ABSTRACT

JEON, JUNHYUN. Characterization of Experimental Evolution and DNA Methylation in the Rice Blast Fungus. (Under the co-direction of Dr. Yong-Hwan Lee and Dr. Ralph A. Dean).

Much progress has been made in understanding the genetic basis of fungal development and pathogenicity during the last few decades. However, there are still wide gap in knowledge on fungal evolution and biology. Recently, rapid development in high throughput sequencing technology made it possible to put forth the hypothesis that couldn't be tested before. Applying molecular biology tools and high throughput sequencing, here I address the experimental evolution and contribution of an epigenetic component-DNA methylation to development of a model plant pathogenic fungus, *Magnaporthe oryzae*.

In chapter 1, I summarize important aspects of *M. oryzae* as a model plant pathogen to study the evolution and DNA methylation in filamentous fungi. I also give a brief brief description of experimental evolution and DNA methylation, which will serve as a general introduction to the next chapters.

In chapter 2, I present the application of method to study of experimental evolution in M. oryzae. Evolved strains (derived strain in chapter 2) under laboratory conditions were obtained and characterized at both phenotypic and molecular level in comparison to the parental strain. Monitoring the phenotypic and genotypic changes revealed that phenotypic changes are not proportional to genotypic changes. Analysis of mutations within the genome of the derived strain showed that genetic drift and selection shape the mutational landscape of the derived genomes, and that main source of mutations under laboratory condition are oxidative damages and/or UV.

In chapter 3, I investigate the DNA methylation during fungal development. DNA methylation in fungi is considered to be a stable mark for transcriptional silencing of transposable element. Analysis of DNA methylation in *M. oryzae* showed that DNA methylation pattern in the fungal genome undergoes massive reprogramming during pathogenic development, suggesting new paradigm to evolution and function of fungal DNA methylation. I argue that such reprogramming of fungal DNA methylation is analogous to epigenetic reprogramming during mammalian development. I conclude the chapter discussing the similarities and differences between two systems.

Characterization of Experimental Evolution and DNA Methylation in the Rice Blast Fungus

by
Junhyun Jeon

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Functional Genomics

Raleigh, North Carolina

2012

APPROVED BY:

_____	_____
Dr. Ralph A. Dean Chair of Advisory Committee	Dr. Yong-Hwan Lee
_____	_____
Dr. Gary A. Payne	Dr. David M. Bird

DEDICATION

I dedicate this accomplishment to my wife Jua and our son Ryan, who have always been there for me, and have never doubted my dream to be a scientist, no matter what it takes. I also dedicate this to my parents who always supported me in every endeavor. I love you all.

이 학위논문을 저를 위해 항상 그 자리에 있어주었고 과학자로서의 꿈을 단 한 순간도 의심치 않았던 사랑하는 아내 주아와 아들 동연이에게 바칩니다.
또한 모든 일에 있어서 저를 아낌없이 후원하고 믿어주셨던 부모님께도 이 학위 논문을 바칩니다.

BIOGRAPHY

Educational history

Junhyun Jeon was born on July 22nd, 1978 in Seoul, Korea. During his early childhood, Junhyun was raised in Yong-In, which was a small city back then. The neighborhood, where plants and insects were diverse and abundant, provided the opportunities to develop curiosity of how the nature works. In 1989, Junhyun moved to Suwon, the capital city of GyenongGi province and attended the middle and high school. In 1997, Junhyun started his undergraduate major in agricultural biology at the Seoul National University (SNU). At the very beginning of his educational journey in university, he was fascinated by molecular aspects of plant pathology – how pathogens can cause diseases on their host plants. This attracted him to Master of Science (M. Sc.) in Plant Pathology under the direction of Dr. Yong-Hwan Lee in 2004. Junhyun received his degree in February 2006 and spent one year as research assistant in the lab. In 2007, he started his doctoral program in the Department of Agricultural Biotechnology at SNU under the direction of Dr. Yong-Hwan Lee. One and half years later, Junhyun took unique opportunity to pursue two degrees at the same time – one at the SNU and the other in Functional Genomics program at North Carolina State University (NCSU). He spent two years in NCSU taking classes and conducting a research project under the direction of Dr. Ralph A. Dean in Center for Integrated Fungal Research. The classes he has taken and years of experiences made him develop new perspectives on study of microbial pathogens. Below is the list of awards and publications that came out of his academic journey.

Awards

1. Recognition of Excellence, KSPP (Korean Society of Plant Pathology)-2006 Poster Competition

2. 2008 BK21 Young Brain Award

Publications (including manuscript in prep)

1. **Junhyun Jeon**, Sook-Young Park, Myoung-Hwan Chi, Jaehyuk Choi, Jongsun Park, Hee-Sool Rho, Soonok Kim, Jaeduk Goh, Sungyong Yoo, Jinhee Choi, Ju-Young Park, Mihwa Yi, Seonyoung Yang, Min-Jung Kwon, Seong-Sook Han, Byeong Ryun Kim, Chang Hyun Khang, Bongsoo Park, Se-Eun Lim, Kyongyong Jung, Sunghyung Kong, Maruthachalam Karunakaran, Hong-Sik Oh, Hyojeong Kim, Seryun Kim, Jaejin Park,Soyoung Kang, Woo-Bong Choi, Seogchan Kang and Yong-Hwan Lee. 2007. Genome-wide Functional Analysis of Pathogenicity Genes in the Rice Blast Fungus. **Nature Genetics** 39. 561-565.

2. Jaehyuk Choi, Jongsun Park, **Junhyun Jeon**, Myoung-Hwan Chi, Jaeduk Goh, Sung-Yong Yoo, Jaejin Park, Kyongyong Jung, Hyojeong Kim, Sook-Young Park, Hee-Sool Rho, Soonok Kim, Byeong Ryun Kim, Seong-Sook Han, Seogchan Kang and Yong-Hwan Lee. 2007. Genome-wide analysis of T-DNA integration into the chromosomes of Magnaporthe oryzae. **Molecular Microbiology** 66:371-382

3. **Junhyun Jeon**, Jaeduk Goh, Sungyong Yoo, Myoung-Hwan Chi, Jaehyuk Choi, Hee-Sool Rho, Jongsun Park, Seong-Sook Han, Byeong Ryun Kim, Sook-Young Park, Soonok Kim and Yong-Hwan Lee. 2008. A putative MAP kinaes kinase kinase, MCK1 is required for cell wall integrity and pathogenicity of the rice blast fungus, Magnaporthe oryzae. **Molecular Plant-Microbe Interactions** 21:525-534.

4. **Junhyun Jeon**, Jaehyun Choi, Jongsun Park, and Yong-Hwan Lee. 2008. Functional genomics in the rice blast fungus to unravel the fungal pathogenicity (Review). **Journal**

of Zhejiang University Science B 9(10): 747-752.

5. Hee-Sool Rho, **Junhyun Jeon**, and Yong-Hwan Lee. 2009. Phospholipase C-mediated signaling is required for fungal development and pathogenicity in *Magnaporthe oryzae*. **Molecular Plant Pathology** 10(3):337 - 346

6. Jaeduk Goh**,** Kyoung Su Kim**,** Jaejin Park, **Junhyun Jeon**, Sook-Young Park, Yong-Hwan Lee. 2011**.** The cell cycle gene *MoCDC15* regulates hyphal growth, asexual development and plant infection in the rice blast pathogen *Magnaporthe oryzae.* **Fungal Genetics and Biology**. 48(8):784-792

7. Jaeduk Goh, **Junhyun Jeon**, Kyoung Su Kim, Jongsun Park, Sook-Young Park and Yong-Hwan Lee. 2011. The Pex7-mediated peroxisomal import system is required for fungal development and pathogenicity in *Magnaporthe oryzae*. **PLoS ONE** 6(12): e28220. doi:10.1371/journal.pone.0028220.

8. Karunakaran Maruthachalam, **Junhyun Jeon**, Yong-Hwan Lee and Krishna V. Subbarao. 2011. Identification of Fungal Pathogenicity Genes by *Agrobacterium tumefaciens*-Mediated Transformation (Book chapter). **Biotechnology of Fungal Genes.** pp1-19

9. Sook-Young Park, Jaeyoung Choi, Se-Eun Lim, Gir-Won Lee, Jongsun Park, Yangsun Kim, Sunghyung Kong, Se Ryun Kim, Hee-Sool Rho, **Junhyun Jeon**, Myung-Hwan Chi, Soonok Kim, Chang Hyun Khang, Seogchan Kang, and Yong-Hwan Lee. 2012. Global expression of transcription factor genes provides new insights into pathogenicity and stress responses in the rice blast fungus (submitted manuscript).

10. **Junhyun Jeon**, Jaeyoung Choi, Gir-Won Lee, Sook-Young Park, Aram Huh, Ralph A. Dean, and Yong-Hwan Lee. Global reprogramming of DNA methylation during pathogenic development in the rice blast fungus (manuscript in preparation).

ACKNOWLEDGMENTS

Above all, I would like to express my sincere gratitude to my academic advisors, Dr. Yong-Hwan Lee and Dr. Ralph A. Dean, for giving me unique opportunities to take part in wonderful scientific voyage of both labs: Fungal Plant Pathology Laboratory in SNU and Fungal Genomics Laboratory in NCSU. Inspiration, encouragement, and unconditional support I had from both mentors made me stand strong against the hardships and challenges. From a scientific point of view, critical and creative thinking is one of the most valuable lessons I have learned from them. I think I have been extremely fortunate to have two great mentors at the same time. Thank you again for many other reasons I can't write down here all.

I would like to also express appreciation to my committee members, Drs. Gary A. Payne and David McK. Bird for their help and support in every way at every stage of PhD program. I have to say to Dr. Bird that 'Journal Club' had been my favorite class. I learned from it more than any other classes.

I thank BK21 program for financial support and also thank all the people who made dual degree program possible for their endeavor. I really thank you all for bringing about life-changing opportunity. I also thank faculties in plant pathology department of both NCSU and SNU for helpful advice and suggestion.

I am very grateful to former and current lab members of both labs: Dr. Yeonyee Oh, Dr. Cristiano Caixeta Nunez, Gregory Bernard, Douglas Brown, Vickie Randleman, Julie Macialek, Dr. William Franck, Dr. Malali Gowda in Dr. Dean lab ; Dr. Sook-Young Park, Dr. Soonok Kim, Dr. Kyoungsu Kim, Dr. Hee-Sool Rho, Dr. Myeong-Hwan Chi, Dr. Jinhee

Choi, Dr. Ju-Young Park, Dr. Jaehyuk Choi, soon-to-be Dr. Jaeduk Goh, Yeokyeong Yoon, and fellow graduate students in Dr. Lee lab. I would never forget the wonderful times I spent with you in the lab. I just can't imagine how I could make it through difficult times without you. Thank you all!

Special appreciation goes to Dr. Sook-Young Park for being so encouraging and understanding, and to Dr. Sangho Park for trustful friendship, long talk, coffees and lunches together. Another round of special appreciation goes to Jaeyoung Choi and Gir-Won Lee for their enthusiasm and passion for data analysis. Without your contribution, this accomplishment weren't possible. Many thanks to Sewon Goh and Jung-Hwan Jeon for making life in USA much easier for me and my family.

Last but not least, I wish to express my deepest gratitude to my wife and son, my parents, my brother and sister for emotional and monetary support. Your very existence gave me the strength to endure difficult moments and made me better person in every possible way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Magnaporthe oryzae as a model plant pathogen for the study of evolution and epigenetics of fungi

Junhyun Jeon[1, 2]

[1]Department of Agricultural Biotechnology, College of Agriculture and Life science, Seoul National University, Seoul 151-921, Korea

[2]Functional Genomics Program, North Carolina State University, Raleigh, NC 27607, USA

INTRODUCTION

Magnaporthe oryzae (anamorph ; *Pyricularia oryzae*), a filamentous heterothallic ascomycetes fungus and casual agent of rice blast disease, is the most serious disease of cultivated rice. Each year enough rice to feed 60 million people is estimated to be destroyed by the disease. Since rice is the staple food that accounts for a significant proportion of the caloric and protein intake by individuals in many countries (http://www.irri.org/), the disease is a major threat to food security world-wide (Talbot, 2003). Indeed, the Center for Disease Control and Prevention listed rice blast as a significant biological weapon.

In the past, control of the disease was mainly through expensive and potentially hazardous chemicals. The study of *M. oryzae* was expected to provide target sites for chemical control. As time has passed, strategies have shifted from chemical control to the control exploiting host resistance. However, fungal genome appeared to be quite unstable and therefore, new races evolved quite quickly (Kang and Lee, 2000). To devise stable and environment-friendly control strategies, a comprehensive understanding of both the pathogen and the host is required.

During the last few decades, remarkable progress has been made on elucidating the fundamental molecular processes of rice blast disease. The genome sequence has greatly facilitated the genetic analysis of plant pathogens at the genomic level (Dean et al., 2005). Unlike most plant-fungal pathogen systems, the genome of its host plant, rice was sequenced as well (Goff et al., 2002; Yu et al., 2002). In addition, enormous amount genomic

information such as T-DNA mutant libraries, EST and microarray is available (Jeon et al., 2007; Mosquera et al., 2009; Kim et al., 2010).

Despite the progress made in understanding the genetic components of rice blast disease, little is known about the evolution and epigenetic aspects of the fungus. Here, I briefly describe the basic biology and life cycle of *M. oryzae* and then give a short introduction to the concept of experimental evolution and epigenetics with focus on DNA methylation.

BASIC BIOLOGY AND LIFE CYCLE OF *M. ORYZAE*

M. oryzae, like many foliar plant pathogens, is highly adapted to penetrate and subsequently colonize its host (Dean, 1997a). Once the asexual spores, conidia land on the surface of the host plant leaf, germ tubes emerge and migrate over the surface. Upon recognition of environmental cues such as surface hydrophobicity (Lee and Dean, 1994), a specialized infection structure, called the appressorium begins to develop. Many genetic components involved in this pre-penetration phase have been characterized (Dean, 1997b). Once the appressorium forms, then enormous turgor pressure up to 8Mpa accumulates within the appressorium. Using this turgor pressure, the fungus ruptures the plant cuticle layer and accesses the underlying tissues. Inside the host cell, the fungal hyphae are morphologically different from hyphae of vegetatively growing fungus: primary hypha differentiates into bulbous invasive hyphae that are known to be enclosed in a host plasma membrane (Kankanala et al., 2007). Next, it establishes a parasitic relationship with and derives nutrients from host cells to grow within the plant by secretion of effector proteins. Recent

work has revealed a special structure, named the biotrophic interfacial complex (BIC), that serves as a hub of protein secretion (Khang et al., 2010). This whole process ultimately leads to the development of disease lesions from which the fungus produces conidia for secondary infection cycles.

All these processes of invasion and colonization are typical of many plant pathogens. Furthermore, the fungus is easy to culture on the standard growth media away from its host and amenable to biochemical and genetic analysis. For these reasons, the fungus serves as an important experimental organism in the studies aimed at understanding the biology of fungal plant pathogens.

CONCEPT OF EXPERIMENTAL EVOLUTION

Mutation is the primary source of new genetic variation driving evolutionary changes in microbial populations (Adams, 2004). The main forces that determine the fate of new mutations are genetic drift and natural selection. Experimental evolution is the evolution of a laboratory population in simple, defined homogeneous environments. Serial passage experiments are a form of experimental evolution where the molecular and phenotypic changes can be monitored in real time, providing insights into the causes and consequences of evolution of a laboratory population (Ebert, 1998). In serial passage experiment, organisms of interest are transferred from one media (either artificial media or living host for the pathogen) to another. During the experiment, the organisms are propagated under defined conditions and their evolved characters are compared with those of the ancestral population.

A special case of serial passage experiment, in which selection is designed to be excluded by population bottleneck on the regular basis, is called a mutation accumulation (MA) experiment. A MA experiment, in which mutations are allowed to drift to fixation, has been a primary means of studying the rates and properties of new spontaneous mutations (Halligan and Keightley, 2009). Since most knowledge on molecular evolution comes from the comparison of extant biological sequences that survived many cycles of natural selection, MA experiment is an excellent way to infer the properties of the original source of variation.

In fungi, serial passage experiments were employed to study fitness of a fungal pathogen in alternative hosts (Cunfer, 1984) or to investigate spontaneous mutations in yeast (Lynch et al., 2008). In combination with the experimental evolution approach, application of rapidly developing high throughput sequencing technology is offering new opportunities to study the phenotypic and molecular evolution at an unprecedented scale in *M. oryzae* (chapter 2).

