).

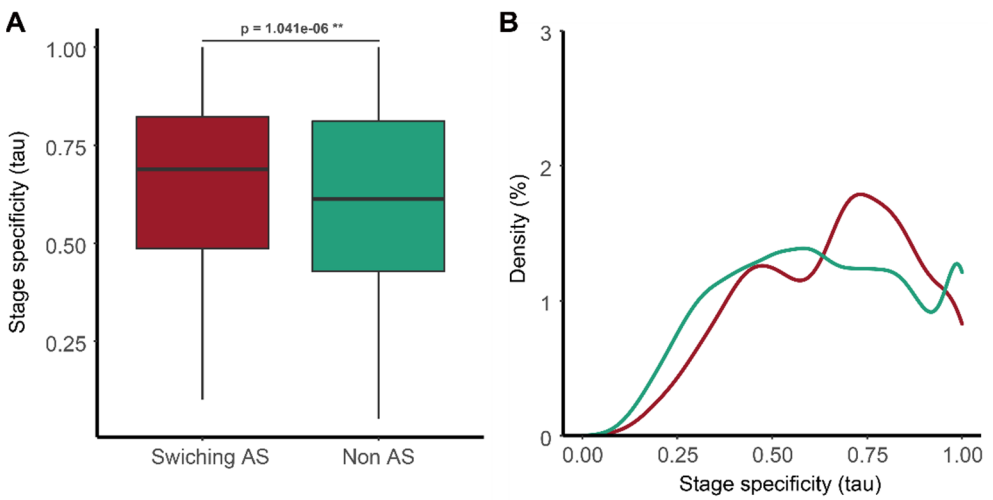


Figure 12. Distribution of tau coefficient of AS producing genes and other genes

(A) Boxplot representing the Tau coefficient value of each transcript. Tau value calculated ((Number of sample - Sum of expression in all stages)/Maximum of expression value)/(Number of sample -1) in each transcript. Switching AS isoforms include the isoforms with relative abundance > 0.5 and are solely expressed during one or more infection stages. Non AS represents genes which were not identified isoforms. Statistically significant differences between group means of tau distribution were determined by Wilcoxon–Mann–Whitney test ($p < 0.01$). (B) Tau value distribution of Switching AS (red) and Non AS genes (green).

Gene Ontology (GO) enrichment analysis was performed to predict likely functions of the genes belonging to Switching AS. They are mostly predicted to be involved in heterocyclic compound binding, ion binding, and/or hydrolase activity (Table 7). Significantly enriched GO terms, compared to the whole gene set, included phospho-transferase, transcription regulator activities, and cyclic compound biosynthetic processes (Figure 11C). We investigated the GO term-associated patterns at each stage. Only the necrotrophic stage showed enriched terms (heterocyclic compound and ion bindings) (Table 8). We performed an in-depth analysis of those associated with infection-enriched phospho-transferases and transcription regulator activities by mining representative gene families. We found 70 kinase genes belonging to phospho-transferases produce AS isoforms. Among them, 20 genes generate Switching AS isoforms (Figure 13A), but no subfamilies were enriched compared to the total kinases (Figure 13B). Among 139 genes that were predicted to encode TFs and also produced AS isoforms (Figure 14A), C_2H_2 zinc finger (28.1%, p=0.04415) and High Mobility Group (HMG; 18.8%, p=0.04427) were enriched in 35 Switching AS isoforms (Figure 14B). We checked whether the proteins encoded by those belonging to Switching AS likely participate in pathogenesis by matching them to the pathogenicity-related genes in PHI-base (Urban et al., 2020) (Table 9). Homologs of 39 genes have been functionally studied (Table 10). Among them, 22 genes, including oxidative stress-related (*ABC1*, *RanBP*, *TIG1*, *MoARK1*) and autophagy-related (*MoAtg1*, *MoAtg2*), have been shown to be involved in pathogenesis.

Table 7. Distribution of gene ontology (GO) in infection-switching genes of *M. oryzae*

GO ID	No. of genes	P-value	Function
GO:1901363	74	0.849	heterocyclic compound binding
GO:0097159	74	0.849	organic cyclic compound binding
GO:0043167	72	0.718	ion binding
GO:0016787	43	0.584	hydrolase activity
GO:0016740	41	0.006	transferase activity*
GO:0036094	40	0.947	small molecule binding
GO:0097367	32	0.513	carbohydrate derivative binding
GO:0005515	30	0.551	protein binding
GO:0016491	29	0.699	oxidoreductase activity
GO:0008144	26	0.833	drug binding
GO:0140096	24	0.142	catalytic activity, acting on a protein
GO:0048037	18	0.491	cofactor binding
GO:0003700	14	0.064	DNA-binding transcription factor*
GO:0022857	13	0.899	transmembrane transporter activity
GO:0140098	5	0.99	catalytic activity, acting on RNA
GO:0016829	5	-	lyase activity
GO:0016874	3	-	ligase activity
GO:0140097	3	-	catalytic activity, acting on DNA
GO:0030234	2	-	enzyme regulator activity
GO:0005085	2	-	guanyl-nucleotide exchange factor
GO:0008289	1	-	lipid binding
GO:0072341	1	-	modified amino acid binding
GO:0033218	1	-	amide binding
GO:0005199	1	-	structural constituent of cell wall
GO:0003712	1	-	transcription coregulator activity
GO:0005319	1	-	lipid transporter activity
GO:0004362	1	-	glutathione-disulfide reductase activity

Asterisk represents enriched GO terms compared to total proteome

Table 8. Distribution of gene ontology (GO) in Switching AS genes of necrotrophic stages

GO ID	No. of genes	P-value	Function
GO:0097159	25	0.002	organic cyclic compound binding*
GO:1901363	25	0.002	heterocyclic compound binding*
GO:0043167	21	0.024	ion binding*
GO:0036094	11	0.234	small molecule binding
GO:0005515	10	0.164	protein binding
GO:0016787	9	0.687	hydrolase activity
GO:0097367	9	0.148	carbohydrate derivative binding
GO:0016740	7	0.481	transferase activity
GO:0008144	6	0.637	drug binding
GO:0003700	5	-	DNA-binding transcription factor
GO:0140096	4	-	catalytic activity, acting on a protein
GO:0016491	4	-	oxidoreductase activity
GO:0140098	3	-	catalytic activity, acting on RNA
GO:0048037	3	-	cofactor binding
GO:0016874	1	-	ligase activity
GO:0022857	1	-	transmembrane transporter activity
GO:0030234	1	-	enzyme regulator activity
GO:0005085	1	-	guanyl-nucleotide exchange factor
GO:0003712	1	-	transcription coregulator activity

Asterisk represents enriched GO term compared to total proteome

Table 9. PHI-base assigned genes in total AS repertoire

PHI-base	Alias	Phenotype
PHI:2051*	*COS1*	increased virulence (hypervirulence)
PHI:8893	*RGS1*	increased virulence (hypervirulence), reduced virulence
PHI:2161*	*SEP1*	loss of pathogenicity
PHI:2176*	*ASD4*	loss of pathogenicity
PHI:2114	Calnexin	loss of pathogenicity
PHI:2098	Ca-transporting ATPase3	loss of pathogenicity
PHI:2302	*CHS5*	loss of pathogenicity
PHI:36	*CPKA*	loss of pathogenicity
PHI:3307	*gpf1*	loss of pathogenicity
PHI:777	MGG_00883	loss of pathogenicity
PHI:5232	*MoARG1*	loss of pathogenicity
PHI:2035*	*Moatg1*	loss of pathogenicity
PHI:2070*	*Moatg2*	loss of pathogenicity
PHI:2071	*Moatg3*	loss of pathogenicity
PHI:2182	*Moatg4*	loss of pathogenicity
PHI:2078	*Moatg10*	loss of pathogenicity
PHI:7117	*MoGSK1*	loss of pathogenicity
PHI:9494	*MoGT2*	loss of pathogenicity
PHI:3161	*Mohik5*	loss of pathogenicity
PHI:3975	*MoIlv2*	loss of pathogenicity
PHI:8928	*MoSCAD2*	loss of pathogenicity
PHI:3174	*MoSip2*	loss of pathogenicity
PHI:3175	*MoSnf4*	loss of pathogenicity
PHI:2180	*MoTea4*	loss of pathogenicity
PHI:4736	*MoYpt7*	loss of pathogenicity
PHI:113	*MPS1*	loss of pathogenicity
PHI:417	*MST7*	loss of pathogenicity
PHI:2050	*OMO1*	loss of pathogenicity
PHI:56	*PMK1*	loss of pathogenicity
PHI:120	*PTH2*	loss of pathogenicity
PHI:4492*	*RanBP*	loss of pathogenicity
PHI:2002*	*TIG1*	loss of pathogenicity
PHI:2115	*AnnA7*	loss of pathogenicity, reduced virulence
PHI:2142	*MoAP1*	loss of pathogenicity, reduced virulence
PHI:2122	*MoHox8, MST12*	loss of pathogenicity, reduced virulence
PHI:2200	*MoSLN1*	loss of pathogenicity, reduced virulence
PHI:2044	*MoSOM1*	loss of pathogenicity, unaffected pathogenicity
PHI:132*	*ABC1*	reduced virulence

PHI:6607	*ARRDC1*	reduced virulence
PHI:519	*CHM1*	reduced virulence
PHI:2116*	*CHS7*	reduced virulence
PHI:3309*	*cnf2*	reduced virulence
PHI:2017	*COM1*	reduced virulence
PHI:35	*CON7*	reduced virulence
PHI:3315*	*conx1*	reduced virulence
PHI:2167*	*EXP5*	reduced virulence
PHI:256	*GAS1*	reduced virulence
PHI:3816*	*GSN1*	reduced virulence
PHI:305	*ICL1*	reduced virulence
PHI:9356	*leu4*	reduced virulence
PHI:81	*MAC1*	reduced virulence
PHI:83	*MAGB*	reduced virulence
PHI:2034	*MFP1*	reduced virulence
PHI:805	MGG_00124	reduced virulence
PHI:888	MGG_01707	reduced virulence
PHI:873	MGG_02443	reduced virulence
PHI:795*	MGG_03451	reduced virulence
PHI:799*	MGG_03530	reduced virulence
PHI:881	MGG_04556	reduced virulence
PHI:2996	MGG_06507	reduced virulence
PHI:807	MGG_06951	reduced virulence
PHI:819	MGG_07061	reduced virulence
PHI:790	MGG_07259	reduced virulence
PHI:883	MGG_08560	reduced virulence
PHI:800*	MGG_13324	reduced virulence
PHI:5083	*MoAcat2*	reduced virulence
PHI:3014	*MoAND1*	reduced virulence
PHI:8039	*Moaos1*	reduced virulence
PHI:2940*	*MoARK1*	reduced virulence
PHI:6928	*MoBRE1*	reduced virulence
PHI:2158	*MoCMK1*	reduced virulence
PHI:6780	*MoCreC*	reduced virulence
PHI:6634	*MoDnm1*	reduced virulence
PHI:7223	*Moend3*	reduced virulence
PHI:7112	*MoERR1*	reduced virulence
PHI:5047	*MoFis1*	reduced virulence
PHI:2189	*MoGIS2*	reduced virulence
PHI:7149	*Moglo3*	reduced virulence
PHI:2128	*MoHox1*	reduced virulence

PHI:2133*	*MoHox3*	reduced virulence
PHI:2130	*MoHox4*	reduced virulence
PHI:2134	*MoHox5*	reduced virulence
PHI:5188	*MoHPX1*	reduced virulence
PHI:8695	*MoImd4*	reduced virulence
PHI:5189	*MoLDS1*	reduced virulence
PHI:3234	*MoLYS20*	reduced virulence
PHI:7793	*Momyo2*	reduced virulence
PHI:8754	*MoPer1*	reduced virulence
PHI:7134	*MoRab5B*	reduced virulence
PHI:4616	*MoSFA1*	reduced virulence
PHI:2405	*MoSSK1*	reduced virulence
PHI:5366	*MoSyn8*	reduced virulence
PHI:8041	*Moubc9*	reduced virulence
PHI:4112*	*MoVELB*	reduced virulence
PHI:7385	*MoVps17*	reduced virulence
PHI:5041*	*MoYAK1*	reduced virulence
PHI:860	*MSP1*	reduced virulence
PHI:404	*PTH11*	reduced virulence
PHI:9290	*RAM1*	reduced virulence
PHI:2154	*RBP35*	reduced virulence
PHI:8752	*RHO2*	reduced virulence
PHI:8049	*Siz1*	reduced virulence
PHI:2117	*SPM1*	reduced virulence
PHI:7194*	*WISH*	reduced virulence
PHI:1017	*ABC4*	reduced virulence, loss of pathogenicity
PHI:2113*	*Kin4*	reduced virulence, loss of pathogenicity
PHI:2188	*MoMCM1*	reduced virulence, loss of pathogenicity
PHI:776	*Mstu1*	reduced virulence, loss of pathogenicity
PHI:3316	*GCC1*	reduced virulence, unaffected pathogenicity
PHI:2994	*FZC14*	reduced virulence, unaffected pathogenicity
PHI:2983	*FZC34*	reduced virulence, unaffected pathogenicity
PHI:9199*	*MoCKb1*	reduced virulence, unaffected pathogenicity
PHI:9032	*Mocrn1*	reduced virulence, unaffected pathogenicity
PHI:3701	*PacC*	reduced virulence, unaffected pathogenicity
PHI:2183	*CPXB*	unaffected pathogenicity
PHI:5601	*FZC11*	unaffected pathogenicity
PHI:5603	*FZC13*	unaffected pathogenicity
PHI:5607	*FZC16*	unaffected pathogenicity
PHI:5609*	*FZC18*	unaffected pathogenicity
PHI:5612	*FZC20*	unaffected pathogenicity

PHI:5618*	*FZC26*	unaffected pathogenicity
PHI:5621	*FZC27*	unaffected pathogenicity
PHI:5622	*FZC28*	unaffected pathogenicity
PHI:5625	*FZC31*	unaffected pathogenicity
PHI:5631	*FZC36*	unaffected pathogenicity
PHI:5632	*FZC37*	unaffected pathogenicity
PHI:5634	*FZC39*	unaffected pathogenicity
PHI:5635	*FZC40*	unaffected pathogenicity
PHI:5638	*FZC41*	unaffected pathogenicity
PHI:5640	*FZC43*	unaffected pathogenicity
PHI:5643	*FZC46*	unaffected pathogenicity
PHI:5651	*FZC54*	unaffected pathogenicity
PHI:5656*	*FZC58*	unaffected pathogenicity
PHI:5657*	*FZC59*	unaffected pathogenicity
PHI:5662*	*FZC62*	unaffected pathogenicity
PHI:5664	*FZC64*	unaffected pathogenicity
PHI:5672	*FZC69*	unaffected pathogenicity
PHI:5673*	*FZC70*	unaffected pathogenicity
PHI:5676	*FZC73*	unaffected pathogenicity
PHI:5681	*FZC77*	unaffected pathogenicity
PHI:5686*	*FZC82*	unaffected pathogenicity
PHI:5690	*FZC85*	unaffected pathogenicity
PHI:5692	*FZC87*	unaffected pathogenicity
PHI:4056	*Ldp1*	unaffected pathogenicity
PHI:2488	*Man1*	unaffected pathogenicity
PHI:2103	MGG_08710	unaffected pathogenicity
PHI:6613	*Mocapn14*	unaffected pathogenicity
PHI:2978*	*MoCel12A*	unaffected pathogenicity
PHI:6638	*Modnm2*	unaffected pathogenicity
PHI:7113	*MoERR2*	unaffected pathogenicity
PHI:5591	*MoNIT4*	unaffected pathogenicity
PHI:5666	*MoPRO1*	unaffected pathogenicity
PHI:3062	*MoRga2*	unaffected pathogenicity
PHI:3063	*MoRga3*	unaffected pathogenicity
PHI:3065*	*MoRga5*	unaffected pathogenicity
PHI:3066	*MoRga6*	unaffected pathogenicity
PHI:2197	*MoRgs6*	unaffected pathogenicity
PHI:2199	*MoRgs8*	unaffected pathogenicity
PHI:3039*	*MoSPA2*	unaffected pathogenicity
PHI:4110	*MoVOSA*	unaffected pathogenicity
PHI:7322*	*PEBP2*	unaffected pathogenicity

PHI:2185*	*PIC1*	unaffected pathogenicity
PHI:2027	*TGL1-1*	unaffected pathogenicity
PHI:2028*	*TGL1-2*	unaffected pathogenicity
PHI:2029	*TGL2*	unaffected pathogenicity
PHI:5610	*TRA1*	unaffected pathogenicity
PHI:7228	*tre1*	unaffected pathogenicity
PHI:2032	*VTL1*	unaffected pathogenicity
PHI:568	*XYL2*	unaffected pathogenicity
PHI:2138	*SGA1*	unaffected pathogenicity, loss of pathogenicity
PHI:2201	*MopdeL*	unaffected pathogenicity, loss of pathogenicity, reduced virulence
PHI:2105	Ca^{2+} permease	unaffected pathogenicity, reduced virulence
PHI:7724*	*cpk2*	unaffected pathogenicity, reduced virulence
PHI:9259*	*MoBIR1*	unaffected pathogenicity, reduced virulence
PHI:6612	*Mocapn9*	unaffected pathogenicity, reduced virulence

*Found in Swiching AS genes

Table 10. PHI-base assigned Switching AS genes

Gene Name	Alias	Phenotype
maker-scaffold000011-augustus-gene-18.11-mRNA-1	*COS1*	increased virulence (hypervirulence)
maker-scaffold000002-augustus-gene-28.131-mRNA-1	*RanBP*	loss of pathogenicity
maker-scaffold000011-augustus-gene-14.80-mRNA-1	*SEP1*	loss of pathogenicity
maker-scaffold000094-augustus-gene-12.5-mRNA-1	*Moatg1*	loss of pathogenicity
maker-scaffold000094-augustus-gene-24.88-mRNA-1	*TIG1*	loss of pathogenicity
maker-scaffold000095-augustus-gene-3.145-mRNA-1	*Moatg2*	loss of pathogenicity
maker-scaffold000095-augustus-gene-4.9-mRNA-1	*ASD4*	loss of pathogenicity
augustus-scaffold000032-processed-gene-11.58-mRNA-1	*conx1*	reduced virulence
augustus-scaffold000061-processed-gene-17.88-mRNA-1	MGG_03451	reduced virulence
maker-scaffold000060-augustus-gene-8.98-mRNA-1	*MoHox3*	reduced virulence
maker-scaffold000060-snap-gene-13.200-mRNA-1	*MoVELB*	reduced virulence
maker-scaffold000061-snap-gene-16.66-mRNA-1	MGG_03530	reduced virulence
maker-scaffold000062-augustus-gene-5.77-mRNA-1	*MoARK1*	reduced virulence
maker-scaffold000093-snap-gene-17.53-mRNA-1	*cnf2*	reduced virulence
maker-scaffold000094-augustus-gene-0.108-mRNA-1	*ABC1*	reduced virulence
maker-scaffold000094-augustus-gene-12.18-mRNA-1	*MoYAK1*	reduced virulence
snap-scaffold000077-processed-gene-12.249-mRNA-1	MGG_13324	reduced virulence
maker-scaffold000064-augustus-gene-1.175-mRNA-1	*WISH*	reduced virulence
maker-scaffold000095-augustus-gene-5.2-mRNA-1	*CHS7*	reduced virulence

maker-scaffold000095-snap-gene-11.72-mRNA-1	*EXP5*	reduced virulence
snap-scaffold000032-processed-gene-4.80-mRNA-1	*GSN1*	reduced virulence
augustus-scaffold000097-processed-gene-15.117-mRNA-1	*Kin4*	reduced virulence, loss of pathogenicity
augustus-scaffold000002-processed-gene-28.58-mRNA-1	*MoCKb1*	reduced virulence, unaffected pathogenicity
maker-scaffold000002-snap-gene-44.15-mRNA-1	*FZC62*	unaffected pathogenicity
maker-scaffold000031-augustus-gene-35.144-mRNA-1	*MoRga5*	unaffected pathogenicity
maker-scaffold000032-augustus-gene-12.12-mRNA-1	*FZC82*	unaffected pathogenicity
maker-scaffold000060-augustus-gene-10.65-mRNA-1	*PIC1*	unaffected pathogenicity
maker-scaffold000061-augustus-gene-6.80-mRNA-1	*FZC26*	unaffected pathogenicity
maker-scaffold000061-augustus-gene-9.9-mRNA-1	*MoSPA2*	unaffected pathogenicity
maker-scaffold000063-augustus-gene-11.91-mRNA-1	*FZC18*	unaffected pathogenicity
maker-scaffold000077-snap-gene-8.115-mRNA-1	*FZC58*	unaffected pathogenicity
maker-scaffold000096-augustus-gene-5.107-mRNA-1	*MoCel12A*	unaffected pathogenicity
maker-scaffold000098-snap-gene-0.131-mRNA-1	*FZC59*	unaffected pathogenicity
augustus-scaffold000110-processed-gene-3.202-mRNA-1	*PEBP2*	unaffected pathogenicity
maker-scaffold000002-augustus-gene-25.5-mRNA-1	*TGL1-2*	unaffected pathogenicity
maker-scaffold000062-augustus-gene-27.130-mRNA-1	*FZC73*	unaffected pathogenicity
maker-scaffold000062-augustus-gene-19.134-mRNA-1	*MoBIR1*	unaffected pathogenicity, reduced virulence
maker-scaffold000093-augustus-gene-11.4-mRNA-1	*cpk2*	unaffected pathogenicity, reduced virulence
maker-scaffold000002-augustus-gene-5.88-mRNA-1	*Mocapn9*	unaffected pathogenicity, reduced virulence