EPIGENETICS – DNA METHYLATION

Epigenetics is the term proposed by Conrad Waddington to describe events that could not be explained by genetic principles (Goldberg et al., 2007). Currently epigenetics means the study of a phenomenon that changes the expression or phenotypes (as a result of expression changes) without changing the underlying DNA sequences. In a broad sense, all developmental processes are thus “epigenetic”. At the molecular level, such epigenetic phenomena refer to covalent and non-covalent modification of DNA and histone proteins.

Recently, non-coding RNAs have been added to the list (Li and Liu, 2011; Rotblat et al., 2011).

Among all the chemical modifications of DNA and histone proteins, DNA methylation is the best characterized modification that is catalyzed by DNA methyltransferases. In plants and animals, DNA methylation is known to be involved in diverse processes including transposon silencing, X-chromosome inactivation, and imprinting (Suzuki and Bird, 2008). The importance of DNA methylation during development was demonstrated by the fact that deletion or disruption of DNA methyltransferases can result in morphological defects and embryonic lethality in plants and mammals, respectively (Finnegan et al., 1996; Okano et al., 1999; Finnegan et al., 2000). In social insects such as the honeybee, it was recently shown that DNA methylation is involved in caste differentiation probably through differential expression of genes and regulation of alternative splicing (Lyko et al., 2010). In fungi, DNA methylation is present at a low level and considered to exclusively target repeat sequences and transposons for transcriptional suppression as a genome defense mechanism (Selker et al., 2003; Suzuki and Bird, 2008).

Recently, high throughput sequencing technology combined with bisulphite sequencing (MethylC-seq) method proved an excellent tool to map genome-wide DNA methylation at single nucleotide resolution (Frommer et al., 1992; Clark et al., 2006; Cokus et al., 2008; Lister et al., 2008; Zemach et al., 2010). MethylC-seq was also used to examine DNA methylation patterns in fungal genomes, revealing evolutionary relationship with eukaryotic DNA methylation (Zemach et al., 2010). In *M. oryzae*, DNA methylation is known to be present in a retrotransposon, MAGGY (Nakayashiki et al., 2001) but genome-wide analysis

of DNA methylation is lacking and its role is obscure. In chapter 3, I will discuss the application of methylC-seq to *M. oryzae* in combination with genetic analysis as an attempt to address the contribution of DNA methylation to fungal development.

REFERENCES

Adams, J. 2004. Microbial evolution in laboratory environments. Res Microbiol 155:311-318.

Clark, S.J., Statham, A., Stirzaker, C., Molloy, P.L., and Frommer, M. 2006. DNA methylation: bisulphite modification and analysis. Nat Protoc 1:2353-2364.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452:215-219.

Cunfer, B.M. 1984. Change of Virulence of Septoria-Nodorum during Passage through Barley and Wheat. Ann Appl Biol 104:61-68.

Dean, R.A. 1997a. Signal pathways and appressorium morphogenesis. Annu. Rev. Phytopathol. 35:211-234.

Dean, R.A. 1997b. Signal pathways and appressorium morphogenesis. Annu Rev Phytopathol 35:211-234.

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.

Ebert, D. 1998. Experimental evolution of parasites. Science 282:1432-1435.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. 1996. Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. Proc Natl Acad Sci U S A 93:8449-8454.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. 2000. DNA methylation, a key regulator of plant development and other processes. Curr Opin Genet Dev 10:217-223.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 89:1827-1831.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus,

J., Macalma, T., Oliphant, A., and Briggs, S. 2002. A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92-100.

Goldberg, A.D., Allis, C.D., and Bernstein, E. 2007. Epigenetics: a landscape takes shape. Cell 128:635-638.

Halligan, D.L., and Keightley, P.D. 2009. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. Annu Rev Ecol Evol S 40:151-172.

Jeon, J., Park, S.Y., Chi, M.H., Choi, J., Park, J., Rho, H.S., Kim, S., Goh, J., Yoo, S., Park, J.Y., Yi, M., Yang, S., Kwon, M.J., Han, S.S., Kim, B.R., Khang, C.H., Park, B., Lim, S.E., Jung, K., Kong, S., Karunakaran, M., Oh, H.S., Kim, H., Kang, S., Choi, W.B., and Lee, Y.H. 2007. Genome-wide functional analysis of pathogenicity genes in the rice blast fungus. Nat Genet 39:561-565.

Kang, S., and Lee, Y.-H. 2000. Population structure and race variation of the rice blast fungus. Plant Pathology Journal 16:1-8.

Kankanala, P., Czymmek, K., and Valent, B. 2007. Roles for rice membrane dynamics and plasmodesmata during biotrophic invasion by the blast fungus. Plant Cell 19:706-724.

Khang, C.H., Berruyer, R., Giraldo, M.C., Kankanala, P., Park, S.Y., Czymmek, K., Kang, S., and Valent, B. 2010. Translocation of Magnaporthe oryzae effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell 22:1388-1403.

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.

Lee, Y.H., and Dean, R.A. 1994. Hydrophobicity of contact surface induces appressorium formation in *Magnaporthe grisea*. FEMS Microbiol. Lett. 115:71-75.

Li, L., and Liu, Y. 2011. Diverse small non-coding RNAs in RNA interference pathways. Methods Mol Biol 764:169-182.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133:523-536.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. 2010. The Honey Bee Epigenomes: Differential Methylation of Brain DNA in Queens and Workers. Plos Biol 8.

Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C.R., Dopman, E.B., Dickinson, W.J., Okamoto, K., Kulkarni, S., Hartl, D.L., and Thomas, W.K. 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc Natl Acad Sci U S A 105:9272-9277.

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.

Nakayashiki, H., Ikeda, K., Hashimoto, Y., Tosa, Y., and Mayama, S. 2001. Methylation is not the main force repressing the retrotransposon MAGGY in Magnaporthe grisea. Nucleic Acids Res 29:1278-1284.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247-257.

Rotblat, B., Leprivier, G., and Sorensen, P.H. 2011. A possible role for long non-coding RNA in modulating signaling pathways. Med Hypotheses 77:962-965.

Selker, E.U., Tountas, N.A., Cross, S.H., Margolin, B.S., Murphy, J.G., Bird, A.P., and Freitag, M. 2003. The methylated component of the Neurospora crassa genome. Nature 422:893-897.

Suzuki, M.M., and Bird, A. 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465-476.

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

Yu, J., Hu, S., Wang, J., Wong, G.K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X., Cao, M., Liu, J., Sun, J., Tang, J., Chen, Y., Huang, X., Lin, W., Ye, C., Tong, W., Cong, L., Geng, J., Han, Y., Li, L., Li, W., Hu, G., Li, J., Liu, Z., Qi, Q., Li, T., Wang, X., Lu, H., Wu, T., Zhu, M., Ni, P., Han, H., Dong, W., Ren, X., Feng, X., Cui, P., Li, X., Wang, H., Xu, X., Zhai, W., Xu, Z., Zhang, J., He, S., Xu, J., Zhang, K., Zheng, X., Dong, J., Zeng, W., Tao, L., Ye, J., Tan, J., Chen, X., He, J., Liu, D., Tian, W., Tian, C., Xia, H., Bao, Q., Li, G., Gao, H., Cao, T., Zhao, W., Li, P., Chen, W., Zhang, Y., Hu, J., Liu, S., Yang, J., Zhang, G., Xiong, Y., Li, Z., Mao, L., Zhou, C., Zhu, Z., Chen, R., Hao, B., Zheng, W., Chen, S., Guo,

W., Tao, M., Zhu, L., Yuan, L., and Yang, H. 2002. A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296:79-92.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328:916-919.

Chapter 2

Characterization of experimental evolution in the rice blast fungus, *Magnaporthe oryzae*

Junhyun Jeon[1, 3], Jaeyoung Choi[2], Gir-won Lee[2], Ralph A. Dean[3, 4], Yong-Hwan Lee[1, 4, 5, 6]

[1]Department of Agricultural Biotechnology, College of Agriculture and Life science, Seoul National University, Seoul 151-921, Korea

[2]Fungal Bioinformatics Laboratory, Seoul National University, Seoul 151-921, Korea

[3]Functional Genomics Program, North Carolina State University, Raleigh, NC 27607, USA

[4]Center for Integrated Fungal Research, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27607, USA

[5]Center for Fungal Pathogenesis, Seoul National University, Seoul 151-921, Korea

[6]Center for Fungal Genetic Resources, Seoul National University, Seoul 151-921, Korea

Correspondence: Yong-Hwan Lee

Phone) +82 2 880 4674

Fax) +82 2 873 2317

E-mail) yonglee@snu.ac.kr

SUMMARY

Little is known about phenotypic and molecular changes in plant pathogenic fungi that result from repeated in-vitro subculture. Here, we followed the changes in the phenotypes and DNA sequences of three independent wild-type cultures of the rice blast fungus Magnaporthe oryzae during 20 generations of growth in a serial passage experiment on oatmeal agar media. Three independent lineages showed no significant differences in growth rate but considerable variations in colony morphology and pigmentation within- and between lineages after 10 generations. Some but not all lineages started to show significant reduction in asexual sporulation and/or appressorium formation at about generation 15. Virulence tended to decrease in all three lineages, though to a varying extent, as subculturing continued. Interestingly, fungal cultures isolated from lesions caused by generation 20 cultures showed lower virulence than the founding culture, although they tended to be more virulent than the cultures from which they were derived. In an attempt to understand the basis for the differences between founding and terminal generation 20 cultures, we employed next-generation sequencing technology to examine genome changes with precision during our serial passage experiment. Analysis of sequencing data revealed that mutations accumulate rapidly prior to phenotypic changes and that both genetic drift and selection seem to contribute to shaping the mutational landscape. Examination of mutational effects on phenotypes through the use of T-DNA insertion mutants showed that the DNA sequence mutations are associated with the observed phenotypic changes. Furthermore, our data point out oxidative damages and UV as major source of mutation during subcultures in laboratory.

Taken together, our results suggest that idea of clonal propagation during subcultures in the laboratory should be reconsidered. We expect our endeavor to provide insights into stability of pathogenicity and genome evolution in plant pathogenic fungi.

INTRODUCTION

Subculture of fungi in various media is a routine practice in many laboratories. One pervasive idea that is rarely tested in the studies of microbial organisms such as fungi is that they can be propagated clonally without selection under laboratory settings. Studies such as Kohn *et al*. (Kohn et al., 2008) and Park *et al.* (Park et al., 2011) supported this view, although their conclusions are based on the use of several markers. This assumption is of particular importance in genetic analysis where functions of gene are inferred from observed phenotypes of mutants having the target gene deleted or mutated. For this inference to hold, the phenotypes should be attributed solely to the mutation introduced. If other mutations occur spontaneously during subculture or preparation of stock cultures, then the new mutations may induce secondary phenotypes or affect phenotypes of interest through genetic interactions with the solicited mutation. However, the rate at which such mutations occur and the molecular spectrum of their effects are largely unknown. Furthermore, selection that may be imposed by growing the fungus in the artificial media has not been appreciated at the molecular level to date.

Magnaporthe oryzae is the causal agent of rice blast disease that is one of the most devastating diseases in rice plants, and is a model plant pathogenic fungus with completely sequenced ~41Mb nuclear genome (Dean et al., 2005). The fungus usually disseminates via an asexual spore, called conidium that on the leaf surface, produces germ tube in the presence of water. The germ tube tip develops into a specialized infection structure, appressorium, in response to environmental cues such as surface hydrophobicity (Lee and

Dean, 1993). Using turgor pressure within the appressorium, the fungus mechanically penetrates into the host plant and grows inside it, leading to formation of disease lesions from which the fungus produces conidia for secondary infection. This infection debilitates host plants, causing yield losses or even death. It was reported that during a single rice growing season, fungal infection cycles 8 to 11 times under field conditions (Han et al., 1997). Although conidia are considered a major source of inoculum, vegetative hyphae can also form appressoria and cause disease when inoculated onto host plants as the recent study demonstrated (Kim et al., 2009).

To control rice blast disease, resistant cultivars have been developed and deployed. However, rice blast resistance often showed breakdown, leading to disease epidemics in the field (Kang and Lee, 2000). Insertion of a transposon into an avirulence gene was demonstrated to be a cause of resistance breakdown (Kang et al., 2001), yet this doesn't seem to be a prevailing mechanism. There are two different but not mutually exclusive explanations for resistance breakdown: 1. Mutation rate is accelerated when selection pressure is imposed by introduction of resistance cultivars; 2. Population of *M. oryzae* in the field is genetically diverse and introduction of resistance cultivars simply confers selective advantage on a particular subpopulation that has mutations in avirulence genes. While the former is technically-challenging to test, the latter possibility can be tested in the laboratory setting by growing the fungus without selection imposed by interaction with the host plant and examining mutations in the genome. If the latter is true, then one may see mutations in the avirulence genes and their regulatory sequences during the experiment.

A method that enables examination of spontaneous mutations in an organism is the mutation accumulation (MA) experiment (Halligan and Keightley, 2009), in which mutations are allowed to drift to fixation. During the MA experiment, genetic drift that is created by population bottleneck prevents the spontaneous mutations from being purged by natural selection, even if those mutations are slightly deleterious (Adams, 2004). The MA experiment has been successfully used to investigate the rate and spectrum of spontaneous mutations in model organisms such as *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Arabidopsis thaliana* (Lynch et al., 2008; Denver et al., 2009; Lynch, 2010; Ossowski et al., 2010). However, we propose that the MA experiment cannot be applied rigorously to filamentous fungi due to the cellular and genetic properties specific to the filamentous fungi. Unlike single-celled fungi such as yeast, individual cells in filamentous fungi do not exist in isolation but are physically connected to one another, forming a network structure, called hyphae. As a result, individual nuclei are not independent from one another. In combination with parasexualism (Noguchi et al., 2006; Schoustra et al., 2007), this network structure provides the means for filamentous fungi to combine mutations from different nuclei, preventing or attenuating the effect of clonal interference (Fogle et al., 2008; Kao and Sherlock, 2008). In addition, many filamentous fungi lack a sexual cycle in the field, rendering it difficult to define a generation in these multi-cellular eukaryotes.

Here we modified the MA experiment to include a mass of fungal tissue on artificial media and used it to characterize the experimental evolution of *M. oryzae*. Phenotypic changes observed in the derived strains suggest that the deleterious effects of mutations seem to be effectively buffered by the fungal genome. Analysis of whole genome sequencing data show

that mutations in the fungus accumulate more rapidly than previously thought and selection as well as genetic drift are acting to shape the mutational spectrum during subculture of the fungus on artificial media. The resulting mutational spectrum points out that main source of mutations during subculture under the laboratory condition are probably oxidative damages and UV, suggesting higher mutation rate in natural environment. To our knowledge, we examined the experimental evolution of filamentous fungi at single-nucleotide resolution for the first time. We believe this work would not only provide basic information in better understanding and predicting evolutionary path of *M. oryzae* but also aid in interpretation of genetic analysis for the fungus.

MATERIALS AND METHODS

Establishment of parental culture and serial passage culture scheme

A laboratory strain of *Magnaporthe oryzae*, 70-15 in Dr. Dean's lab was used to establish the parental culture as follows: This laboratory strain was grown on oatmeal agar plate for a week under constant fluorescent light. Spore suspension (~$5x10^4$ spores/ml) was prepared from this culture and used for spray-inoculation on a susceptible rice variety M2O2. One of the resulting susceptible lesions was excised and surface-sterilized, and transferred onto oatmeal agar plate. This plate culture was grown for 10 days and designated as parental culture (S_0). A patch of S_0 (~1.5x1.5cm) including both mycelia mass and conidia was taken and streaked on new oatmeal agar plate, and grown for a week to establish the next generation of culture (S_1). This successive subculturing was repeated until it reaches S_{20}. Starting from S_3, subculture was done in three plates, establishing three independent lineages (S_{3-1}, S_{3-2} and S_{3-3}) to S_{20} (S_{20-1}, S_{20-2} and S_{20-3}). Stock culture was prepared for every culture from each generation and stored at -70℃.