KJ201 genome annotation has been deposited in CFGP(http://cfgp.riceblast.snu.ac.kr/main.php)

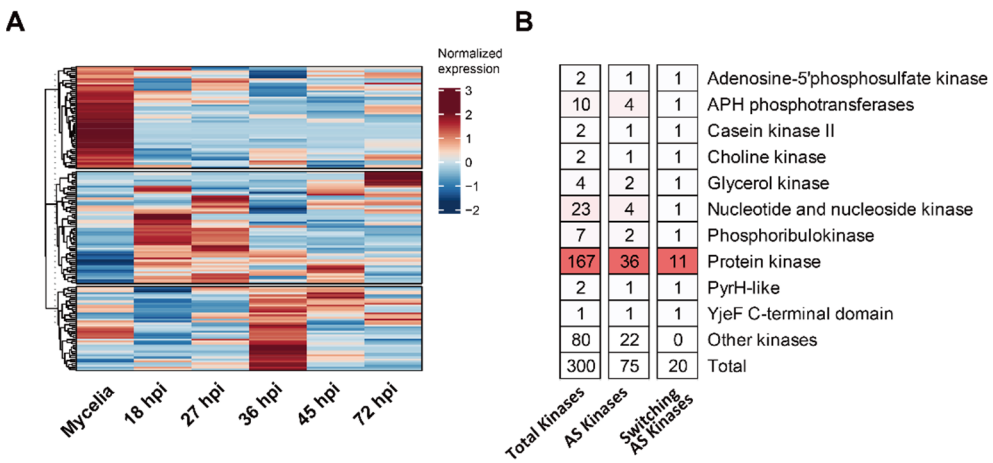


Figure 13. Expression patterns of alternatively spliced kinase genes

(A) The heatmap shows the expression patterns of AS isoforms produced by kinase-encoding genes. The rows are clustered by the Pearson correlation method, and Z-score normalization was applied to illustrate the expression levels across all stages. (B) The number of isoforms produced by each kinase-encoding gene is highlighted using relative color density. Switching AS isoforms include the isoforms with relative abundance > 0.5 and solely expressed during one or more infection stages.

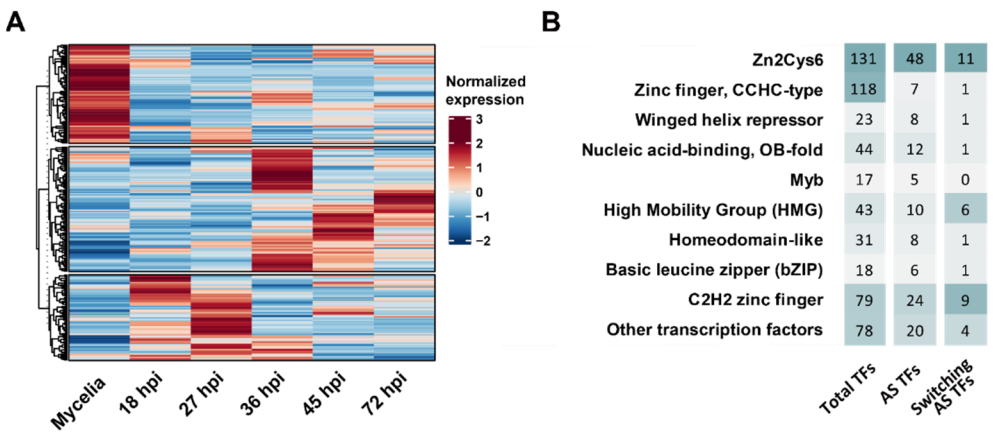


Figure 14. Expression patterns of alternatively spliced TF genes

(A) The heatmap shows the expression patterns of AS isoforms produced by TF-encoding genes. The rows are clustered by the Pearson correlation method, and Z-score normalization was applied to illustrate the expression levels across all stages. (B) The number of isoforms produced by each TF-encoding gene is highlighted using relative color density. Switching AS isoforms include the isoforms with relative abundance > 0.5 and solely expressed during one or more infection stages.

IV. Intron retention is the most common type of AS in *M. oryzae*

The number of isoforms produced from 2,413 AS genes ranged from one to ten, with 1,365 genes expressing one novel isoform and 1,048 genes expressing more than one novel isoform (Table 11). These genes showed multiple isoforms with differential expression patterns (Figure 15). Intron retention (IR) was the predominant type under all conditions (Figure 16). Alternative donor (A3SS), alternative acceptor (A5SS), and exon skipping (ES) were also observed. We compared intron splicing sites between annotated mRNAs and their isoforms to determine variation in their splicing sites. The pattern of AS in other fungal species showed the same order of type. Moreover, the order of type is preserved in the plant kingdom, whereas exon skipping is predominant in AS type of human (Figure 17). In annotated mRNAs, 6,362 introns have conserved di-nucleotides (GT|AG) at their splicing sites, and only two introns have alternative splicing sites (GC|AG) (Figure 18). In contrast, 269 introns had alternative splicing sites (GC|AG) among the isoforms.

Table 11. Distribution of novel isoform transcripts of *M. oryzae*

No. of novel isoforms	Number of genes
10	1
9	5
8	6
7	9
6	19
5	57
4	97
3	244
2	610
1	1,365

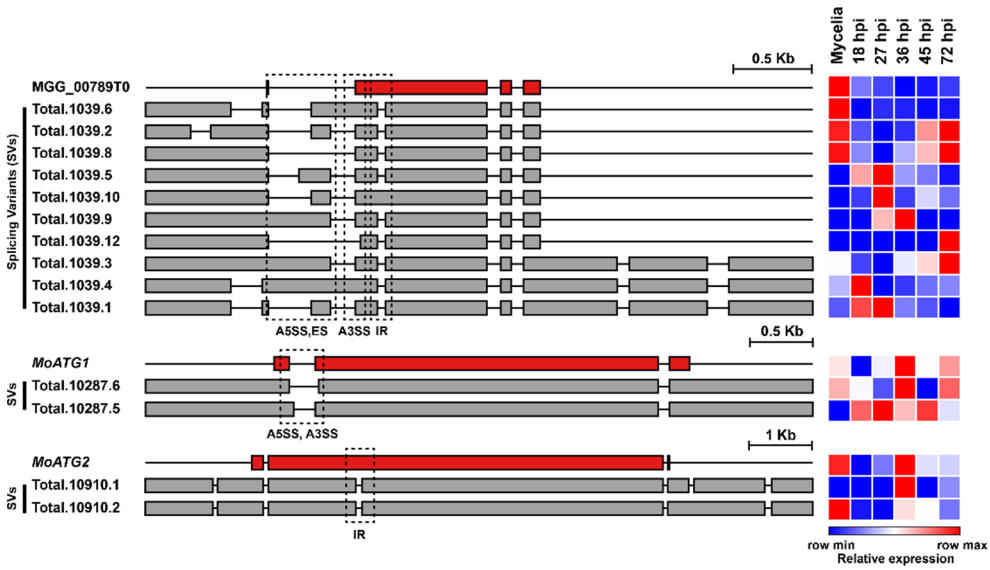


Figure 15. Example of AS type in the variable gene and pathogenicity genes

Occurrence patterns of four types of AS, including intron retention (IR), exon skipping (ES), alternative 3' splice site (A3SS), and alternative 5' splice site (A5SS) existing in one gene (MGG_00789T0). Illustration of two *ATG* genes depicts the AS type of functionally studied genes. The red box represents the annotated isoforms. The grey box represents identified novel transcripts in this study. The heatmap represents the expression patterns of AS isoforms, and Z-score normalization was applied to represent the expression levels across all stages.

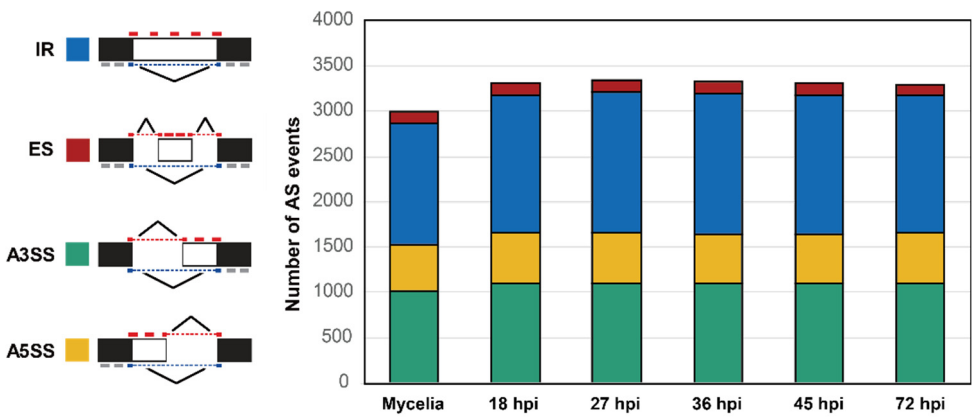


Figure 16. Different types of AS observed under different conditions

Occurrence patterns of four types of AS, including intron retention (IR), exon skipping (ES), alternative 3' splice site (A3SS), and alternative 5' splice site (A5SS), during the vegetative and infection stages. The black box denotes exon.