Vegetative growth, conidiation, and appressorium formation

Vegetative growth rate was measured on complete media, oatmeal agar and V8 juice agar plates on days 3 and 5 with three replicates. Ability to produce conidia was measured by counting the number of conidia from 5-day-old V8 agar plate. Conidia were collected by flooding the plate with 5ml of sterilized distilled water. The number of conidia was counted using hemacytometer under a microscope. Conidial germination and appressorium formation

were measured on plastic coverslip. Conidia were harvested from 8 to 10-day-old culture on oatmeal agar plate with sterilized distilled water. Conidial suspension of 40µl was pipetted onto plastic coverslip following adjustment of its concentration to ~ 2×10^4 spores/ml. Drops were placed in a moistened box and incubated at 25ºC. After 8h incubation, the percentage of conidia germinating and germinated conidia forming appressoria was determined by microscopic examination of at least 100 conidia per replicate in at least three independent experiments, with three replicates per experiment.

Pathogenicity assay

Conidia were harvested from 8 to 10-day old culture on oatmeal agar plate with sterilized distilled water. Conidial suspension was adjusted to ~5×10^4 spores/ml in concentration and spray-inoculated on susceptible rice seedlings (cv. M2O2) of 3 to 4 leaf stage. Inoculated plants were kept in dew chamber for 24h and allowed to grow within growth chamber for 5 days. Virulence was evaluated using three most diseased leaves, based on the following formula: virulence = 9 (for large lesions) × number of lesions + 3 (for medium-size lesions) × number of lesions + 1 (for small lesions or specs) × number of lesions.

DNA isolation and manipulation

Genomic DNA for Southern hybridization and sequencing was prepared by standard method (Sambrook and Russell, 2001). Restriction enzyme digestion, agarose gel separation, and DNA gel blotting were performed following the standard protocols (Sambrook and Russell, 2001). Hybridization was carried out in solution containing 6×SSC, 5×Denhardt's

solution, 0.5% SDS and 100µl ml^{-1} denatured salmon sperm DNA, at 65ºC. Blots were exposed to X-ray film.

DNA sequencing and read processing

We sequenced genomes from three independent cultures derived by 10 and 20 generations of plate cultures of the reference strain 70-15, for which genome sequence information was published in 2005 (Dean et al., 2005). Illumina (Illlumina, San Diego, CA) Genome Analyzer platform was used to obtain a depth of sequence coverage of between 30 and 100 in each culture.

In order to ensure the accuracy of further analysis, the following several steps are performed to filter the raw data (according to actual condition can be chosen). (1) Remove reads with a certain proportion of N's bases or low complexity reads (less than 10%, parameter setting at 9 bp); (2) Remove reads with a certain proportion of low quality (≤Q20) bases. (less than 40 bases); (3) Remove adapter contamination. (at least 15 bp overlap between adapter and reads, and at most 3 bp mismatches); (4) S_{10} and S_{20} lineage sequence was adjusted to the same size using random selection method. The above processes are applied to read1 and read2. After that, 1.5%-33.4% of the data is eliminated.

Correction of reference genome with S_0 mutations and mapping raw reads on the new reference genome

To compare S_{10} and S_{20} lineages against S_0, a new reference genome is needed. When S_0 was aligned to MG8, a total of 504 variations were found in S_0. These were applied on MG8

to make a new reference genome, MG8_S0. This will be used for calling variations in S_{10} and S_{20} lineages. Annotation of the new modified reference genome, MG8_S0, was resolved by using annotation migration tool, Scipio v1.4 (Keller et al., 2008). To align Illumina sequencing reads on the MG8 reference genome, BWA (Burrows-Wheeler Aligner) 0.6.1-r104 (Li and Durbin, 2010) was used.

Variant calling and analysis of variants

To find structural variations, SAMtools 0.1.18 (Li et al., 2009) was used. Among all the possible variants, we restricted our analysis to the one with at least four sequencing depth and over 0.5 of frequency. To calculate mapping coverage on the reference genome, Bed file format was used. To convert BAM file to Bed, bamToBed in BEDtools package was used with default options. The snpEff package v2.0.5d (http://snpeff.sourceforge.net/) was used to investigate the followings: 1) to identify where SNPs/InDels are located (Intergenic/Exon/Intron); 2) to investigate transitions and transversions; 3) to find synonymous and non-synonymous changes when a codon was changed by SNP.

Validation of mutations by Sanger sequencing

The loci to be checked by Sanger sequencing were randomly selected from the dataset and subjected to PCR amplification. Successfully amplified loci were treated with ExoSAP-IT® For PCR Product Clean-UP (Affimetrix) and directly capillary-sequenced with primers used in PCR.

RESULTS

Modified mutation accumulation experiment

The previous studies using the MA experiment in model organisms subjected the population to the extreme degree of bottleneck. For example, only a single seed for the plant or a single cell for the yeast was used to establish the next generation. This enables any mutations within the genome of selected individual to be fixed. New mutations accumulate in addition to the fixed mutations. Such extreme bottleneck is useful in estimating mutation parameters such as mutation rate. However, this extreme bottleneck is unlikely in natural and laboratory settings where filamentous fungi are studied. In many laboratories, subculture of the fungus is a routine practice carried out by transferring a portion of existing culture, leading inevitably to genetic drift to some extent.

In order to understand what mutations occur at what rate and how these mutations affects the traits of a fungal pathogen, *M. oryzae*, we modified the MA experiment by serially transferring a portion of cultures instead of single conidia (asexual spore with three cells) on oatmeal agar plates. The laboratory strain 70-15 was used due to the availability of genome sequence information. The modified MA experiment was done as described in Figure 1. Initial culture was derived from a disease lesion caused by the fungus on a rice leave to ensure full virulence and was designated S_0. The S_0 was subcultured on oatmeal agar plate, forming S_1. The subculturing continued up to S_{20}. Starting from S_4, the fungus was subcultured in three independent lineages (named lineage 1, lineage 2, and lineage 3). Having multiple lineages is expected to give more information on mutation processes of this MA

experiment than using a single lineage. All the derived strains were stored as permanent stock cultures in -70℃ for later uses. After the derived strains were established, phenotypic changes were first monitored during the experimental evolution of the fungus.

Colony morphology and growth rate of the derived strains

Since it is cumbersome and time-consuming to check the phenotypic changes of all the derived strains, the phenotypes of S_0 were compared with every fifth derived strains (S_5, S_{10}, S_{15}, and S_{20}). Colony morphology was checked first. Colony morphologies are subtle phenotype that changes frequently in a single genotype depending on environments. Due to the subtle nature, the norm of reaction for colony morphology of wild-type was explored. Based on the wild-type norm of reaction, seemingly aberrant morphology of the derived strains was repeatedly checked for reproducibility. We confirmed that in the lineage 1, aerial hyphae and mycelia melanin decreased, and that in the lineage 2, mycelia melanin was excessive (Figure 2). Colony morphology of the lineage 3 didn't change, compared to that of wild-type. The observed changes in the lineages 1 and 2 started to appear at S_{15}. However, the growth rate didn't show dramatic differences among the derived strains (Figure 3).

Production of asexual spores, germination and appressorium formation in the derived strains

When the ability of the derived strains to produce asexual spores was monitored, it was found that the lineage 1 of S_{15} and S_{20} barely produced conidia (Figure 4). To find out when this impairment began, conidiation of the lineage 1 in S_{11} to S_{14} was quantified. I found that

S_{11-1} already showed significant reduction in conidiation (25% of S_0) and conidiation of S_{14-1} was comparable to that of S_{15-1}. Germination from conidia was normal for every derived strain (Figure 5). Appressorium formation started to be defective from S_{15}. Especially in the lineage 1, almost no conidia developed appressoria.

Virulence of the derived strains

Virulence of the derived strains was tested using a susceptible rice cultivar M2O2. Virulence of S_{15-1} was not tested due to the lack of spores. Unlike colony morphology, conidiation, and appressorium formation, reduction in virulence was first observed in S_{10-2} (Figure 6). The defect in virulence was exacerbated in S_{15} and S_{20}, regardless of the lineages (Figures 7 and 8). The derived strains in S_{15} and S_{20} didn't cause as many lesions as S_0 and the size of lesions were smaller than those of S_0. One observation worth noting is that the virulence of the lineage 3 was greatly reduced in S_{15} but significantly restored in S_{20}, though it was still less virulent than S_0. Except for lineage 1 whose appressorium formation rate was less than 10%, the other derived strains showed appressorium formation rate of over 70%. This suggests that the reduced virulence cannot be attributed solely to defects in appressorium formation. It is conceivable that the mutations that have occurred during the MA experiment may be involved in appressorium-mediated penetration of the cuticle layer or invasive growth inside host cells.

Virulence and genetic heterogeneity of population

A few lesions observed on leaves inoculated with the derived strains pose an interesting question about the relationship between virulence and genetic heterogeneity of a population. By nature, our modified MA experiment would allow unfixed mutations to remain in the population. Given the presence of unfixed mutations that may affect virulence, are the lesions caused by subpopulation that doesn't have those mutations? If that's the case, one would predict that the fungus isolated from the lesions would be more virulent than the strain from which it was derived. To test this, we first isolated the fungus separately from the lesions caused by $S_{20\text{-}1}$, $S_{20\text{-}2}$, and $S_{20\text{-}3}$. The lesions caused by S_0 and S_1 were included as a control. We tried to isolate fungus from at least three lesions for each derived strain but obtained only one for $S_{20\text{-}1}$ or two isolates for $S_{20\text{-}2}$ and $S_{20\text{-}3}$, suggesting that the fungus in some lesions was not viable. Spores were harvested from these lesion isolates and used for spray-inoculation onto rice plants. Interestingly, the five lesion isolates appeared to be more virulent than the derived strains themselves, although the differences were not statistically significant (Figure 9). This result supports our prediction about genetic heterogeneity of the population.

Fingerprinting of the derived strains and their lesion isolates with MGR586

In an attempt to investigate the mutations that have occurred during the MA experiment, DNA fingerprinting was carried out for the derived strains and those isolated from lesions using MGR586 (Romao and Hamer, 1992) as a probe. Our DNA fingerprinting analysis, however, revealed little differences among strains tested (Figure 10). This suggests that the

derived strains are really 'derived' from S_0 and resolution of fingerprinting approach may be too low to detect mutations.

Genome-sequencing of the derived strains

To catalogue the mutations within the genome of the derived strains, we employed next generation sequencing technology for sequencing of S_0 (1 strain), S_{10} (3 strains), and S_{20} (3 strains). Genomic DNAs of mycelia tissues were extracted and sequenced, yielding sequencing reads that correspond to about 30 to 100× coverage (Table 1). Large differences in amount of sequencing reads among the derived strains are mainly due to the use of different version of sequencing platforms: Illumina GAIIx for S_0 and $S_{20\text{-}1}$, and HiSeq system for the rest. To deal with possible biases resulting from vastly uneven sequencing reads, we randomly selected the sequencing reads to make raw read depth equal (~40× except S_0) (Table 2). Our sequencing data covered at least 96% of nucleotide sites within the fungal genome (Table 3). Since it is not known how identical S_0 is to the strain used for establishing reference genome sequences, we aligned the resulting S_0 sequences to the reference genome and subtracted the existing mutations, making a corrected reference genome. Sequencing reads from the derived strains were aligned to this corrected reference genome.

To call mutations, only the sites covered with at least four but less than 100 sequencing reads were considered. This allowed us to call mutations with high confidence for on average 96.3% of nucleotide sites in the genome (Table 4). By the nature of the experiment, we expected to see different types of mutations in terms of fixation: Some of the mutations are fixed in the population and the others are not yet fixed or in the course of fixation. We

reasoned that low frequency mutations would complicate our analysis because they are likely to be discarded during genetic drift and are unlikely to have major effects on phenotypic changes of the derived strains. Furthermore, calls of low frequency mutation can be dubious due to possible sequencing error. Therefore, we decided to restrict our analysis to mutations having medium frequency of 0.5.

Validation of predicted mutations by Sanger sequencing

To validate our high throughput sequencing data, we randomly selected 18 predicted mutations for Sanger sequencing (Table 5). One of the loci was not amplified in the PCR. For the remaining 17 loci that were successfully amplified, we sequenced the PCR products and compared the resulting sequences with the corrected references at the corresponding loci by BLAST search. From fifteen loci, we confirmed the variations predicted by high throughput sequencing method. For the other two loci, the two independent methods didn't agree. One of the two loci was an unfixed mutation, providing the possibility that PCR was somehow biased toward the one base against the other. Our validation experiment suggests that our dataset is reliable enough to carry out further analysis.

Mutations that occurred during MA experiment among the derived strains

By applying the read depth and frequency criteria, we were able to identify varying number of mutations from the derived strains (Figure 11). Overall, the number of mutations lies between 200 and 350 per strain. The smallest number was found in $S_{20\text{-}2}$ and the largest was in $S_{20\text{-}1}$. Interestingly, the number of mutations didn't seem to be proportional to the

number of subcultures. Regardless of lineages, the number of mutations in the derived strain reached about 250 at S_{10} and didn't increase significantly in S_{20} except in $S_{20\text{-}1}$. Among all the mutations detected, more than 90% of mutations were single nucleotide polymorphisms (SNP) and, insertions and deletions were less than 10% of overall mutations. Around 30 to 60% of mutations were fixed in the population, depending on the lineages (Figure 12). This mixture of fixed and unfixed mutation is what we expected to see and reflects the experimental design involving population of cells rather than a single cell. Most of insertions and deletions were one or two bases long (Figure 13).

Mutational bias among chromosomes

Given the number of mutations, we asked if mutations were evenly distributed among chromosomes by comparing the number of observed mutations with the expected number among chromosomes (Figures 14 and 15). The expected number of mutations in each chromosome was calculated based on the length of corresponding chromosome. We found that there was a significant bias in the distribution of mutations among chromosomes (χ^2 test, 6 df, $P = 2.71\times10^{-17}$ for S_{10} and $P = 1.27\times10^{-16}$ for S_{20}). Chromosomes 1, 4, and 6 tended to have more mutations than expected, whereas chromosomes 2 and 5 tended to have less than expected. A large standard deviation in S_{20} is due to the higher number of mutations in $S_{20\text{-}1}$ than in other lineages. It is not clear what caused this bias among chromosomes: no association was detected between mutations and transposable elements.

Dynamics of mutations among and between lineages

To explore the dynamics of mutations, we examined how many mutations overlap between S_{10} and S_{20} or among lineages (Figure 16). Over 50% of mutations were shared between S_{10} and S_{20} except the case between $S_{10\text{-}1}$ and $S_{20\text{-}1}$. As expected, many mutations that were present in S_{10} but lost in S_{20} were not fixed ones, showing that even the medium frequency mutations have high probability of being lost during subcultures. Notably, new mutations occurred in S_{20}, while some of the mutations that existed in S_{10} were being lost (left column of Figure 16). Since the number of mutations appeared to increase rapidly before S_{10} and stabilize in S_{20}, we speculate that mutation rate is initially high during the course of subcultures and later decreases to reach equilibrium where the rate of occurrence and extinction become equal.

In our experiment, we kept three independent lineages. Comparison of the mutations in three lineages showed that a significant portion of mutations (~30 to 70%) are conserved among them (right column of Figure 16). This suggests two possibilities that are not mutually exclusive. First, the genome of *M. oryzae* has mutational hotspots that mutate more frequently than other sites. Alternatively, selection is acting on the genome to retain mutations at particular set of sites. As a way to test whether selection was affecting mutations, we examined the distribution of mutations in coding sequences and non-coding sequences. Expectation under no selection is that the mutations would be evenly distributed between coding and non-coding sequences in proportion to their length. We calculated the expected number of mutations in coding and non-coding sequences and compared it with the observed number of mutations in each region. The result showed strong bias toward the non-coding

sequences (χ^2 test, 3 df, $P < 1\times10^{-10}$ for all cases), indicating that selection might have driven out the mutations in the coding sequences. It is uncertain whether selection is due to oatmeal agar plate or within-population competition. Although the selection might exert its influence, we cannot rule out the possibility that there are mutational hotspots especially in the non-coding part of genome of *M. oryzae*. However, we don't know whether, under different experimental settings, the observed mutations can be replicated by repeated subcultures.

Transition vs. Transversion

We analyzed the mutational spectrum in the derived strains by considering the six base-substitution types. We found that, in all derived strains, the rate of transition (Ts) base-substitution are nearly four to five times as many as transversion (Tv) base substitutions, leading to the Ts/Tv ratio between 4 and 5 (Table 7), which is much greater than the random expectation of 0.5. This indicates the overabundance of transitions, compared to transversions. In transitions, G:C→ A:T and A:T → G:C occurred at similar rate. Although we don't know the particular sources of DNA damage or error leading to A:T↛ G:C, it is known that spontaneous deamination of cytosine or 5-methylcytosine is the major source of G:C → A:T transitions (Wagner et al., 1992; Lindahl, 1996). However, DNA methylation is unlikely to be a major cause of this because methylated cytosine accounts for only 0.2% of total cytosines. We propose that UV-induced mutation and/or oxidative damage that can cause G:C→ A:T transitions (Kreutzer and Essigmann, 1998) played a major role in mutagenesis of the derived strains.