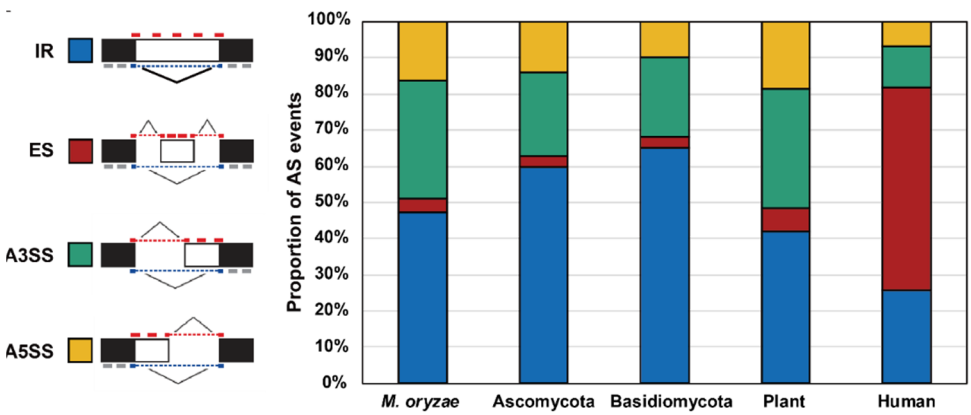


Figure 17. Distribution of different types of AS in the other organisms

Simplified diagram of four major isoform structures: Intron retention (IR), exon skipping (ES), alternative 3' splice site (A3SS), and alternative 5' splice site (A5SS). The relative abundance of AS patterns in four major internal events of AS in *M. oryzae* and the average proportion of other species from previous studies

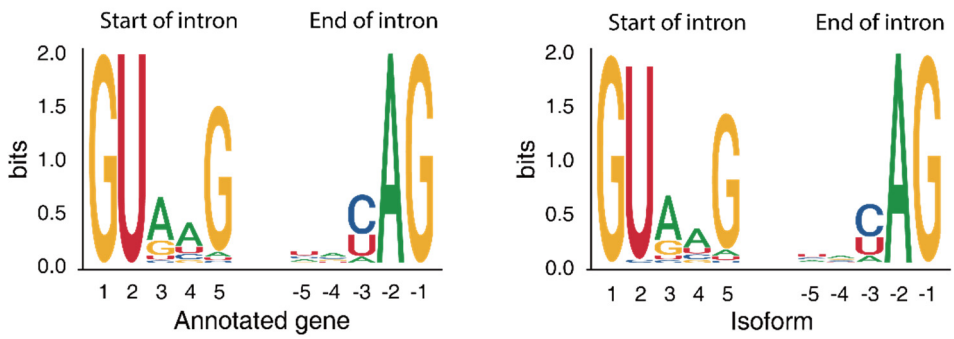


Figure 18. Pattern distribution of splicing site

Sequence patterns of splicing sites in annotated genes (left) and isoforms (right). The LOGO diagrams show 5 bp downstream and upstream from the splicing junctions.

V. AS-mediated changes in protein domain structure

We evaluated the protein-coding ability of individual AS isoforms using two methods, including coding potential and Premature Termination Codon (PTC) (Figure 19A). The coding potential was used to remove potentially non-coding transcripts. Among the 4,270 isoforms examined, 3,926 (91.9%) have the protein-coding potential. Some mRNAs produced via AS are unlikely to be translated because of NMD. Among the isoforms with coding potential, 2,344 isoforms had PTC. Whereas 71.6% of IR generated PTC, less than 60% of the other splicing patterns (exon skipping, alternative 3' or 5' splice site) created PTC (Figure 19B). A GO analysis based on the presence (PTC+) or absence (PTC-) of PTC showed distinct patterns. The PTC+ group contained five enriched terms, whereas 18 enriched terms were predicted in the PTC- group (Table 12). The remaining 1,582 isoforms did not have PTC, suggesting their translation (Table 13). These transcripts were contained more Kozak sequence, which was nucleic acid motif that functions as the protein translation initiation site in eukaryotic mRNA transcripts (Figure 20). Among the merged *ab initio* protein predictions, 1,220 (77%) isoforms were translated without PFAM domain change. However, 368 (22%) isoforms were predicted to encode proteins with altered PFAM domain(s) (Figure 19C) and could be categorized into multiple types. The most frequent one was the loss of domain(s) (263 isoforms). Domain copy variation (52 isoforms), gaining a domain (44 isoforms), and domain alteration (3 isoforms) were also observed (Figure 19C).

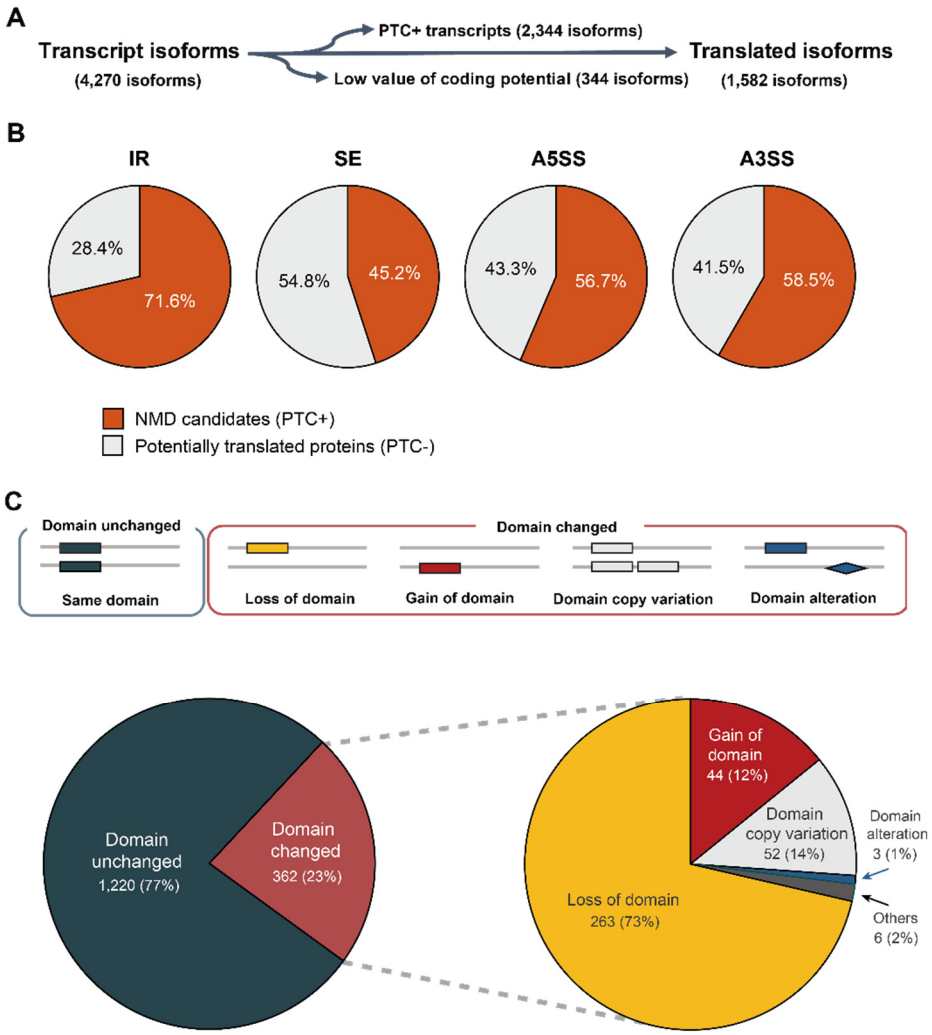


Figure 19. Domain transition patterns among the translated isoforms

(A) Filtration pipeline to remove non-translated isoforms and isoforms with low coding potential (B) Distribution patterns of NMD candidates (PTC+) in four major isoform types: Intron retention (IR), exon skipping (ES), alternative 3' splice site (A3SS), and alternative 5' splice site (A5SS). (C) The pie graphs show the distribution patterns of domain transition patterns among translated AS transcripts of *M. oryzae*

Table 12. Distribution of gene ontology (GO) in PTC+ and PTC- genes

GO ID	No. of genes	P-value	Function	Isoform type
GO:0140110	45	<0.001	transcription regulator activity*	PTC-
GO:0003700	42	0.001	DNA-binding transcription factor activity*	PTC-
GO:0009058	142	0.001	biosynthetic process	PTC-
GO:0009889	52	0.001	regulation of biosynthetic process	PTC-
GO:0019222	55	0.002	regulation of metabolic process	PTC-
GO:0080090	53	0.002	regulation of primary metabolic process	PTC-
GO:0051171	53	0.002	regulation of nitrogen compound metabolic process	PTC-
GO:0060255	55	0.002	regulation of macromolecule metabolic process	PTC-
GO:0031323	52	0.003	regulation of cellular metabolic process	PTC-
GO:0000981	32	0.004	DNA-binding transcription factor activity, RNA polymerase II-specific*	PTC-
GO:1901576	132	0.004	organic substance biosynthetic process	PTC-
GO:0044249	131	0.004	cellular biosynthetic process	PTC-
GO:0065007	74	0.015	biological regulation	PTC-
GO:0050789	69	0.015	regulation of biological process	PTC-
GO:0034641	148	0.018	cellular nitrogen compound metabolic process	PTC-
GO:0019637	29	0.019	organophosphate metabolic process	PTC-
GO:0050794	66	0.02	regulation of cellular process	PTC-
GO:0006139	116	0.044	nucleobase-containing compound metabolic process	PTC-
GO:0009058	226	<0.001	biosynthetic process	PTC+
GO:0009987	419	0.006	cellular process	PTC+
GO:0044237	353	0.022	cellular metabolic process	PTC+
GO:0051179	156	0.045	localization	PTC+
GO:0051234	153	0.046	establishment of localization	PTC+

Asterisk represents GO of molecular function

Table 13. Statistics of domain distribution of two different *ab initio* translation method

Ab initio translation	Statistics	No domain	Domain unchanged	upstream uORF	Domain changed			
					Domain altered	Domain appeared	Domain disappeared	Domain copy variation
Same start codon ORF	Isoform number	304	934	0	3	44	250	47
	Proportion (%)	19.22%	59.04%	0.00%	0.19%	2.78%	15.80%	2.97%
The longest ORF	Isoform number	252	803	313	3	33	131	47
	Proportion (%)	15.93%	50.76%	19.79%	0.19%	2.09%	8.28%	2.97%

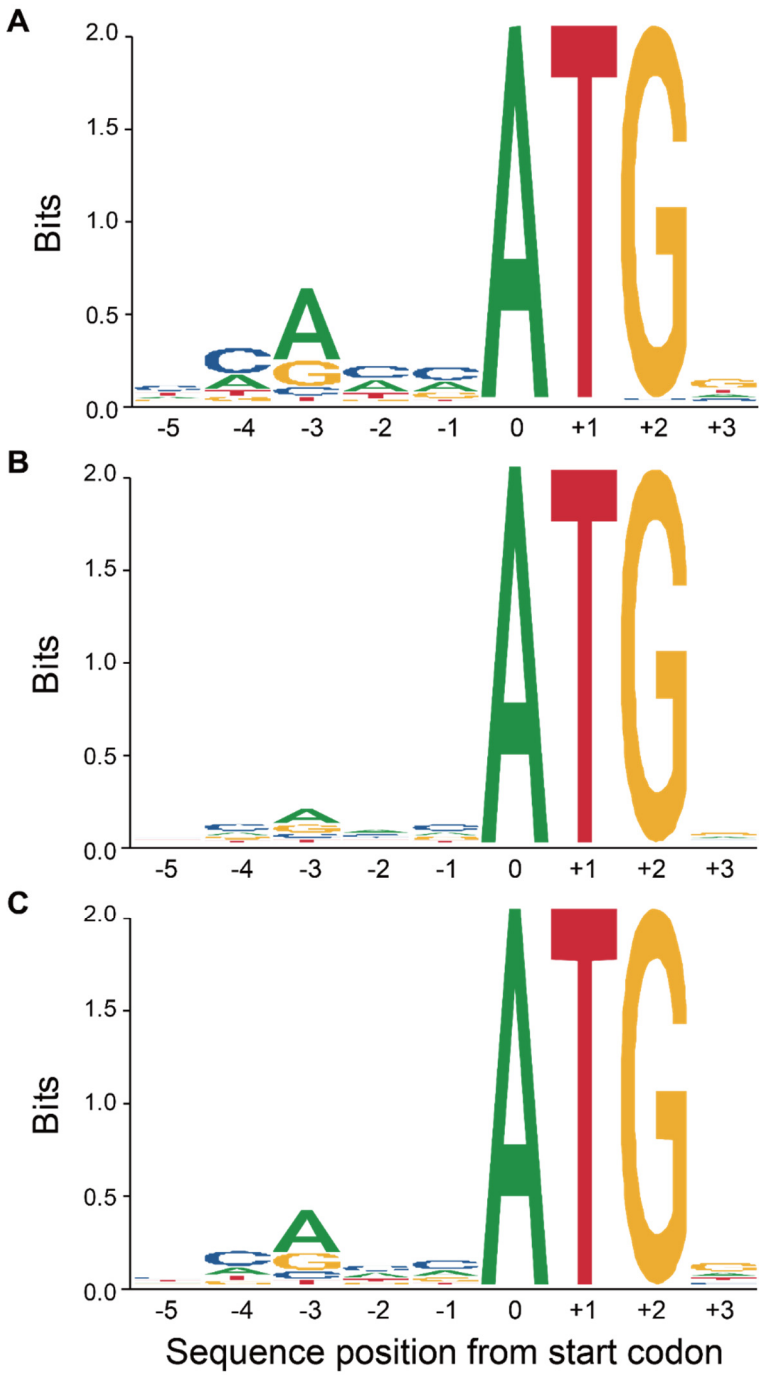

A
B
C
Bits
2.0
1.5
1.0
0.5
0.0
-5
-4
-3
-2
-1
0
+1
+2
+3
Sequence position from start codon

Figure 20. Sequence contexts of Kozak sequences

The sequence LOGO diagrams show the probability of base composition for sequence contexts of Kozak sequence (5 bp upstream and 1 bp downstream from the start codon). (A) sequence contexts of start codon region in the annotated gene model (B) sequence contexts of start codon in total isoforms (C) sequence contexts of start codon in subset of isoforms which removed PTC+ isoforms.

VI. Structural changes caused by AS among secreted proteins

Pathogenic fungi secrete some proteins as virulence factors. We assessed whether AS affects the production of any secreted proteins. The genome of KJ201 contains 1,340 genes predicted to encode secreted proteins. Among 168 genes that produced AS isoforms, 71 genes (42.3%) did not have PTC. The ability of secretion for 44 proteins seems to be affected by AS (Figure 21A), with 43 isoforms being predicted to produce proteins without the signal peptide and one gaining the signal peptide. Additionally, proteins produced from 23 isoforms were predicted to gain a transmembrane motif without any change in their signal peptide. We also identified that 28 genes encoding small secreted proteins, candidate effectors, produce AS isoforms. They were expressed at specific stages of infection (Figure 21B).

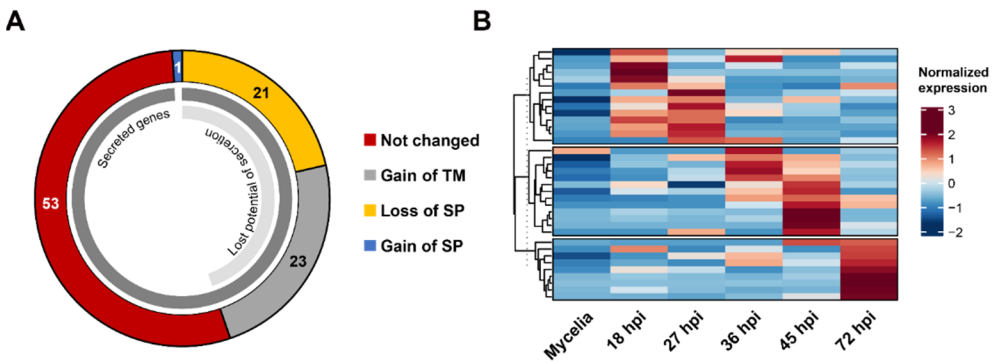


Figure 21. Predicted modifications of secreted proteins due to AS

(A) Three types of changes that potentially affect protein secretion. Two types, the loss of signal peptide (SP) and the gain of transmembrane motifs (TM), are expected to disrupt protein secretion. The gain of SP was also observed. (B) The expression pattern of AS isoforms of secreted protein-coding genes.

VII. Validation of the production and translation of some AS isoforms

We used RNA samples from mycelia and infected rice leaves (3 dpi and 6 dpi) to validate the production of predicted mRNA isoforms. Both annotated and isoforms were detected for six out of 10 genes tested (Figure 22). RT-PCR confirmed intron retention in the isoforms of five genes (orthologs of MGG_07681T0, MGG_11426T0, MGG_01368T0, MGG_011132T0, and MGG_08019T0) and exon skipping in one gene (MGG_17060T0). We also conducted a label-free quantitative mass spectrometry to validate the translation of some isoforms. We identified the proteins encoded by 266 genes in the proteome data. Most of the proteins were produced from annotated transcripts, but 24 proteins were translated from isoforms (Table 14).

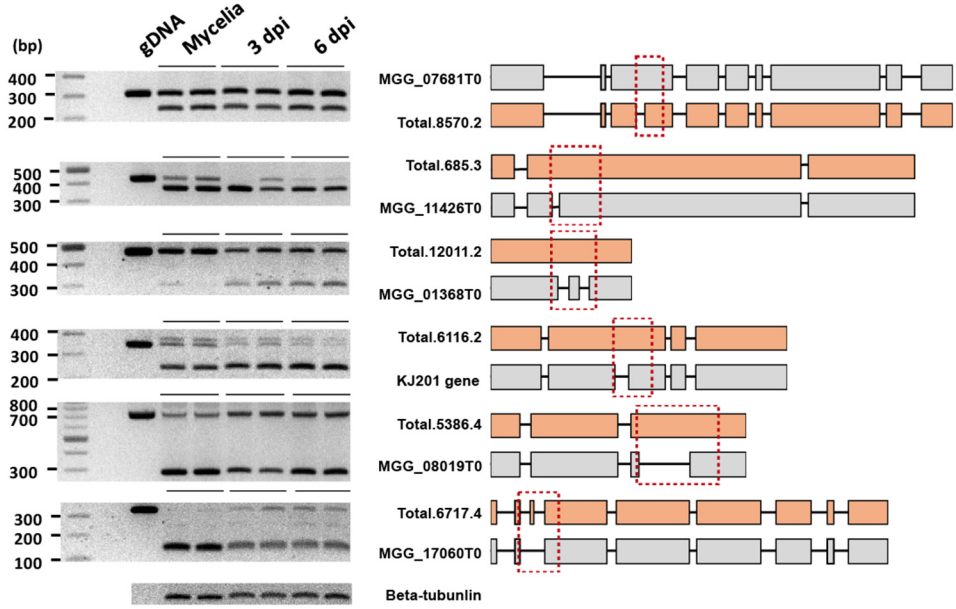


Figure 22. Validation of the transcripts produced via AS and resulting structural changes

AS of randomly selected genes was validated using RT-PCR. The samples analyzed are mycelia and rice leaves collected at 3 days post-inoculation (dpi) and 6 dpi. Genomic DNA was used as the control. The structural differences between the annotated form (grey) and isoform (orange) are shown. The red boxes indicate the amplified regions of DNA and cDNA.