Contribution of mutations to phenotypic changes

To relate the mutations to the phenotypic changes observed in the derived strains, we took advantage of information from the literature and T-DNA insertions mutants we have in the lab. We found that two genes among the ones having mutations in the derived strains were reported in the literature. Those were *THL1* and MgCdc42 (Gao et al., 2002; Zheng et al., 2009). The mutations in two genes were found in all the derived strains examined. Unfortunately, the mutations were either in promoter region (*THL1*) or intron (*MgCdc42*), rendering it difficult to make direct connection between mutations and phenotypes, although we cannot exclude the possibility that the mutations can interfere with transcription or splicing.

We therefore turned to T-DNA resources. Over 20,000 strains of T-DNA insertion mutants and their phenotypes are archived in Center for Fungal Genetics Resources (http://knrrb.knrrc.or.kr/index.jsp?rrb=cfgr) and ATMT database (http://atmt.snu.ac.kr) (Jeon et al., 2007). To make the use of these resources, we extracted the annotated genes that have mutations in exon sequences of the derived strains and are tagged by T-DNA at the same time (Table 8). A total of seven annotated genes were found to meet the criteria. Most of mutations were amino-acid changing i.e. non-synonymous mutations. However, we also found a synonymous mutation, a frame-shift mutation by insertion, and a codon deletion mutation by nucleotides deletion. The derived strain that had the most mutations was $S_{20\text{-}1}$, which is consistent with the fact that $S_{20\text{-}1}$ displayed more phenotypic defects than any other strains. T-DNA insertions in genes having mutations caused defects in growth, appressorium formation, and pathogenicity. Since T-DNA insertions are mostly in promoter regions of the

genes, they would probably affect the transcription of downstream genes. So T-DNA insertions would have different effects at the molecular level from mutations we found in the derived strains. However, phenotypic consequences of those effects are likely to be similar to each other. We thus conclude that these mutations would have likely contributed to the phenotypic changes observed in the derived strains.

DISCUSSION

For genetic studies of fungi, subcultures on artificial media under laboratory conditions are routine practice. The idea or assumption that underlies subculture practice is that fungal culture regenerated from clonally related isogenic founders is stable and thus invariant at both phenotypic and genotypic level. However, this idea has been rarely tested and no whole-genome scale investigation has been undertaken. In this study, we monitored the changes in phenotypes and genome sequences that occurred during repeated *in-vitro* subcultures of *M. oryzae* on oatmeal agar media.

We expected that mutations in DNA sequences involved in conidation, appressorium formation, and virulence would be nearly neutral because the fungus is propagated by human intervention and is away from its host plants. Consistent with our expectation, we observed the defects in aforementioned traits. However, phenotypes were quite stable, since little changes were detected until S_{10}. This observation is consistent with the previous report (Park et al., 2011), though they used single-spore as an inoculums instead of a patch of fungal culture. In our high throughput sequencing data, 200-350 mutations per the derived strain were detected. The number of mutations seemed to increase rapidly up to S_{10} and reach equilibrium somewhere between S_{10} and S_{20}. Although it is difficult to interpolate with a few data points, it is tempting to speculate that accumulation of mutations and their phenotypic changes display opposing pattern during the course of experimental evolution. This discordance may reflect the capacity of fungal genome to buffer mutations. That is, a certain

number of mutations is somehow tolerated and more mutations than a critical number are deleterious to the fungus, probably through genetic interactions.

After S_{10}, all three lineages showed varying degree of defects in different traits and these defects were exacerbated in following generations, suggesting the effect of genetic drift and accumulation of deleterious mutations during subcultures. Indeed, our sequencing data showed the effect of genetic drift: distinct mutations in different lineages. However, a significant portion of mutations were shared among lineages, suggesting the action of selection. We reasoned that if selection is acting on the genome of the derived strains, then we would see the bias in distribution of mutations between coding and non-coding sequences. Comparison of number of mutations in coding and non-coding sequences revealed strong mutational bias toward the non-coding sequences, implying the presence of selection against the mutations in coding sequences. These results indicate that mutational landscape of the fungal genome under subculture scheme is shaped and maintained by both genetic drift and selection.

One question that arises from our data is what would happen if S_{10} or S_{20} cultures become founding population and are sub-cultured for a long enough time. Will they accumulate new mutations rapidly up to next 10 generations and reach equilibrium again as they did before? Or will they stay in equilibrium as they do now? One clue to the question is that there were already many mutations (over 500 mutations) in S_0, compared to the reference genome sequences of 70-15. Although we don't know the exact history of S_0 before the experiment was undertaken, it is obvious that the fungus went through repeated subcultures since the genome sequencing of original strain. This suggests that mutations randomly occur in the

genome but selection from growing the fungus on artificial media can largely determine the set of mutations that are retained in the genome.

Our virulence test with the lesion isolates of the derived strains suggested the genetic heterogeneity of population in the fungal culture. It has been recognized that the fungus can cause different types of lesions on the single leaf of host plant (Valent et al., 1991). Different lesion types can be attributed either to differences in micro-environment of plant surface where the fungal spores are placed or to difference in genotypes of fungal spores inoculated. The relative contribution of the two factors to the lesion formation is unclear, but our data suggest that genetic heterogeneity at the population level can explain variation in disease lesions to some extent.

One of our predictions during the experiment was that we will be able to see mutations in known pathogenicity genes and/or avirulence genes if we effectively eliminate host selection by growing the fungus away from it. However, we found mutations only in an intron of *MgCdc42* and no mutations in the avirulence gene, *Avr-Pita*. This is probably because many known pathogenicity genes are the one required for normal growth as well host infection and thus supports that pathogenicity of *M. oryzae* is not a simple phenotype that can be determined by several factors but complicated biological phenomena that involves essential components of cells. No mutations in the avirulence gene may suggest either that the time given during 20 subcultures wasn't long enough for mutations to occur or that there are unknown mechanisms keeping mutation in avirulence genes in check. No avirulence genes were shown to be virulence factors in compatiable interaction of *M. oryzae*, providing an

evolutionary conundrum for their maintenance in the fungal genome. Such stability of avirulence genes is the topic that deserves further work in the future.

We found that major source of mutations in *M. oryzae* during subcultures under laboratory conditions are oxidative damage and UV probably coupled with fidelity error in repair. This implies that the mutation rate in nature can be higher than that reported here, because UV radiation during our experiment was probably much lower than in natural conditions.

Taken together, we investigated the experimental evolution of the model plant pathogenic fungus, *M. oryzae* at both phenotypic and genotypic level. We are aware of no other similar analysis of experimental evolution of filamentous fungi, so the generality of our observations is unknown. However, it is clear that filamentous fungi pose unique challenges in terms of life style and biology. Therefore, we believe that our endeavor would provide not only the basis for future work to be compared but also information that can aid in genetic analysis of this fungus. Our results also call attention to new opportunities for experimental evolution to explore the long-term dynamics of mutations in relation to compatible and incompatible interactions of fungi with their host plants.

REFERENCES

Adams, J. 2004. Microbial evolution in laboratory environments. Res Microbiol 155:311-318.

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.

Denver, D.R., Dolan, P.C., Wilhelm, L.J., Sung, W., Lucas-Lledo, J.I., Howe, D.K., Lewis, S.C., Okamoto, K., Thomas, W.K., Lynch, M., and Baer, C.F. 2009. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc Natl Acad Sci U S A 106:16310-16314.

Fogle, C.A., Nagle, J.L., and Desai, M.M. 2008. Clonal interference, multiple mutations and adaptation in large asexual populations. Genetics 180:2163-2173.

Gao, W.M., Khang, C.H., Park, S.Y., Lee, Y.H., and Kang, S.C. 2002. Evolution and organization of a highly dynamic, subtelomeric helicase gene family in the rice blast fungus *Magnaporthe grisea*. Genetics 162:103-112.

Halligan, D.L., and Keightley, P.D. 2009. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. Annu Rev Ecol Evol S 40:151-172.

Han, S.S., Ra, D.S., Choi, S.H., and Kim, C.K. 1997. Population dynamics of *Pyricularia grisea* during leaf and panicle blast stages in the same field. Korean J. Plant Pathol. 13:408-415.

Jeon, J., Park, S.Y., Chi, M.H., Choi, J., Park, J., Rho, H.S., Kim, S., Goh, J., Yoo, S., Park, J.Y., Yi, M., Yang, S., Kwon, M.J., Han, S.S., Kim, B.R., Khang, C.H., Park, B., Lim, S.E., Jung, K., Kong, S., Karunakaran, M., Oh, H.S., Kim, H., Kang, S., Choi, W.B., and Lee, Y.H. 2007. Genome-wide functional analysis of pathogenicity genes in the rice blast fungus. Nat Genet 39:561-565.

Kang, S., and Lee, Y.-H. 2000. Population structure and race variation of the rice blast fungus. Plant Pathol. J. 16:1-8.

Kang, S., Lebrun, M.H., Farrall, L., and Valent, B. 2001. Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. Mol Plant Microbe Interact 14:671-674.

Kao, K.C., and Sherlock, G. 2008. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. Nat Genet 40:1499-1504.

Keller, O., Odronitz, F., Stanke, M., Kollmar, M., and Waack, S. 2008. Scipio: using protein sequences to determine the precise exon/intron structures of genes and their orthologs in closely related species. BMC Bioinformatics 9:278.

Kim, S., Park, S.Y., Kim, K.S., Rho, H.S., Chi, M.H., Choi, J., Park, J., Kong, S., Goh, J., and Lee, Y.H. 2009. Homeobox transcription factors are required for conidiation

and appressorium development in the rice blast fungus *Magnaporthe oryzae*. PLoS Genet 5:e1000757.

Kohn, L.M., Schaffer, M.R., Anderson, J.B., and Grunwald, N.J. 2008. Marker stability throughout 400 days of in vitro hyphal growth in the filamentous ascomycete, *Sclerotinia sclerotiorum*. Fungal Genet Biol 45:613-617.

Kreutzer, D.A., and Essigmann, J.M. 1998. Oxidized, deaminated cytosines are a source of C --> T transitions in vivo. Proc Natl Acad Sci U S A 95:3578-3582.

Lee, Y.H., and Dean, R.A. 1993. cAMP Regulates Infection Structure Formation in the Plant Pathogenic Fungus *Magnaporthe grisea*. Plant Cell 5:693-700.

Li, H., and Durbin, R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589-595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.

Lindahl, T. 1996. The Croonian Lecture, 1996: Endogeneous damage to DNA. Phil. Trans. R. Soc. Lond. B 351:1529-1538.

Lynch, M. 2010. Rate, molecular spectrum, and consequences of human mutation. Proc Natl Acad Sci U S A 107:961-968.

Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C.R., Dopman, E.B., Dickinson, W.J., Okamoto, K., Kulkarni, S., Hartl, D.L., and Thomas, W.K. 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc Natl Acad Sci U S A 105:9272-9277.

Noguchi, M.T., Yasuda, N., and Fujita, Y. 2006. Evidence of Genetic Exchange by Parasexual Recombination and Genetic Analysis of Pathogenicity and Mating Type of Parasexual Recombinants in Rice Blast Fungus, *Magnaporthe oryzae*. Phytopathology 96:746-750.

Ossowski, S., Schneeberger, K., Lucas-Lledo, J.I., Warthmann, N., Clark, R.M., Shaw, R.G., Weigel, D., and Lynch, M. 2010. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science 327:92-94.

Park, S.-Y., Chi, M.-H., Milgroom, M.G., Kim, H., Han, S.-S., Kang, S., and Lee, Y.-H. 2011. Genetic stability of *Magnaporthe oryzae* during successive passages through rice plants and on artificial medium. The Plant Pathology Journal 26:313-320.

Romao, J., and Hamer, J.E. 1992. Genetic organization of a repeated DNA sequence family in the rice blast fungus. Proc Natl Acad Sci U S A 89:5316-5320.

Sambrook, J., and Russell, D.W. 2001. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schoustra, S.E., Debets, A.J., Slakhorst, M., and Hoekstra, R.F. 2007. Mitotic recombination accelerates adaptation in the fungus Aspergillus nidulans. PLoS Genet 3:e68.

Valent, B., Farrall, L., and Chumley, F.G. 1991. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics 127:87-101.

Wagner, J.R., Hu, C.C., and Ames, B.N. 1992. Endogeneous oxidative damage of deoxycytidine in DNA. Proc Natl Acad Sci U S A 89:3380-3384.

Zheng, W., Zhao, Z.Y., Chen, J.S., Liu, W.D., Ke, H.Y., Zhou, J., Lu, G.D., Darvill, A.G., Albersheim, P., Wu, S.C., and Wang, Z.H. 2009. A *Cdc42* ortholog is required for penetration and virulence of *Magnaporthe grisea*. Fungal Genetics and Biology 46:450-460.

Table 1. Basic statistics of whole genome sequencing

Strain	Size	Raw Read Depth	Depth of Aligned Reads
S0	1295012592	31.56	29.23
S10-1	4095725448	99.83	82.34
S10-2	4156334754	101.31	76.97
S10-3	4090114358	99.69	75.53
S20-1	1706770800	41.60	37.56
S20-2	4095473811	99.82	76.81
S20-3	4048043505	98.67	80.15

Table 2. Average sequencing depth of raw and mapped reads

Lineage	Size	Raw Read Depth	Depth of Aligned Reads
S0	1275597288	31.06	28.97
S10-1	1643712344	40.03	35.73
S10-2	1643712279	40.03	35.06
S10-3	1643712279	40.03	35.17
S20-1	1643712264	40.03	36.83
S20-2	1643712279	40.03	35.03
S20-3	1643712279	40.03	36.18

Table 3. Percentage of covered nucleotide sites present within the genome of *M. oryzae*

Lineage	Reference genome size (bp)	Covered sites (bp)	Ratio
S0	41062686	40958442	99.75%
S10-1.again	41062692	40878231	99.55%
S10-2	41062692	40702514	99.12%
S10-3	41062692	40569126	98.80%
S20-1	41062692	40952069	99.73%
S20-2	41062692	39476371	96.14%
S20-3	41062692	40501159	98.63%

Table 4. Proportion of nucleotide sites covered by at least four sequencing reads

Lineage	Reference genome size (bp)	Over 4x (bp)	Ratio
S0	41062686	40906171	99.62%
S10-1	41062690	40528236	98.70%
S10-2	41062690	39766656	96.84%
S10-3	41062690	39316364	95.75%
S20-1	41062690	40860128	99.51%
S20-2	41062690	36333733	88.48%
S20-3	41062690	38988961	94.95%

Table 5. List of the mutations that were validated by Sanger sequencing

Strain	Chromosome	Position	Change[a]		Frequency	Sanger Sequencing[b]
S10-1	supercont8.1	811343	T	A	<1	Yes
	supercont8.4	1774729	G	A	1	Yes
	supercont8.7	3158861	C	T	<1	Yes
	supercont8.1	3158922	C	T	<1	Yes
S10-2	supercont8.1	811329	T	A	<1	Yes
	supercont8.4	5213688	G	A	1	Yes
	supercont8.7	28760	T	G	1	Yes
	supercont8.6	4127585	A	G	<1	Yes
S103	supercont8.2	5876403	T	C	1	Yes
	supercont8.3	3001414	A	G	<1	No
S20-1	supercont8.6	3254939	A	G	<1	Yes
	supercont8.6	3254962	T	C	<1	Yes
S20-2	supercont8.1	811617	G	A	<1	Yes
	supercont8.1	811624	C	T	<1	Yes
	supercont8.1	811641	T	G	<1	Yes
	supercont8.6	3256089	T	C	<1	Yes
S20-3	Supercont8.1	5068128	T	C	<1	NA[c]
	supercont8.6	2352336	T	C	1	No

[a]Change of nucleotide from corrected reference to the evolved genome.

[b]Whether the changes are confirmed by Sanger sequencing.

[c]Not amplificed in PCR.