Table 14. Detected protein list of label-free quantitative proteome analysis

# of detected proteins		Function
Annotation	Isoform	
-	1	Uncharacterized
-	1	HECT-like Ubiquitin-conjugating enzyme (E2)
-	1	Chitin recognition protein
1	1	Uncharacterized
-	1	Tyrosinase 65
1	2	Ras family
-	1	Ras family
-	1	U6 snRNA phosphodiesterase
-	1	Uncharacterized
-	1	Uncharacterized
-	1	Uncharacterized
-	1	Uncharacterized
-	1	Ribosome-associated complex
-	1	G-protein alpha subunit 14
-	1	Uncharacterized
-	1	Ubiquitin system component Cue
-	1	Peroxidase
1	1	Helicase
1	1	Histidyl-tRNA synthetase
1	1	Tetrahydrofolate dehydrogenase
-	1	RNA-dependent RNA polymerase
-	2	UAA transporter family
-	3	Oxysterol-binding protein
-	3	Calreticulin/calnexin

DISCUSSION

AS occurs widely in eukaryotes and expands the transcriptomic and proteomic diversity without increasing the number of genes. AS in plants is known to get involved in regulating disease resistance (Yang et al., 2014). Fungal transcriptome can also be diversified via AS during host invasion (Sieber et al., 2018). In this study, we performed comprehensive profiling of mRNAs generated via AS in the rice blast fungus to investigate if and how AS operates during plant infection. Instead of using 70-15, a widely employed laboratory strain, we used a field strain KJ201 to analyze the type and production pattern of AS isoforms. The RNA-seq data used in this study are highly enriched with fungal reads compared to the transcriptome data derived from plants infected with other *M. oryzae* isolates, enabling us to dissect AS profiles at multiple stages of infection in depth.

A previous study based on sequencing of Expressed Sequence Tags (ESTs) found AS isoforms in *M. oryzae* (Ebbole et al., 2004, Kim et al., 2010). However, because only a very limited number of transcripts were sequenced, this analysis could not fully reveal the global patterns of AS. We used high-depth RNA-seq data (Table 5) and a two-step AS profiling pipeline (Figure 4) to identify more AS isoforms accurately. Compared to recent studies on AS in other fungi (Burkhardt et al., 2015; Xia et al., 2017; Ibrahim et al., 2020), our analysis involved AS profiling at multiple time-points during infection. Because AS is known to be regulated by environmental conditions (Johnson and Vilardell, 2012), it is critical to analyze transcripts produced under diverse conditions to reveal the full spectrum of AS and resulting products. Our analysis identified 4,270 novel isoforms derived from 2,413 genes in *M. orzyae*

KJ201. Moreover, 87.3% of the annotated gene model transcript was confirmed by expression in AS repertoire. Among them, 696 isoforms transcribed from 499 genes were only produced during infection (Figure 5).

Plants undergo AS during pathogen and symbiotic microbes interactions (Rigo et al., 2019). Plant AS repertoire is known to be reprogrammed by fungal and oomycete infections (Huang et al., 2017), suggesting that both plants and pathogens respond to each other via AS. During infection, the extent of AS appears to increase, and the production of AS isoforms is regulated differentially depending on infection stage (Figure 7). The low depth of reads corresponding to *M. oryzae* transcripts during the biotrophic stage (27 and 36 hpi) likely caused the underestimation of the extent of AS (Liu et al., 2013). Detection of more genes subjected to AS during this stage than those undergoing during the vegetative stage suggested more active AS during infection (Figure 11B). During the biotrophic stage, molecular interactions occur between rice and *M. oryzae*. Since the outcome of such interactions determines whether infection can progress or not, the AS repertoire diversification in *M. oryzae* may modulate its adaptability by increasing the complexity of both transcriptome and proteome.

To assess the potential role of AS during infection and identify pathogenesis-associated genes regulated via AS, we queried PHI-base (Urban et al., 2020) using the identified genes subjected to AS. The genes for two autophagy-related proteins and four oxidative stress regulators in *M. oryzae* were identified to undergo AS. The anti-oxidation defense mechanism in *M. oryzae* is known to confer tolerance to host oxidative burst (Samalova et al., 2014). Autophagy is another adaptation strategy to low nutrient availability during the early stage of infection (Deng and Naqvi, 2019).

These findings suggest that the function of stress regulators could be modulated via AS. Moreover, the genes subjected to AS are enriched with TFs and phospho-transferases (Figure 11C, Table 7). The human pathogen *Candida albicans* utilizes TFs and kinases (one of the most well-characterized groups of phospho-transferases) to manage stresses and environmental adaptation (Brown et al., 2014). Considering the well-established role of TFs and phospho-transferases in regulating infection processes (Turra et al., 2014), the observed regulation of their transcription via AS also supports the role of AS in managing stresses in hosts.

AS is regulated by regulatory factors such as SR proteins and *trans*-acting hnRNP proteins (Chen and Manley, 2009), and the complexity of AS has been proposed to correlate with the number of these regulatory factors (Busch and Hertel, 2012). Relative to seven genes encoding SR proteins in *M. oryzae*, *Saccharomyces cerevisiae* has only one gene, and *Schizosaccharomyces pombe* carries two (Grutzmann et al., 2014). These regulatory factors are known to function in tissue and stress-specific manners (Duque, 2011). For example, two *Fusarium graminearum* SR proteins, FgSrp1 and FgSrp2, modulate AS at different stages of growth. FgSrp1 is involved in infection, whereas FgSrp2 is only involved in vegetative growth. (Zhang et al., 2017; Zhang et al., 2020) The genes encoding splicing regulatory factors of *M. oryzae* were also differently expressed at the vegetative and infection stages (Figure 7B), which is consistent with the distinct AS patterns observed (Figure 9).

The dominant AS mechanism in *M. oryzayzaesaaye* was intron retention (IR), and exon skipping did not occur frequently throughout all stages analyzed (Figure 15). This pattern has also been observed in other fungi and plants (Janbon et al., 2014;

Gonzalez-Hilarion et al., 2016). IR prevalently produces mRNAs containing PTC in plants (Chaudhary et al., 2019). IR generates nonsense isoforms harboring premature terminal codons (PTC+) subjected to NMD pathway or nuclear sequestration (Syed et al., 2012; Chaudhary et al., 2019). Three-fourth of the isoforms generated via IR contained PTC, and the proportion of PTC containing isoforms produced via other types of AS was less than that of IR (Figure 19B). The genes generating NMD-sensitive AS isoforms perform diverse functions in plants (Drechsel et al., 2013). The predicted functions of the genes producing isoforms harboring PTC were only enriched in localization, whereas the genes producing isoforms without PTC perform more diversified biological functions (Table 12).

AS isoforms are known to cause NMD or increase the proteome complexity. Production of alternative proteins created via AS was observed in *Lachancea kluyveri* and *P. cubensis* (Savory et al., 2012; Marshall et al., 2013). We found that 24 proteins were translated from the isoforms using a label-free quantitative mass spectrometry analysis, confirming the production of heterogeneous proteins from a single gene. Our *in silico* analysis of the domains of predicted proteins (Figure 19C) indicated that functional changes might occur due to AS, suggesting that protein complexity may increase during host infection.

During infection, secreted proteins play diverse critical roles, including the regulation and modification of host-associated environments and processes (Rep, 2005). The signal peptide and transmembrane motifs affect the secretion of individual proteins. Their modification via AS could alter their secretion, serving as one of the regulatory mechanisms for their function. This type of alteration of a small secreted protein by AS was previously reported in cucurbit downy mildew fungus *P.*

cubensis (Savory et al., 2012). In *M. oryzae*, 28 small secreted protein-encoding genes generated isoforms, including the gene encoding the well-known effector BAS3 (Mosquera et al., 2009). The alteration of protein secretion via AS suggests the involvement of AS in modulating effector secretion during infection.

In summary, our study revealed the temporal patterns of AS while *M. oryzae* infects rice. Functional analyses of the infection stage-specific AS isoforms and their products will advance our understating of how *M. oryzae* regulates the production of various factors to infect rice.

LITERATURE CITED

Almagro Armenteros, J.J., Tsirigos, K.D., Sonderby, C.K., Petersen, T.N., Winther, O., Brunak, S., Von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol*. 37, 420-423.

Angelini, C., De Canditiis, D., and De Feis, I. (2014). Computational approaches for isoform detection and estimation: good and bad news. *BMC Bioinf*. 15, 135.

Barta, A., Kalyna, M., and Reddy, A.S. (2010). Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants. *Plant Cell* 22, 2926-2929.

Brown, A.J., Budge, S., Kaloriti, D., Tillmann, A., Jacobsen, M.D., Yin, Z., Ene, I.V., Bohovych, I., Sandai, D., Kastora, S., Potrykus, J., Ballou, E.R., Childers, D.S., Shahana, S., and Leach, M.D. (2014). Stress adaptation in a pathogenic fungus. *J. Exp. Biol*. 217, 144-155.

Brown, J.W., Simpson, C.G., Marquez, Y., Gadd, G.M., Barta, A., and Kalyna, M. (2015). Lost in translation: pitfalls in deciphering plant alternative splicing transcripts. *Plant Cell* 27, 2083-2087.

Burkhardt, A., Buchanan, A., Cumbie, J.S., Savory, E.A., Chang, J.H., and Day, B. (2015). Alternative splicing in the obligate biotrophic oomycete pathogen *Pseudoperonospora cubensis*. *Mol. Plant Microbe Interact*. 28, 298-309.

Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* 3, 1-12.

Carpenter, S., Ricci, E.P., Mercier, B.C., Moore, M.J., and Fitzgerald, K.A. (2014).

Post-transcriptional regulation of gene expression in innate immunity. *Nat. Rev. Immunol*. 14, 361-376.

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A.S.N., Byrne, L.J., Wilson, C.M., and Syed, N.H. (2019). Alternative splicing and protein diversity: plants versus animals. *Front. Plant Sci*. 10, 708.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol*. 10, 741-754.

Choi, J., Park, J., Kim, D., Jung, K., Kang, S., and Lee, Y.H. (2010). Fungal secretome database: integrated platform for annotation of fungal secretomes. BMC Genomics 11, 105.

Dean, R., Van Kan, J.A., Pretorius, Z.A., Hammond-Kosack, K.E., Di Pietro, A., Spanu, P.D., Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., and Foster, G.D. (2012). The top 10 fungal pathogens in molecular plant pathology. *Mol. Plant. Pathol*. 13, 414-430.

Dean, R.A., Talbot, N.J., Ebbole, D.J., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J.R., Pan, H., Read, N.D., Lee, Y.H., Carbone, I., Brown, D., Oh, Y.Y., Donofrio, N., Jeong, J.S., Soanes, D.M., Djonovic, S., Kolomiets, E., Rehmeyer, C., Li, W., Harding, M., Kim, S., Lebrun, M.H., Bohnert, H., Coughlan, S., Butler, J., Calvo, S., Ma, L.J., Nicol, R., Purcell, S., Nusbaum, C., Galagan, J.E., and Birren, B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434, 980-986.

Deng, Y.Z., and Naqvi, N.I. (2019). Metabolic basis of pathogenesis and host adaptation in rice blast. *Annu. Rev. Microbiol*. 73, 601-619.

Dong, C., He, F., Berkowitz, O., Liu, J., Cao, P., Tang, M., Shi, H., Wang, W., Li, Q., Shen, Z., Whelan, J., and Zheng, L. (2018). Alternative splicing plays a critical role in maintaining mineral nutrient homeostasis in rice (*Oryza sativa*). *Plant Cell* 30, 2267-2285.

Drechsel, G., Kahles, A., Kesarwani, A.K., Stauffer, E., Behr, J., Drewe, P., Ratsch, G., and Wachtera, A. (2013). Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the *Arabidopsis* steady state transcriptome. *Plant Cell* 25, 3726-3742.

Duque, P. (2011). A role for SR proteins in plant stress responses. *Plant Signal Behav*. 6, 49-54.

Ebbole, D.J., Jin, Y., Thon, M., Pan, H.Q., Bhattarai, E., Thomas, T., and Dean, R. (2004). Gene discovery and gene expression in the rice blast fungus, *Magnaporthe grisea*: Analysis of expressed sequence tags. *Mol. Plant-Microbe Interact*. 17, 1337-1347.

Feng, S., Xu, M., Liu, F., Cui, C., and Zhou, B. (2019). Reconstruction of the full-length transcriptome atlas using PacBio Iso-Seq provides insight into the alternative splicing in *Gossypium australe*. *BMC Plant Biol*. 19, 365.

Filichkin, S.A., Priest, H.D., Givan, S.A., Shen, R., Bryant, D.W., Fox, S.E., Wong, W.K., and Mockler, T.C. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res*. 20, 45-58.

Foissac, S., and Sammeth, M. (2015). Analysis of alternative splicing events in custom gene datasets by AStalavista. *Methods Mol. Biol*. 1269, 379-392.

Franceschetti, M., Bueno, E., Wilson, R.A., Tucker, S.L., Gomez-Mena, C., Calder, G., and Sesma, A. (2011). Fungal virulence and development is regulated by

alternative pre-mRNA 3'end processing in *Magnaporthe oryzae*. *PLoS Pathog*. 7, e1002441.

Gao, X., Yin, C., Liu, X., Peng, J., Chen, D., He, D., Shi, W., Zhao, W., Yang, J., Peng, Y.-L. (2019). A glycine-rich protein MoGrp1 functions as a novel splicing factor to regulate fungal virulence and growth in *Magnaporthe oryzae*. *Phytopathol. Res.* 1, 1-15

Gehrmann, T., Pelkmans, J.F., Lugones, L.G., Wosten, H.A., Abeel, T., and Reinders, M.J. (2016). *Schizophyllum commune* has an extensive and functional alternative splicing repertoire. *Sci. Rep.* 6, 33640.

Gonzalez-Hilarion, S., Paulet, D., Lee, K.T., Hon, C.C., Lechat, P., Mogensen, E., Moyrand, F., Proux, C., Barboux, R., Bussotti, G., Hwang, J., Coppee, J.Y., Bahn, Y.S., and Janbon, G. (2016). Intron retention-dependent gene regulation in *Cryptococcus neoformans*. *Sci. Rep.* 6, 32252.

Grutzmann, K., Szafranski, K., Pohl, M., Voigt, K., Petzold, A., and Schuster, S. (2014). Fungal alternative splicing is associated with multicellular complexity and virulence: a genome-wide multi-species study. *DNA Res*. 21, 27-39.

Gupta, R., Min, C.W., Kramer, K., Agrawal, G.K., Rakwal, R., Park, K.H., Wang, Y., Finkemeier, I., and Kim, S.T. (2018). A multi-omics analysis of *Glycine max* leaves reveals alteration in flavonoid and isoflavonoid metabolism upon ethylene and abscisic acid treatment. *Proteomics* 18.

Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinf*. 12, 491.

Huang, J., Gu, L., Zhang, Y., Yan, T., Kong, G., Kong, L., Guo, B., Qiu, M., Wang, Y., Jing, M., Xing, W., Ye, W., Wu, Z., Zhang, Z., Zheng, X., Gijzen, M., Wang, Y., and Dong, S. (2017). An oomycete plant pathogen reprograms host pre-mRNA splicing to subvert immunity. *Nat. Commun*. 8, 2051.

Ibrahim, H.M.M., Kusch, S., Didelon, M., and Raffaele, S. (2020). Genome-wide alternative splicing profiling in the fungal plant pathogen *Sclerotinia sclerotiorum* during the colonization of diverse host families. *Mol. Plant Pathol*.

Janbon, G., Ormerod, K.L., Paulet, D., Byrnes, E.J., 3rd, Yadav, V., Chatterjee, G., Mullapudi, N., Hon, C.C., Billmyre, R.B., Brunel, F., Bahn, Y.S., Chen, W., Chen, Y., Chow, E.W., Coppee, J.Y., Floyd-Averette, A., Gaillardin, C., Gerik, K.J., Goldberg, J., Gonzalez-Hilarion, S., Gujja, S., Hamlin, J.L., Hsueh, Y.P., Ianiri, G., Jones, S., Kodira, C.D., Kozubowski, L., Lam, W., Marra, M., Mesner, L.D., Mieczkowski, P.A., Moyrand, F., Nielsen, K., Proux, C., Rossignol, T., Schein, J.E., Sun, S., Wollschlaeger, C., Wood, I.A., Zeng, Q., Neuveglise, C., Newlon, C.S., Perfect, J.R., Lodge, J.K., Idnurm, A., Stajich, J.E., Kronstad, J.W., Sanyal, K., Heitman, J., Fraser, J.A., Cuomo, C.A., and Dietrich, F.S. (2014). Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. grubii reveals complex RNA expression and microevolution leading to virulence attenuation. *PLoS Genet*. 10, e1004261.

Jeon, J., Lee, G.W., Kim, K.T., Park, S.Y., Kim, S., Kwon, S., Huh, A., Chung, H., Lee, D.Y., Kim, C.Y., and Lee, Y.H. (2020). Transcriptome profiling of the rice blast fungus *Magnaporthe oryzae* and its host *Oryza sativa* during

infection. *Mol. Plant Microbe Interact*. 33, 141-144.

Jeong, S. (2017). SR Proteins: binders, regulators, and connectors of RNA. *Mol. Cells* 40, 1-9.

Jin, L., Li, G., Yu, D., Huang, W., Cheng, C., Liao, S., Wu, Q., and Zhang, Y. (2017). Transcriptome analysis reveals the complexity of alternative splicing regulation in the fungus *Verticillium dahliae*. *BMC Genomics* 18, 130.

Johnson, T.L., and Vilardell, J. (2012). Regulated pre-mRNA splicing: the ghostwriter of the eukaryotic genome. *Biochim. Biophys. Acta* 1819, 538-545.

Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L., and Gao, G. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res*. 45, W12-W16.

Kannan, N., Taylor, S.S., Zhai, Y.F., Venter, J.C., and Manning, G. (2007). Structural and functional diversity of the microbial kinome. *PLoS Biol*. 5, 467-478.

Kim, K.T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., Kang, S., and Lee, Y.H. (2016). Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Front. Plant Sci*. 7: 186

Kim, K.T., Ko, J., Song, H., Choi, G., Kim, H., Jeon, J., Cheong, K., Kang, S., and Lee, Y.H. (2019). Evolution of the genes encoding effector candidates within multiple pathotypes of *Magnaporthe oryzae*. *Front. Microbiol*. 10: 3389

Kim, S., Park, J., Park, S.Y., Mitchell, T.K., and Lee, Y.H. (2010). Identification and analysis of *in planta* expressed genes of *Magnaporthe oryzae*. BMC Genomics 11: 104

Kuang, Z., Boeke, J.D., and Canzar, S. (2017). The dynamic landscape of fission

yeast meiosis alternative-splice isoforms. *Genome Res*. 27, 145-156.

Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A., and Murphy, J.W. (2004). Introns and splicing elements of five diverse fungi. *Eukaryot. Cell* 3, 1088-1100.

Li, Z.Q., Wu, L.Y., Wu, H., Zhang, X.X., Mei, J., Zhou, X.P., Wang, G.L., and Liu, W.D. (2020). Arginine methylation is required for remodelling pre-mRNA splicing and induction of autophagy in rice blast fungus. *New Phytol*. 225, 413-429.

Liu, Y., Ferguson, J.F., Xue, C., Silverman, I.M., Gregory, B., Reilly, M.P., and Li, M. (2013). Evaluating the impact of sequencing depth on transcriptome profiling in human adipose. *PLoS One* 8, e66883.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D.W., Yiu, S.M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T.W., and Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 1, 18.

Marquez, Y., Brown, J.W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Res*. 22, 1184-1195.

Marshall, A.N., Montealegre, M.C., Jimenez-Lopez, C., Lorenz, M.C., and Van Hoof, A. (2013). Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. *PLoS Genet*. 9, e1003376.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* 17, 10-12.

Miranda-Saavedra, D., and Barton, G.J. (2007). Classification and functional annotation of eukaryotic protein kinases. *Proteins*. 68, 893-914.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.Y., El-Gebali, S., Fraser, M.I., Gough, J., Haft, D.R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A.P., Paysan-Lafosse, T., Pesseat, S., Potter, S.C., Qureshi, M.A., Rawlings, N.D., Redaschi, N., Richardson, L.J., Rivoire, C., Salazar, G.A., Sangrador-Vegas, A., Sigrist, C.J.A., Sillitoe, I., Sutton, G.G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Yong, S.Y., and Finn, R.D. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. 47, D351-D360.

Moller, S., Croning, M.D., and Apweiler, R. (2001). Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646-653.

Mosquera, G., Giraldo, M.C., Khang, C.H., Coughlan, S., and Valent, B. (2009). Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as biotrophy- associated secreted proteins in rice blast disease. *Plant Cell* 21, 1273-1290.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet*. 40, 1413-1415.

Park, J., Park, J., Jang, S., Kim, S., Kong, S., Choi, J., Ahn, K., Kim, J., Lee, S., Kim,

S., Park, B., Jung, K., Kim, S., Kang, S., and Lee, Y.H. (2008). FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. *Bioinformatics* 24, 1024-1025.

Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Res* 9.

Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc*. 11, 1650-1667.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol*. 33, 290-295.