Table 6. Distribution of mutations in coding sequences (CDS) and non-coding sequences (nonCDS)

	Type	S10-1		S10-2		S10-3		S20-1		S20-2		S20-3	
		CDS	**nonCDS**	**CDS**	**nonCDS**	**CDS**	**nonCDS**	**CDS**	**nonCDS**	**CDS**	**nonCDS**	**CDS**	**nonCDS**
Observed	**Insertion**	4	8	2	13	3	16	5	23	1	15	2	12
	Deletion	1	13	1	13	1	20	2	22	0	16	2	22
	SNP	66	257	64	237	69	261	83	377	61	248	66	246
Expected	**Insertion**	4.98	7.02	6.22	8.78	7.88	11.12	11.62	16.38	6.64	9.36	5.81	8.19
	Deletion	5.81	8.19	5.81	8.19	8.71	12.29	9.96	14.04	6.64	9.36	9.96	14.04
	SNP	134.02	188.98	124.89	176.11	136.92	193.08	190.86	269.14	128.21	180.79	129.45	182.55

Table 7. Transitions and transversions in the derived strains

Lineage	Transition	Transversion	Ts/Tv Ratio
S10-1	205	45	4.56
S10-2	193	39	4.95
S10-3	207	43	4.81
S20-1	239	58	4.12
S20-2	158	37	4.27
S20-3	198	41	4.83

Table 8. List of mutations that are found in coding sequences with T-DNA insertion lines available

Locus	Strain	Chr.	Position	Change		Type	Freq	Exon Rank	Effect	T-DNA			
										ID	Position	Distance	Phenotypic defects
MGG_01344T0	**S20-1**	supercont8.2	3187723	A	C	SNP	<1	1	Nonsynonymous	ATMT0207B6	5'	1337bp	Pathogenicity
MGG_03703T0	**S10-1**	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous	ATMT0047B3	5'	306bp	Pathogenicity
	S10-2	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous				
	S10-3	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous				
	S20-1	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous				
	S20-2	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous				
	S20-3	supercont8.4	969548	T	C	SNP	1	2	Nonsynonymous				
MGG_04923T0	**S20-1**	supercont8.3	3965142	G	A	SNP	1	2	Nonsynonymous	ATMT0057B2	5'	839bp	Appressorium, Pathogenicity
MGG_06064T0	**S10-1**	supercont8.3	2445604	G	A	SNP	1	3	Nonsynonymous	ATMT0045B1	5'	859bp	Appressorium, Pathogenicity
	S10-2	supercont8.3	2445604	G	A	SNP	1	3	Nonsynonymous				
	S10-3	supercont8.3	2445604	G	A	SNP	1	3	Nonsynonymous				
	S20-1	supercont8.3	2445604	G	A	SNP	1	3	Nonsynonymous				
MGG_08208T0	**S20-1**	supercont8.2	6082401	A	C	SNP	<1	1	Synonymous	ATMT0457B6	Exon	0	No defect
MGG_13324T0	**S20-1**	supercont8.3	67254	*	+G	INS	1	2	Frame shift	ATL0379	Exon	0	Not available
MGG_15990T0	**S20-1**	supercont8.1	194398	*	-ACA	DEL	1	1	Codon deletion	ATMT0215B6	5'	362bp	Growth, Pathogenicity

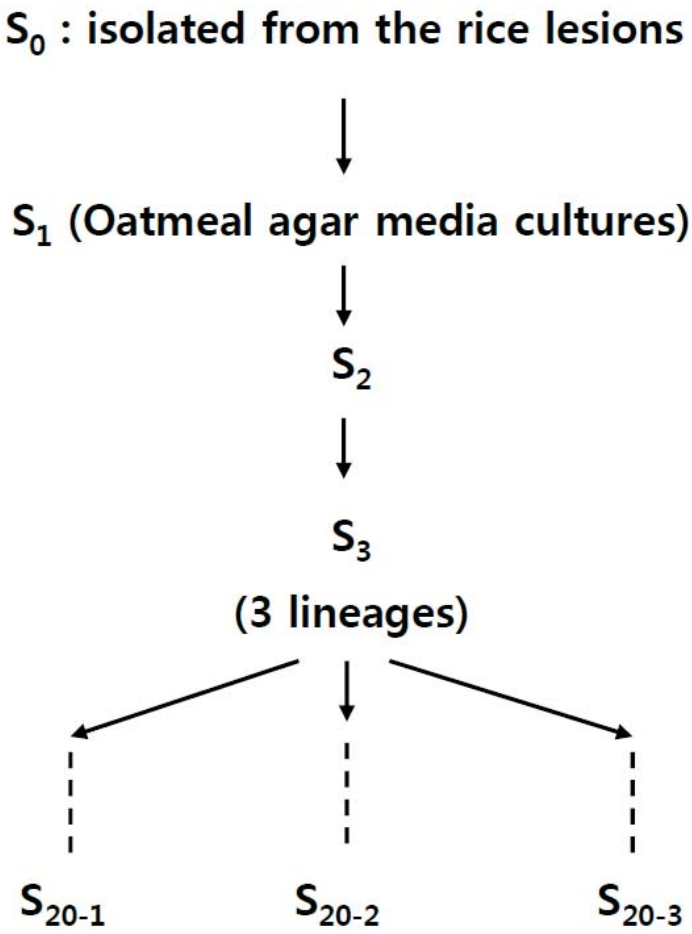


Figure 1. Schematic diagram of the modified MA experiment in *M. oryzae.* S_0 is the culture of the fungus isolated from a disease lesion on a rice leave. A patch of culture was transferred to the new oatmeal agar plates for the subculture (S_1). Subculture continued until it reaches S_{20}. From S_4, three independent cultures were maintained.

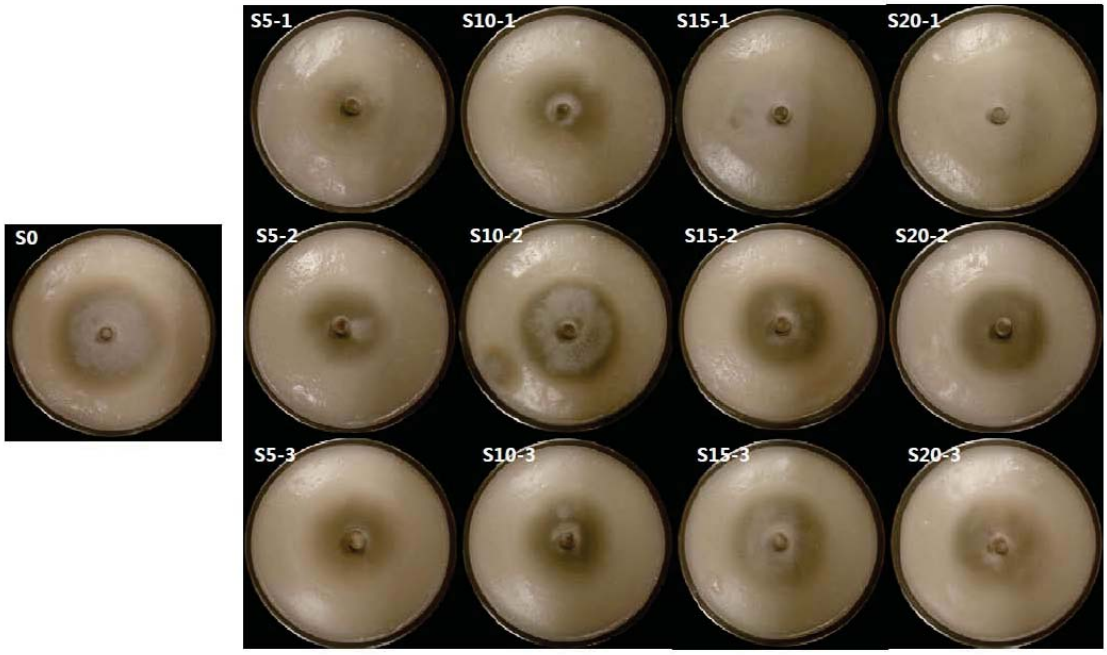


Figure 2. Colony morphologies of the derived strains on oatmeal agar plates. S_0 and every fifth derived strains are shown

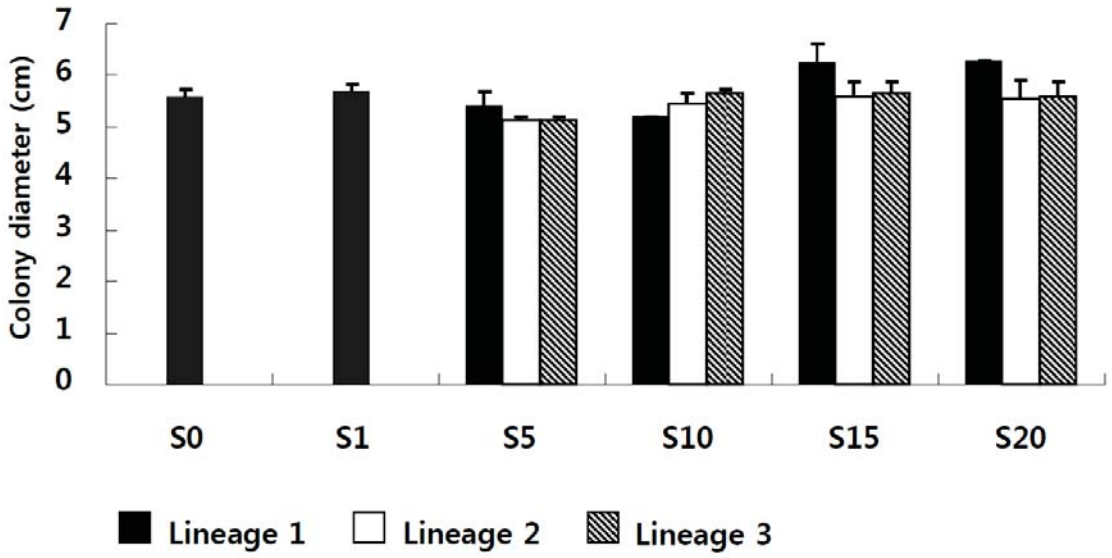


Figure 3. Growth rate of the derived strains on oatmeal agar plates. Growth rate was measured as the colony diameter. At least three replicates were used for each of three independent measurements.

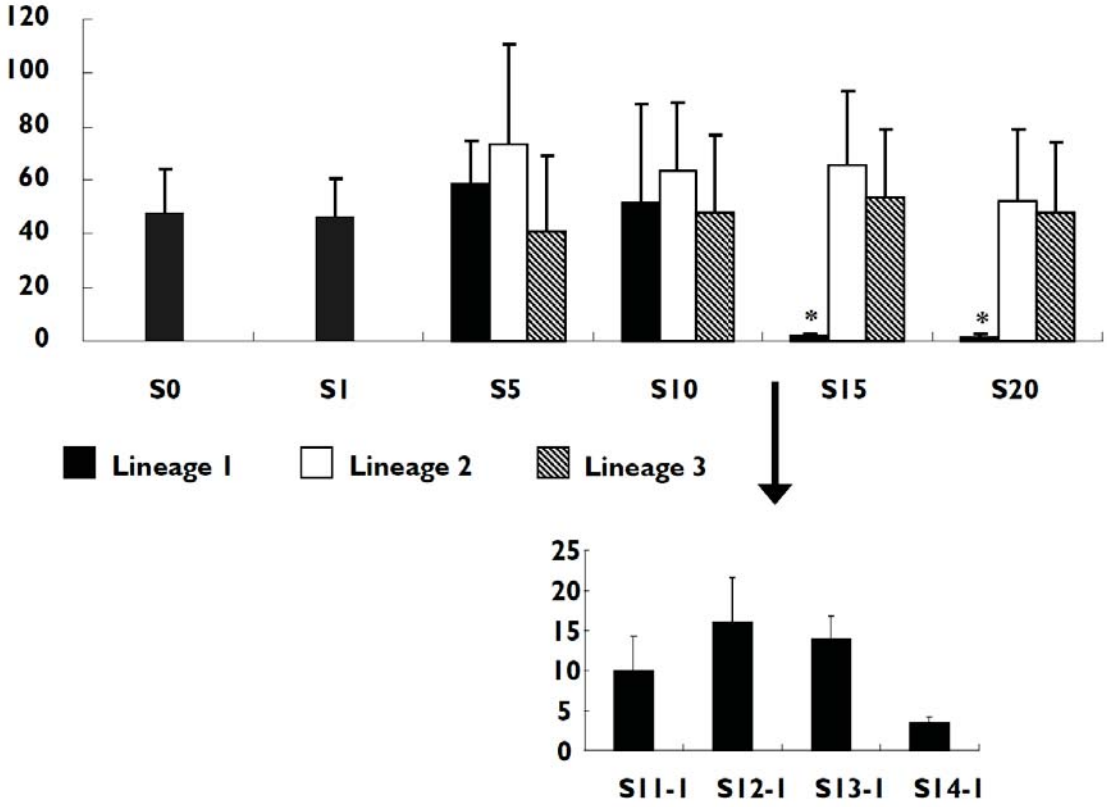


Figure 4. Asexual reproduction of the derived strains. Asexual spores were harvested from the plates and their numbers were counted using haemacytometer (y-axis represent the number of conidia per ml of conidia suspension). At least three replicates were used for each of three independent experiments. Asterisk indicates statistically significant differences, compared to S_0. Asexual spore production of S_{11} to S_{14} was shown below.

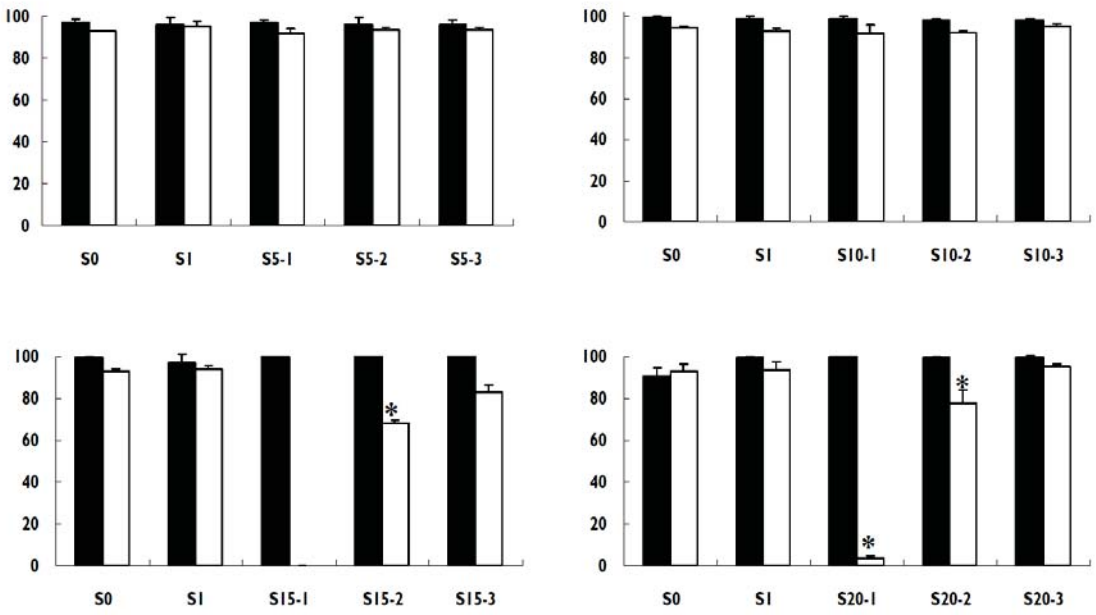


Figure 5. Germination (black bar) and appressorium formation (white bar) of the derived strain (both measured in percentage). At least three replicates were used for each of three independent experiments. Asterisk indicates statistically significant differences, compared to S_0.