Rep, M. (2005). Small proteins of plant-pathogenic fungi secreted during host colonization. *FEMS Microbiol. Lett*. 253, 19-27.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*. 16, 276-277.

Rigo, R., Bazin, J.R.M., Crespi, M., and Charon, C.L. (2019). Alternative splicing in the regulation of plant-microbe interactions. *Plant Cell Physiol*. 60, 1906-1916.

Samalova, M., Meyer, A.J., Gurr, S.J., and Fricker, M.D. (2014). Robust anti-oxidant defences in the rice blast fungus *Magnaporthe oryzae* confer tolerance to the host oxidative burst. *New Phytol*. 201, 556-573.

Savory, E.A., Zou, C., Adhikari, B.N., Hamilton, J.P., Buell, C.R., Shiu, S.H., and Day, B. (2012). Alternative splicing of a multi-drug transporter from *Pseudoperonospora cubensis* generates an RXLR effector protein that elicits

a rapid cell death. *PLoS One* 7.

Scotti, M.M., and Swanson, M.S. (2016). RNA mis-splicing in disease. *Nat. Rev. Genet*. 17, 19-32.

Shi, Y.G. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol*. 18, 655-670.

Sibthorp, C., Wu, H., Cowley, G., Wong, P.W., Palaima, P., Morozov, I.Y., Weedall, G.D., and Caddick, M.X. (2013). Transcriptome analysis of the filamentous fungus *Aspergillus nidulans* directed to the global identification of promoters. *BMC Genomics* 14, 847.

Sieber, P., Voigt, K., Kammer, P., Brunke, S., Schuster, S., and Linde, J. (2018). Comparative study on alternative splicing in human fungal pathogens suggests its involvement during host invasion. *Front. Microbiol*. 9, 2313.

Staiger, D., and Brown, J.W.S. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell* 25, 3640-3656.

Sultan, M., Schulz, M.H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., Schmidt, D., O'keeffe, S., Haas, S., Vingron, M., Lehrach, H., and Yaspo, M.L. (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321, 956-960.

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J.W. (2012). Alternative splicing in plants--coming of age. *Trends Plant Sci*. 17, 616-623.

Talbot, N.J. (2003). On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. *Annu. Rev. Microbiol*. 57, 177-202.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol*. 28, 511-515.

Tress, M.L., Abascal, F., and Valencia, A. (2017). Alternative splicing may not be the key to proteome complexity. *Trends Biochem. Sci*. 42, 98-110.

Trincado, J.L., Entizne, J.C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D.J., and Eyras, E. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol*. 19, 40.

Turra, D., Segorbe, D., and Di Pietro, A. (2014). Protein kinases in plant-pathogenic fungi: conserved regulators of infection. *Annu. Rev. Phytopathol*., Vol 52 52, 267-288.

Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc*. 11, 2301-2319.

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S.Y., De Silva, N., Martinez, M.C., Pedro, H., Yates, A.D., Hassani-Pak, K., and Hammond-Kosack, K.E. (2020). PHI-base: the pathogen-host interactions database. *Nucleic Acids Res*. 48, D613-D620.

Wang, B., Guo, G., Wang, C., Lin, Y., Wang, X., Zhao, M., Guo, Y., He, M., Zhang, Y., and Pan, L. (2010). Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. *Nucleic Acids Res*. 38, 5075-5087.

Wang, B., Tseng, E., Regulski, M., Clark, T.A., Hon, T., Jiao, Y., Lu, Z., Olson, A., Stein, J.C., and Ware, D. (2016). Unveiling the complexity of the maize

transcriptome by single-molecule long-read sequencing. *Nat. Commun*. 7, 11708.

Wang, K., Wang, D.H., Zheng, X.M., Qin, A., Zhou, J., Guo, B.Y., Chen, Y.J., Wen, X.P., Ye, W., Zhou, Y., and Zhu, Y.X. (2019). Multi-strategic RNA-seq analysis reveals a high-resolution transcriptional landscape in cotton. *Nat. Commun*. 10.

Wilson, D., Pethica, R., Zhou, Y.D., Talbot, C., Vogel, C., Madera, M., Chothia, C., and Gough, J. (2009). SUPERFAMILY-sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Res*. 37, D380-D386.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359-362.

Xia, Y., Fei, B., He, J., Zhou, M., Zhang, D., Pan, L., Li, S., Liang, Y., Wang, L., Zhu, J., Li, P., and Zheng, A. (2017). Transcriptome analysis reveals the host selection fitness mechanisms of the *Rhizoctonia solani* AG1IA pathogen. *Sci. Rep*. 7, 10120.

Yang, S., Tang, F., and Zhu, H. (2014). Alternative splicing in plant immunity. *Int. J. Mol. Sci.* 15, 10424-10445.

Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., Xu, H., Huang, X., Li, S., Zhou, A., Zhang, X., Bolund, L., Chen, Q., Wang, J., Yang, H., Fang, L., and Shi, C. (2018). WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Res*. 46, W71-W75.

Yu, H., Tian, C., Yu, Y., and Jiao, Y. (2016). Transcriptome survey of the contribution of alternative splicing to proteome diversity in *Arabidopsis thaliana*. *Mol.*

Plant 9, 749-752.

Zhang, Y., Dai, Y., Huang, Y., Wang, K., Lu, P., Xu, H., Xu, J.R., and Liu, H. (2020). The SR-protein FgSrp2 regulates vegetative growth, sexual reproduction and pre-mRNA processing by interacting with FgSrp1 in *Fusarium graminearum*. *Curr. Genet*. 66, 607-619.

Zhang, Y.M., Gao, X.L., Sun, M.L., Liu, H.Q., and Xu, J.R. (2017). The *FgSRP1* SR-protein gene is important for plant infection and pre-mRNA processing in *Fusarium graminearum*. *Environ. Microbiol*. 19, 4065-4079.

기주식물 침입시 벼도열병균의 선택적 스플라이싱 다양성 분석

전 종 범

초 록

최근 RNA 시퀀싱(RNA-sequencing) 분석 기술의 발달과 전사체 데이터의 축적으로 인해 심층적인 전사체 분석이 가능하게 되었다. 이런 전사체 분석은 각각의 유전자들이 전사서열의 조절을 통해 조절되고 있음을 밝혀주었다. 그 조절기작 중 대표되는 것은 선택적 스플라이싱(alternative splicing)으로 이러한 방법들은 하나의 유전자에서 이형의 전사서열(isoform)을 만듦으로써 유전자의 발현과 단백질 기능을 조절하는 것으로 알려져 있다. 인간과 식물에서 유전체의 선택적 스플라이싱 메커니즘은 세포 분화, 환경 적응 및 각종 스트레스에 대응하기 의해 조절되는 것으로 밝혀졌었는데, 병원균 감염 역시 선택적 스플라이싱을 사용하여 식물에서 전사체 네트워크를 재구성하는 요인 중 하나로 알려져 있다. 그러나 선택적 스플라이싱 다양화는 주로 식물 전사체에 초점을 맞추고 있으며, 미생물-식물 상호작용에서 식물병원균 전사체의 선택적 스플라이싱 연구는 거의 이루어지지 않았다. 그러므로 위 조절기작의 곰팡이 유전체에서 연구는 곰팡이 병원체가 식물 방어 시스템을 어떻게 상쇄하거나 극복하는지에 대한 단서를 얻는 데 필요하다.

해당 연구는 곰팡이 시스템의 반응을 이해하기 위해 벼 감염시 벼도열병균(*Magnaporthe oryzae*)의 선택적 스플라이싱 레퍼토리를

프로파일링(profiling)했다. 특히 곰팡이 균사생장 또는 다양한 기주감염 단계에서의 레퍼토리의 동정은 병 진전 과정을 시간대별로 이해하기 위한 다각적인 전사체를 구축했다. 그 결과 선택적 스플라이싱에 의해 생기는 이형 전사서열은 균사생장시기에 비해 기주감염시기 동안 증가하는 것으로 밝혀졌다. 더욱이, 특히 기주감염 단계에서 특이적인 이형 전사서열은 곰팡이가 기주 세포의 침투한 직후 단계부터 활물기생성 시기동안 증가하였는데, 이후 사물기생성 단계에서는 특이적인 이형 전사서열의 숫자가 감소한다는 것을 밝혔다. 이러한 조성의 차이를 기반하여 병 진전 과정 동안 특이적 스플라이싱 유도는 기주 세포와의 상호작용과 상관관계가 있음을 밝혔다.

이 기작을 잘 이해하기 위해서는 선택적 스플라이싱이 단백질 기능을 변화하는 방법에 대한 이해가 필요하다. 선택적 스플라이싱의 역할에는 두 가지 주요 개념이 있다. 하나는 논센스 서열에 의한 전사서열의 분해(Nonsense-Mediated Decay)이고 다른 역할은 단백체의 다양성을 증가시키는 것이다. 이 연구에서는 선택적 스플라이싱의 두 가지 주요 역할을 고려해 잠재적 NMD 후보와 단백질 번역 후보를 나눴다. 특히, 다수의 인트론 잔류(intron retention)는 NMD를 유발하는 요인임을 밝혔다. 또한 NMD가 일어나지 않는 서열인 단백질 번역후보에서 우리는 기능 도메인(protein domain) 구조 변화 및 분비와 관련된 서열의 변화를 가진 단백질을 확인함으로써 기능 도메인에서의 단백질 변이와 분비단백질의 분비구조의 변이를 찾았다. 이는 감염 동안 단백질의 잠재적인 기능의 변화를 유도할 것으로 판단된다.

이 연구는 전체 기주감염 과정에 걸쳐 선택적 스플라이싱 레퍼토리를 제공하며 벼도열병균에서 병원성 및 환경 신호의 적응에 대한 분자생물학적 연구의 기반을 제공할 것이다. 더욱이 위 레퍼토리의 단백질 예측은 곰팡이 유전체에서 다면발현(pleiotrophy)의 개념을 제안한다. 종합적으로, 이러한 선택적 스플라이싱 전사체는 기주-병원균 상호 작용 동안 곰팡이 선택적 스플라이싱 전사체의 복잡성에 대한

식견을 제공할 것이다.

주요어 : 선택적 스플라이싱, 곰팡이-병원성 상호작용, 벼도열병균, 전사체 다양화, 병원성

학번 : 2013-31043