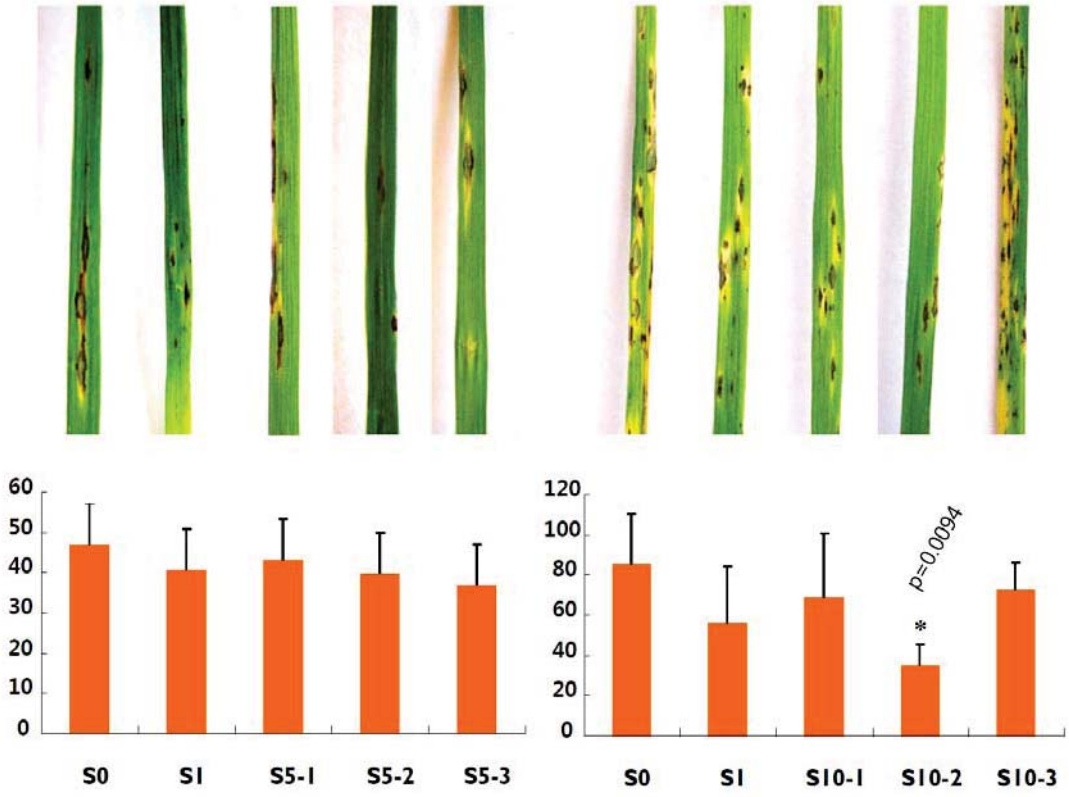


Figure 6. Virulence of the derived strains on a susceptible rice cultivar M2O2 (see Pathogenicity assay in Materials and Methods for virulence quantification). At least three diseased leaves were sampled for evaluation of virulence for each of three independent tests.

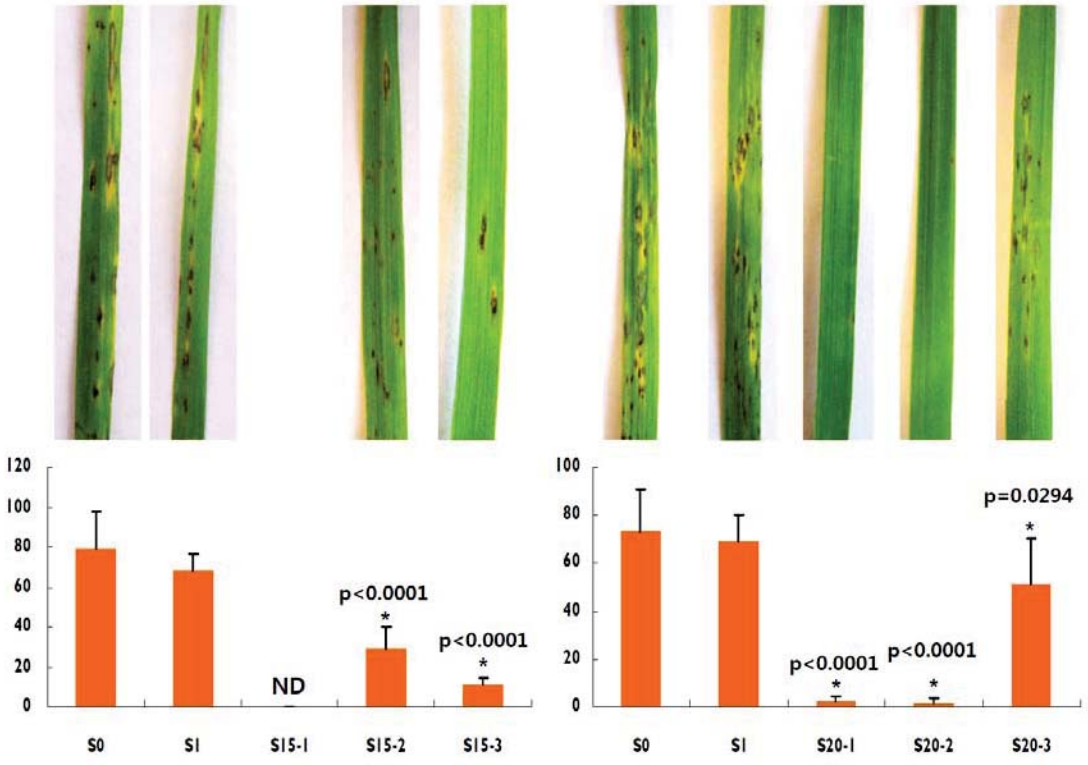


Figure 7. Virulence of the derived strains on a susceptible rice cultivar M2O2 (see Pathogenicity assay in Materials and Methods for virulence quantification). At least three diseased leaves were sampled for evaluation of virulence for each of three independent tests.

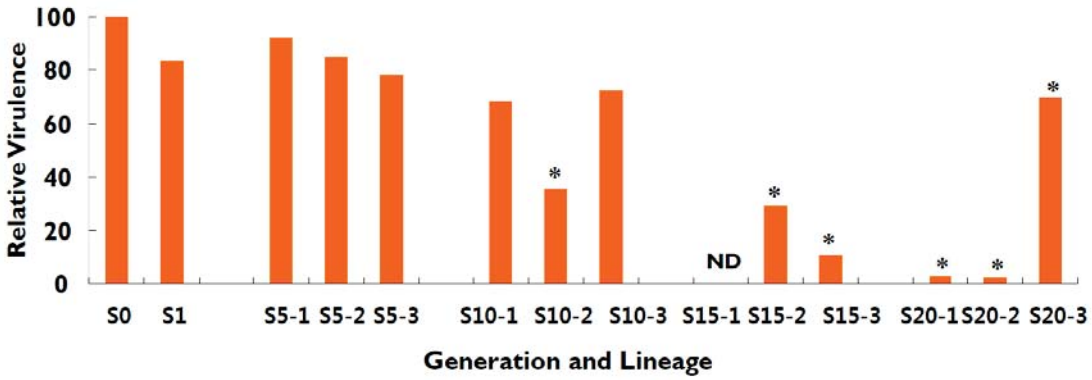


Figure 8. Relative virulence of the derived strains to that of S_0. Taking virulence of S_0 as 100, the relative virulence of the derived strains were calculated and plotted.

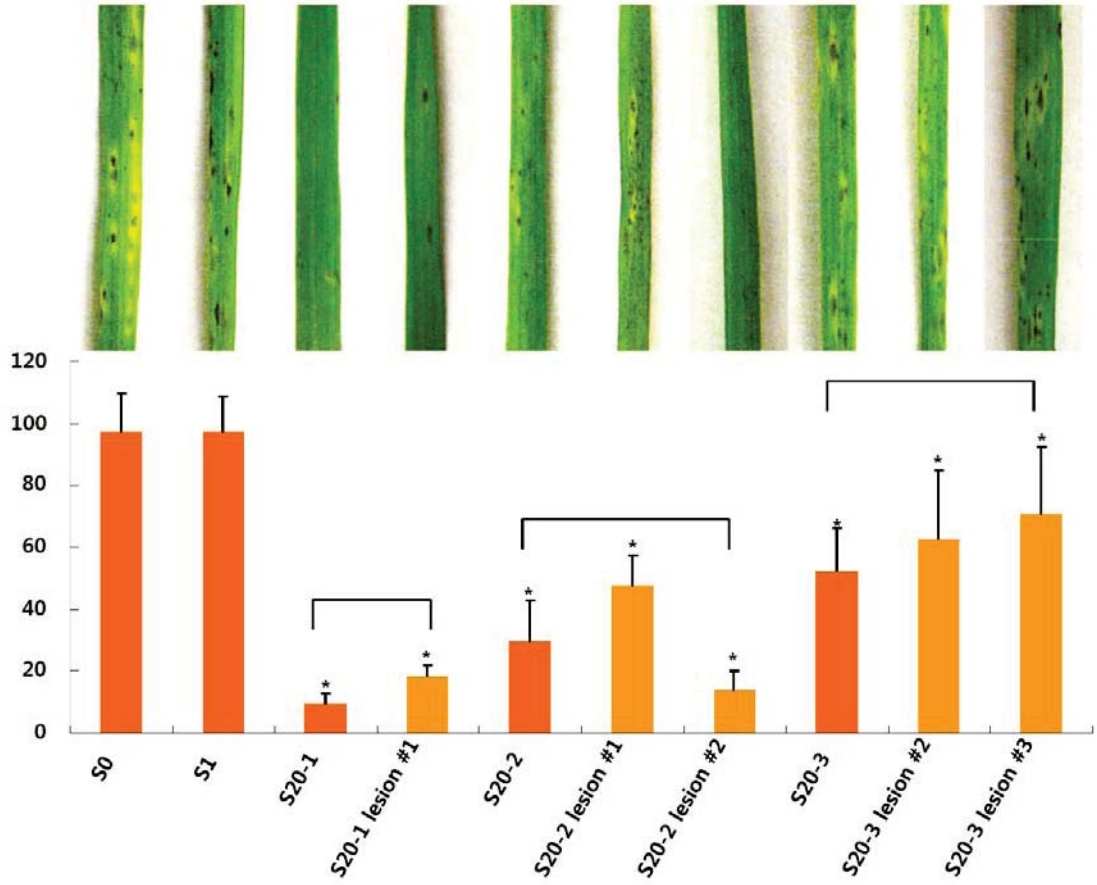


Figure 9. Virulence of the lesion isolates (yellow bars) from the derived strains (orange bars) (see Pathogenicity assay in Materials and Methods for virulence quantification). Asterisks indicate statistically significant difference between S_0 and the derived strains and their lesion isolates. There is no difference within a group of strains demarcated by vertical brackets.

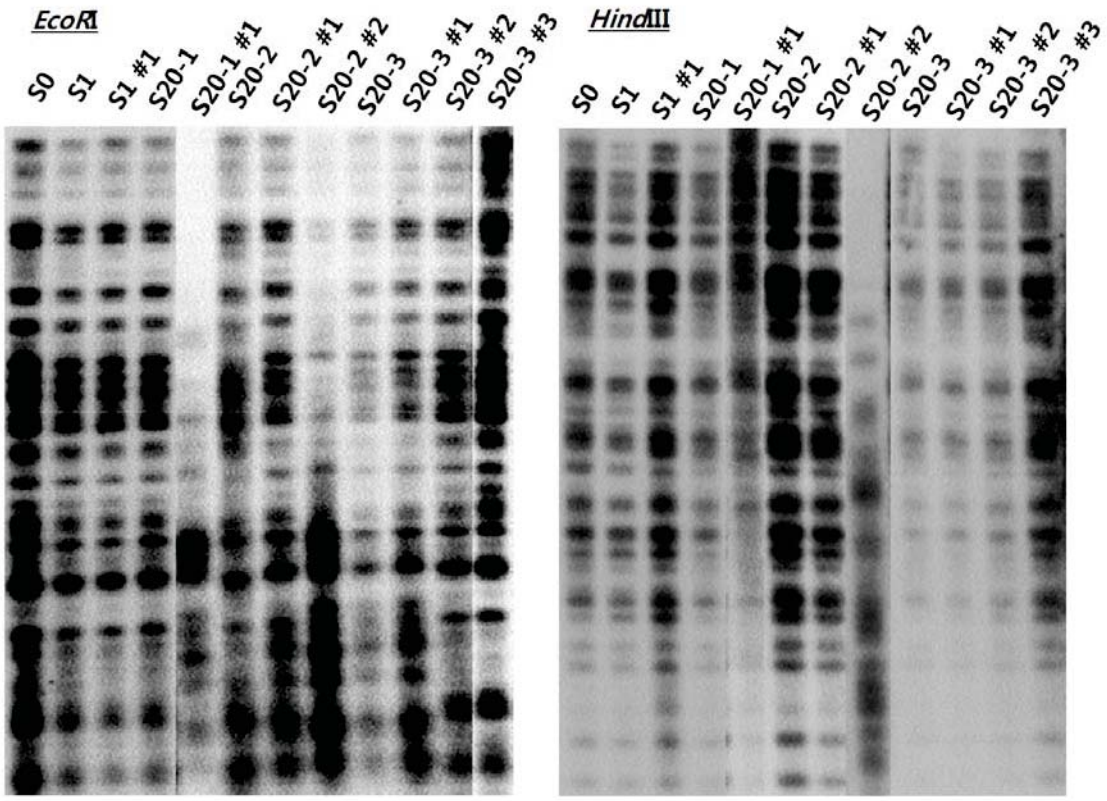
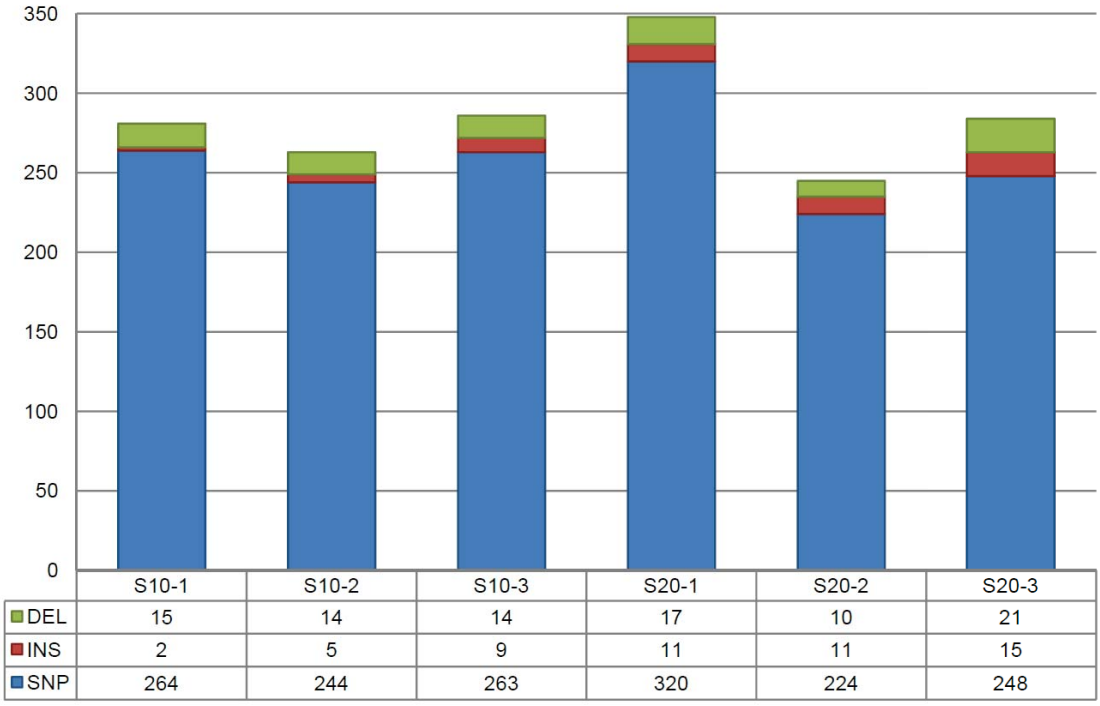
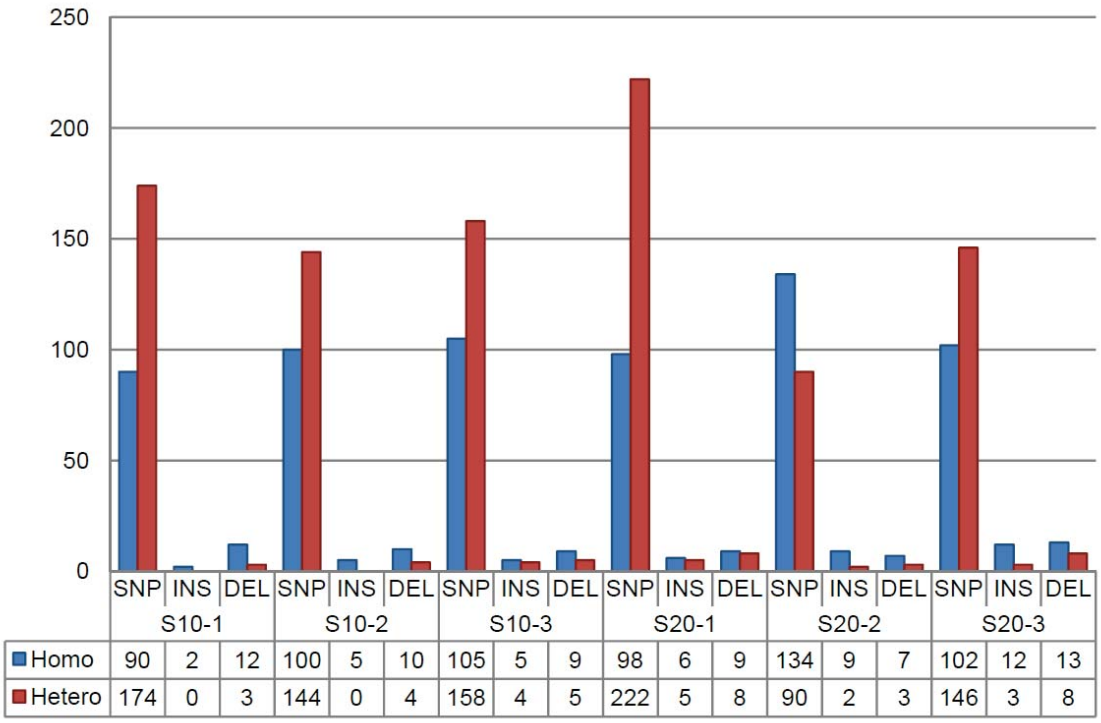


Figure 10. DNA fingerprinting of the derived strains and their lesion isolates. Two different enzymes were used with MGR586 as probe. $S_{20\text{-}1}$ #1 in *EcoR*I panel and $S_{20\text{-}2}$ #2 in *Hind*III panel show degradation of genomic DNA after digestion.



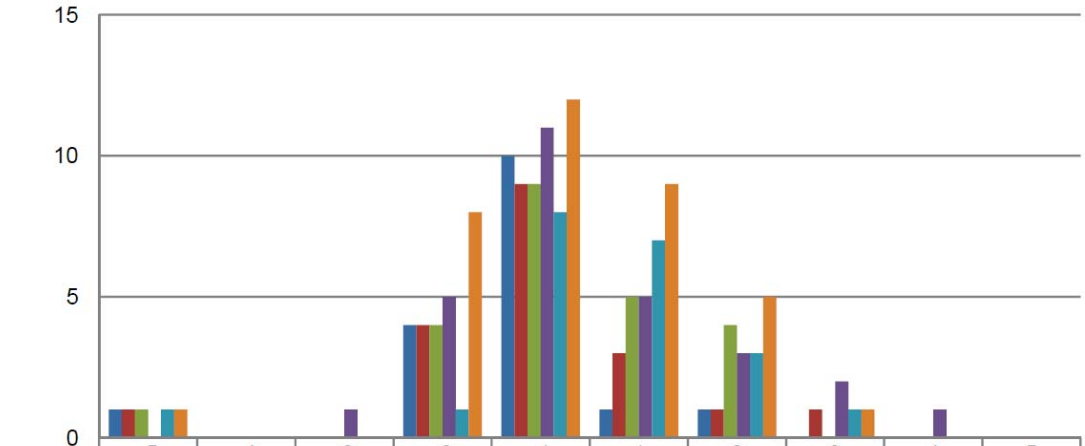
	S10-1	S10-2	S10-3	S20-1	S20-2	S20-3
DEL	15	14	14	17	10	21
INS	2	5	9	11	11	15
SNP	264	244	263	320	224	248

Figure 11. The number of mutations in the derived strains. Del, deletion; INS, insertion; SNP, single nucleotide polymorphism.



	SNP	INS	DEL	SNP	INS	DEL	SNP	INS	DEL	SNP	INS	DEL	SNP	INS	DEL	SNP	INS	DEL
	S10-1			S10-2			S10-3			S20-1			S20-2			S20-3		
Homo	90	2	12	100	5	10	105	5	9	98	6	9	134	9	7	102	12	13
Hetero	174	0	3	144	0	4	158	4	5	222	5	8	90	2	3	146	3	8

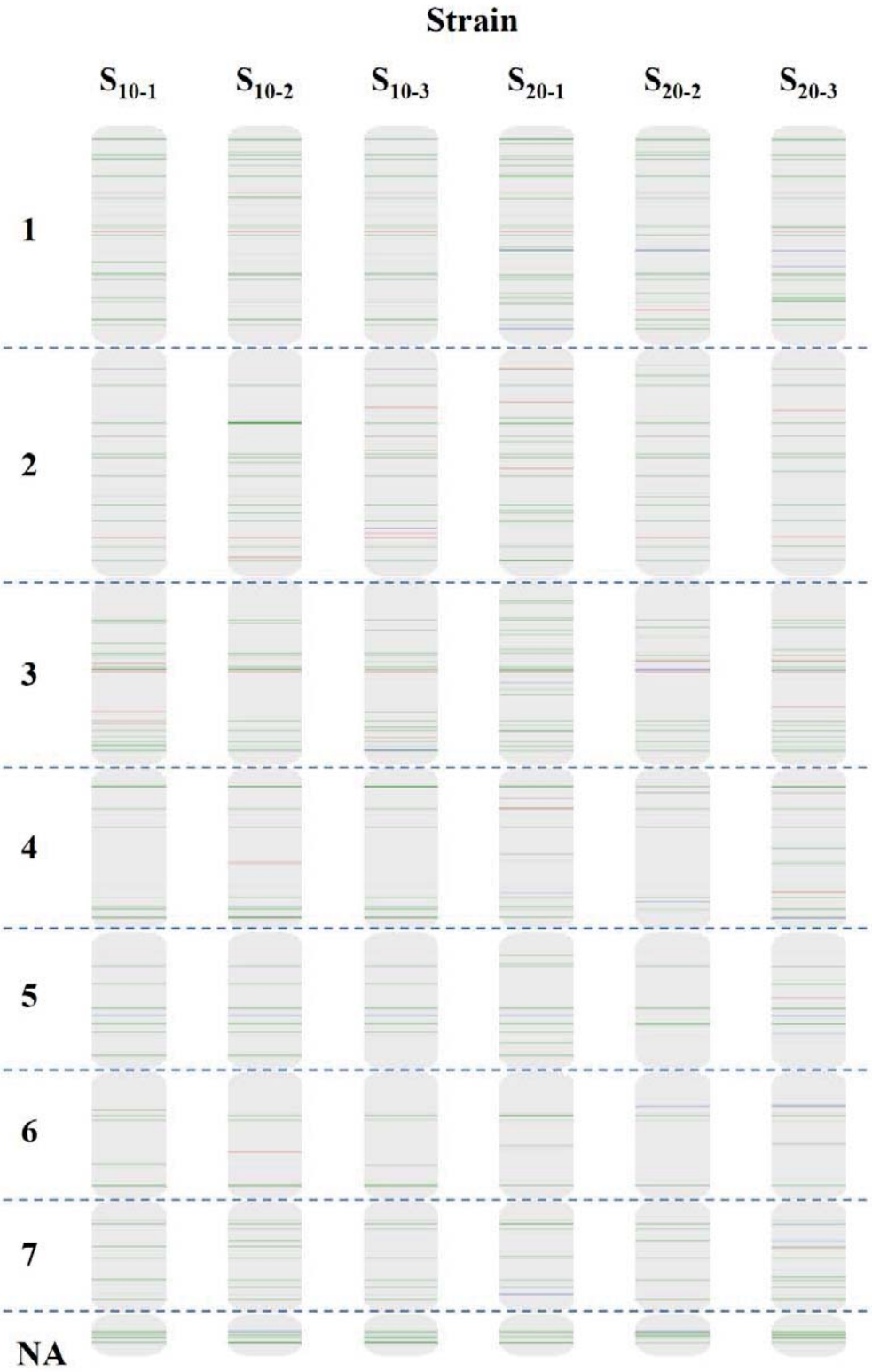
Figure 12. Composition of mutations in the derived strains. ‘Homo’ indicates the mutations with frequency of 1 (fixed mutation) and ‘Hetero’ indicates the mutations with frequency between 0.5 and 1.



	-5	-4	-3	-2	-1	1	2	3	4	5
S10-1	1	0	0	4	10	1	1	0	0	0
S10-2	1	0	0	4	9	3	1	1	0	0
S10-3	1	0	0	4	9	5	4	0	0	0
S20-1	0	0	1	5	11	5	3	2	1	0
S20-2	1	0	0	1	8	7	3	1	0	0
S20-3	1	0	0	8	12	9	5	1	0	0

Figure 13. Length distribution of insertions and deletions (InDel) in the derived strains. X-axis represent length of InDels (positive value for insertion and negative value for deletion). Y-axis represent the number of InDels having corresponding lengh in X-axis.

Figure 14. Distribution of mutations across chromosomes among the derived strains. Each vertical bar represents chromosomes within the derived strain. Horizontal lines within the bar represent the mutations found in the particular position of chromosomes. Green , blue, and red indicate SNP, insertion, and deletion, respectively.


Strain
S10-1
S10-2
S10-3
S20-1
S20-2
S20-3
1
2
3
4
5
6
7
NA

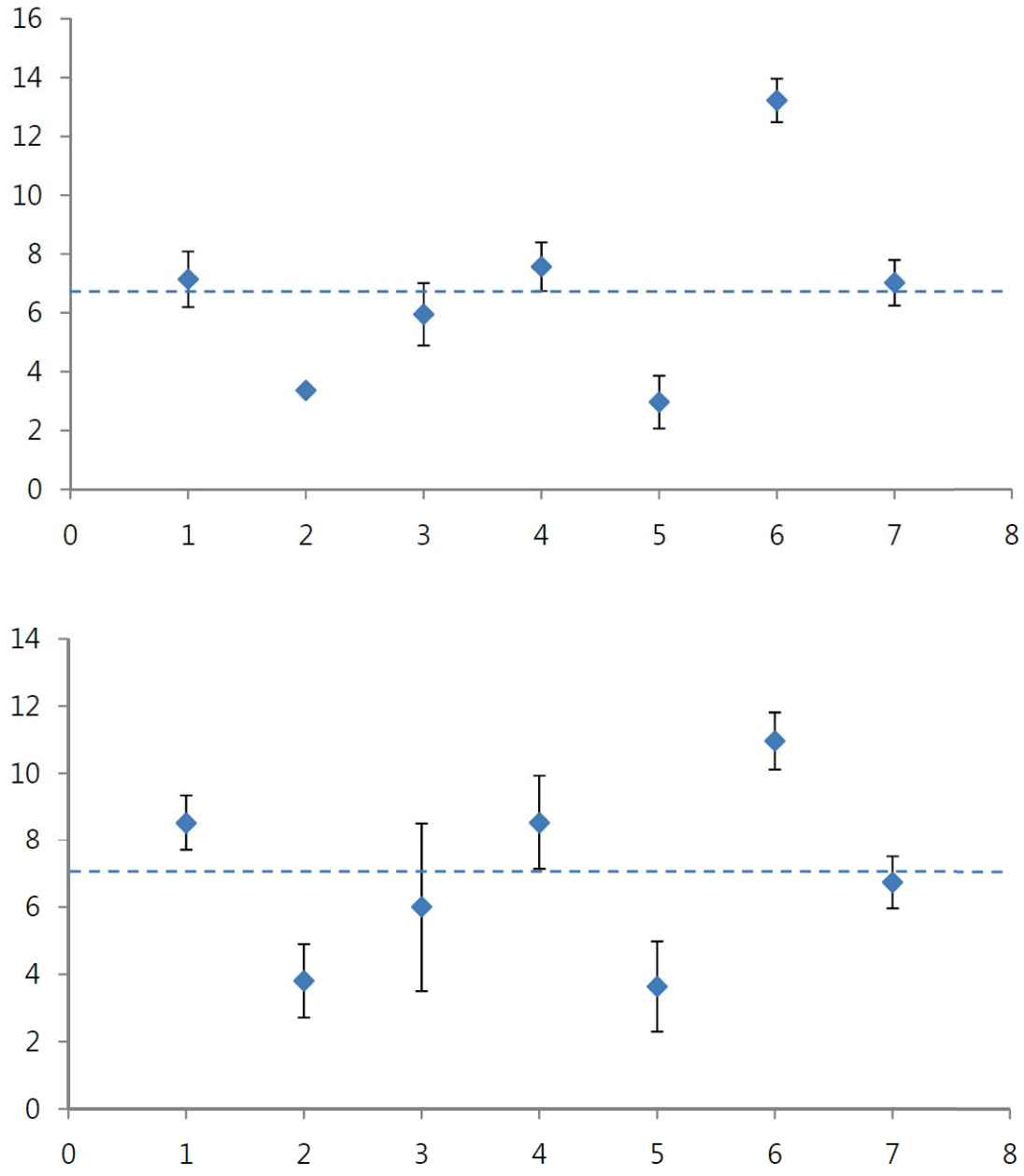


Figure 15. Mutational bias among chromosomes of the derived strains S_{10} (upper panel) and S_{20} (lower panel). X-axis indicates seven chromosomes and y-axis indicates the number of mutations per Mb. The dotted lines represents the expected number of mutation.

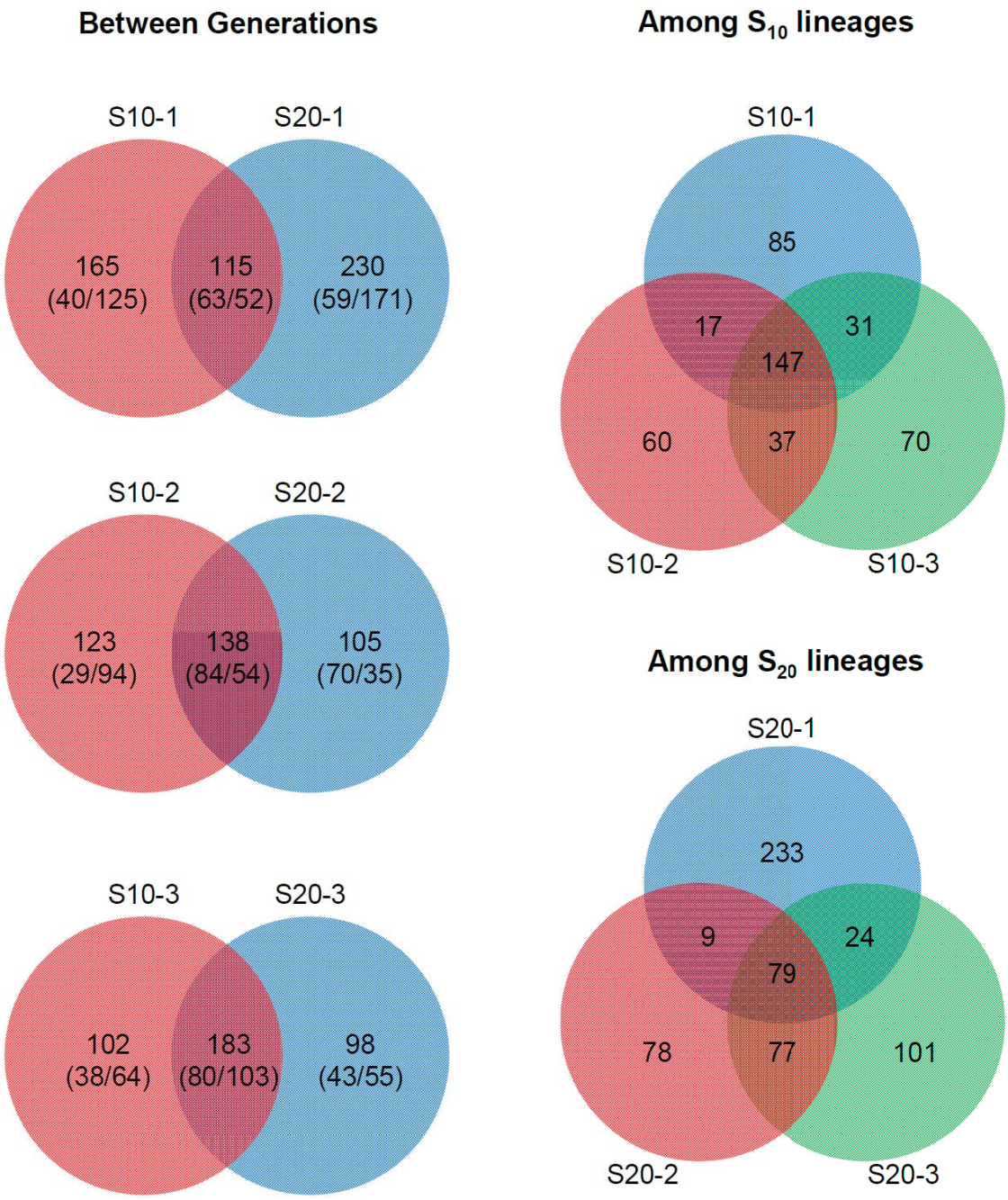


Figure 16. Mutations that overlap among lineages or between generation. (#1/#2): #1, number of fixed mutations; #2, number of unfixed mutations.

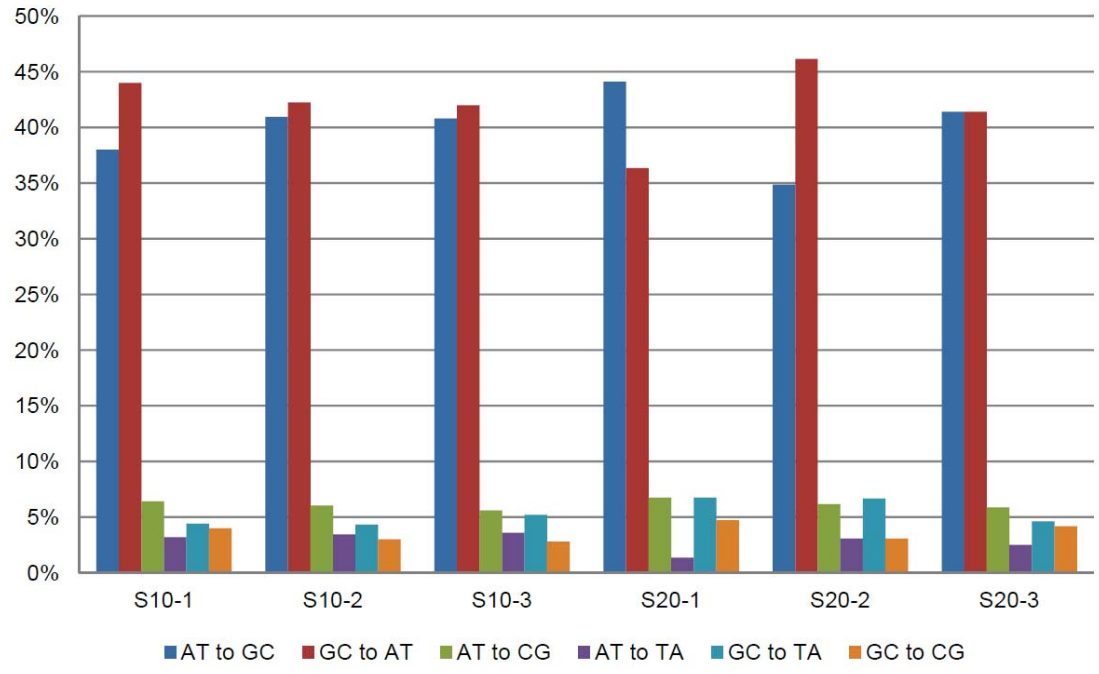


Figure 17. Percentage of types of nucleotide substitutions in the derived strains.

Chapter 3

Genome-wide analysis of DNA methylation in the rice blast fungus, *Magnaporthe oryzae*

Junhyun Jeon[1, 2], Jaeyoung Choi[4], Gir-Won Lee[4], Sook-Young Park[5], Aram Huh[1], Ralph A. Dean[2, 3], Yong-Hwan Lee[1, 4, 5, 6]

[1]Department of Agricultural Biotechnology, College of Agriculture and Life science, Seoul National University, Seoul 151-921, Korea

[2]Functional Genomics, North Carolina State University, Raleigh, NC 27607, USA

[3]Center for Integrated Fungal Research, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27607, USA

[4]Fungal Bioinformatics Laboratory, Seoul National University, Seoul 151-921, Korea

[5]Center for Fungal Pathogenesis, Seoul National University, Seoul 151-921, Korea

[6]Center for Fungal Genetic Resources, Seoul National University, Seoul 151-921, Korea

Correspondence: Yong-Hwan Lee

Phone) +82 2 880 4674

Fax) +82 2 873 2317

E-mail) yonglee@snu.ac.kr

SUMMARY

A key determinant of microbial pathogenesis is an ability of a pathogen to undergo morphological switching in response to environmental stimuli. This capacity to change form and structure is critical for survival, dispersal, and successful infection of hosts. Genetic pathways that regulate morphological transitions have been extensively studied in many microbial pathogens, yet the contribution of the epigenetic component is largely unknown. Here we use genetic manipulations and high-throughput bisulphite sequencing (methylC-seq) on the model plant pathogenic fungus, *Magnaporthe oryzae* to elucidate the dynamics and mechanics of DNA methylation during fungal development at single-nucleotide resolution. We show that two genes encoding DNA methyltransferases are responsible in a cooperative fashion for DNA methylation in this fungus and that progression of fungal development correlates with genome-wide reduction and reprogramming of the DNA methylome. Detailed analysis of methylC-seq data shows that reduction and reprogramming is commonly associated with upstream and downstream regions of annotated genes, suggesting a role for DNA methylation in regulating transcription. RNA-seq analysis of wild-type and DNA methyltransferase deletion mutants support that transcript abundance of genes, transposable elements, and unannotated intergenic transcripts are altered by DNA methylation. Our work provides new insights into the evolution of DNA methylation, revealing that DNA methylation in fungi is a dynamic epigenetic entity that may contribute to morphological switching driven by environmental cues.

INTRODUCTION

Genetic information is encoded and inherited primarily by four bases of DNA. An additional layer of heritable information is provided by chemical modifications of chromatin that consists of DNA and histones. These modifications bear information that can influence expression of underlying genes and ultimately cellular phenotypes without altering primary DNA sequences. The total sum of epigenetic information borne by chemical modifications called the epigenome. One such modification of the epigenome is DNA methylation that in eukaryotes, occurs exclusively at the 5′ position of cytosine pyrimidine ring (Suzuki and Bird, 2008). DNA methylation is the best characterized epigenetic modification in eukaryotes. In plants and animals, it has been shown that DNA methylation has important roles in myriad of cellular processes ranging from transposon silencing (Slotkin and Martienssen, 2007), genetic imprinting (Autran et al., 2005; Charalambous et al., 2007; Ferguson-Smith, 2011), and X-chromosome inactivation (Hellman and Chess, 2007) to embryo development and differentiation (Li et al., 1992; Okano et al., 1999; Illingworth et al., 2008; Laurent et al., 2010).

DNA methylation patterns are established and maintained by DNA methyltransferases (DNMTs) that catalyze the transfer of a methyl group to cytosine, leading to formation of 5-methylcytosine (5mC). In animals, much of what we know about DNA methylation has been learned from studies of mammals, in which DNA methylation occurs almost exclusively at symmetrical CpG dinucleotide sites. In mammalian systems, DNA methylation patterns are established by DNMT3 family of *de-novo* methyltransferases and maintained by DNMT1

family of maintenance methyltransferases. Unlike animals having methylation predominantly in a CG context, plants possess distinct DNA methylation machineries that operate in different sequence contexts. In plants, d*e-novo* methylation is carried out by DRM2, a homologue of the DNMT3 methyltransferases, and maintenance methylation is accomplished by three different DNA methyltransferases: MET1, the plant homologue of DNMT1 for CG context; CMT3, the plant-specific DNA methyltransferase for CHG context (where H = A, C, T); DRM2, through persistent de-novo methylation for CHH context.

To understand the biological significance of DNA methylation, it is important to map the genomic distribution of methylated sequences, so-called methylome. The gold standard technology for detecting 5mC is bisulphite sequencing (Frommer et al., 1992). This method uses a chemical reaction that converts unmethylated cytosine to uracil, leaving methylated cytosines unaffected. Chemical conversion is followed by PCR, cloning and sequencing to identify 5mC that remain as cytosine. Due to the need to use of locus-specific primers and multiple cloning steps, bisulphite sequencing is not well-suited for DNA methylation profiling at the genomic scale. In the past, approaches to mapping the methylome therefore depended on genomic microarrays combined with either the use of sensitivity of enzymes to CpG methylation within their recognition sites or the use of protein affinity to enrich methylated sequences as probes (Weber et al., 2005; Schumacher et al., 2006; Zilberman et al., 2007). These approaches are quite useful but have two important constraints. First, the use of microarray poses a serious technical challenge on application when dealing with organisms having large genome size. Second, resolution of readouts from these approaches is relatively low. Recently it has been demonstrated that the bisulphite sequencing method can

be successfully combined with next-generation sequencing technology (methylC-seq) to produce a genome-wide DNA methylation profile at the single base-pair resolution (Cokus et al., 2008; Lister et al., 2008; Laurent et al., 2010; Zemach et al., 2010).

Such high-throughput approaches have led to new observations that challenge the conventional view of DNA methylation. In the past, DNA methylation was thought to be predominantly present in the promoter regions of genes and transposons and play roles in irreversibly silencing transcription. Now, it is clear that DNA methylation is found in gene bodies of many actively transcribed genes and that in some species, transposons are not the preferential target for methylation (Zemach et al., 2010). In addition, studies for *Arabidopsis thaliana* revealed that DNA methylation is not static but dynamic from generation to generation, providing a possible mechanism for phenotypic diversity in the absence of genetic mutation.

In fungi, it has long been thought that DNA methylation mainly plays a role in genome defense via silencing of repeat sequences and transposons. Recent genome-wide analysis of five fungal species corroborated that average methylation level of the fungal genome is very low and that transposons and repeat sequences are indeed heavily methylated, compared to gene bodies (Zemach et al., 2010). Considering that none of the five fungal species are pathogenic and fungi inhabit extremely diverse environmental niches, a complete picture of DNA methylation in the fungal kindom, however, is yet to emerge. In fact, a recent study on a fungal pathogen of human, *Candida albicans* showed that repeat sequences are largely free of methylation and that differential expression of many structural genes during dimorphic switching is associated with DNA methylation (Mishra et al., 2011).

Magnaporthe oryzae is a causal agent of rice blast disease that is one of the most devastating diseases in rice plants, and is a model plant pathogenic fungus with a completely sequenced ~41Mb nuclear genome (Dean et al., 2005). The fungus usually disseminates via asexual spores, called conidia that on leaf surface, produce germ tubes in the presence of water. The germ tube tip develops into a specialized infection structure, appressorium, in response to environmental cues such as surface hydrophobicity (Lee and Dean, 1993). Using turgor pressure within the appressorium, the fungus mechanically penetrates and grows inside the plant host, leading to formation of disease lesions where the fungus produces conidia for secondary infections. This infection debilitates host plants, causing yield losses or even death. In *M. oryzae*, the retro-transposon, MAGGY is the only DNA sequence that has been experimentally shown to be methylated to date. Methylation of MAGGY is involved in repressing transcription but not transposition (Nakayashiki et al., 2001). Despite the presence of DNA methylation in *M. oryzae*, the impact of DNA methylation on fungal biology has not been addressed so far. Here, I report the first genetic and genome-wide analysis of DNA methylation in a plant pathogenic fungus, *M. oryzae* and discuss the implication of DNA methylation in the biology and evolution of this fungus.

MATERIALS AND METHODS

Strains and culture conditions

M. oryzae strain KJ201 was used as wild-type throughout this work. Wild-type strain and all the mutants generated in this study were maintained on either V8 (4% V8 juice) or oatmeal (50g of oatmeal per liter) agar plates at 25℃ under continuous fluorescent light to promote conidiation. Selection of hygromycin- or geneticin-resistant mutants were performed using TB3 media supplemented with 200µl ml^{-1} hygromycin B or 800 µl ml^{-1} geneticin. For genomic DNA and total RNA extraction from mycelia, the fungus was cultured in complete medium (0.6% yeast extract, 0.6% casamino acid, 1% glucose) for three days. Appressorium induction for genomic DNA extraction was done in bulk on the hydrophobic side of a green-mirror using conidia harvested from an oatmeal agar plate culture.

DNA isolation and Southern hybridization

Genomic DNA was isolated using two methods depending on follow-up experiments. Genomic DNA for Southern hybridization and bisulphite sequencing purposes was prepared by standard method (Sambrook and Russel, 2001). Genomic DNA from mutants for PCR–based screening was isolated by 'quick and safe method' (Chi, 2009) as follows: a patch of mycelial mat on TB3 media was taken using sterilized tooth-picks and ground in 1.5ml e-tube with 500µl of extraction buffer (100mM Tris-HCl, 10mM EDTA and 1M KCl). Cell debris was removed by centrifugation at 5,000 rpm for 15min and genomic DNA was

pelleted by isopropyl alcohol. The pellet was washed with 70% alcohol and dissolved in strerilzed-distilled water.

Restriction enzyme digestion, agrose gel separation, cloning and DNA gel blotting were performed following the standard protocols (Sambrook and Russel, 2001). Hybridization was carried out in solution containing 6×SSC, 5×Denhardt's solution, 0.5% SDS and 100µl ml^{-1} denatured salmon sperm DNA, at 65ºC. Blots were exposed to phosphoimager analyzer (BAS-2040, Fuji Photo Film Co., Ltd., Tokyo, Japan) and visualized by phosphoimager analyzer software.

Phylogenetic analysis of DNA methyltransferases

Phylogenetic relationships among DNA methyltransferases were inferred using neighbor-Joining method. (Saitou and Nei, 1987). The bootstrap consensus tree inferred from 250 replicates is taken to represent the evolutionary history of the taxa analyzed. The evolutionary distances were computed using the JTT matrix-based method (Jones et al., 1992) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). Phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007).

Targeted disruption of genes encoding DNA methyltransferases and their complementation

Targeted disruption of a gene was achieved by transformation of fungal protoplasts with knockout construct containing flanking sequences (~1kb in length for both 5′ and 3′ flanking)

of a target gene and selectable marker. Knockout construct was prepared using modified double-joint PCR (Yu et al., 2004). PCR products were directly used for transformation of fungal protoplast. Double knockout mutants were obtained by transformation of *Δdmt1* protoplast with knockout construct for *MoDMT2*. Resulting transformants were selected on TB3 media containing hygromycin (200ppm) or geneticin (800ppm) and screened by PCR for correct replacement event. The mutants were further confirmed by Southern blot analysis. Primers used in preparation of the knockout construct were listed in table 1.

Real-time PCR analysis of DNA methyltransferase expression

Total RNAs were extracted from powder into which cells and tissues from different developmental stages or cultures grown under various stress conditions were ground, using Easy-spin™ total RNA extraction kit (iNtRON Biotechnology, Seoul, Korea). For each sample, 2μg of total RNA was reverse transcribed into first strand cDNA synthesis with ImProm-II^{TM} Reverse Transcription System (Promega, Madison, WI). Real-time PCR were performed in a 20μl volume containing 2 μl of cDNA (25 ng of input RNA), 100 nM of each primer and 12.5 μl of 2× Power SYBR® Green PCR Master Mix (Applied Biosystems, Warrington, UK). Reaction was run on the Applied Biosystems 7500 Real Time PCR System (Applied Biosystems, Foster City, CA) for 40 cycles of 15s at 95℃, 30s at 60℃ and 30s at 72℃. Resulting values of mean threshold cycles (Ct) was normalized as previously described (Kim et al., 2005). Primers used in real-time PCR analysis were listed in table 1.

Vegetative growth, conidiation, and appressorium formation

Vegetative growth rate was measured as colony diameter on complete media, minimal media, oatmeal agar and V8 juice agar plates with at least three replicates. Ability to produce conidia was measured by counting the number of conidia from 8-days-old oatmeal agar cultures in 6-well plates. Conidia were collected by flooding the plate with 5ml of sterilized distilled water and number of conidia was counted using a hemacytometer under a microscope. Conidial germination and appressorium formation were measured on plastic coverslips. Conidia were harvested from 8 to 10-day-old cultures on oatmeal agar plate with sterilized distilled water. Conidial suspension of 40µl was pipetted onto plastic coverslips following adjustment of its concentration to ~ 2×10^4 spores/ml. Drops were placed in a moistened box and incubated at 25ºC. The percentage of conidia that germinated and the percentage that formed appressoria were determined by microscopic examination of at least 100 conidia per replicate in at least two independent experiments with three replicates.

Pathogenicity assay

Conidia were harvested from 8 to 10-days old culture grown on oatmeal agar plates with sterilized distilled water. Conidial suspensions with the concentration adjusted to $5x10^4$spores/ml were spray-inoculated on susceptible rice seedlings (cv. Nagdong) at the 3 to 4 leaf stage. Disease lesions were assessed 6 days post inoculation.

Bisulphite sequencing library construction and high-throughput sequencing

Five microgram of genomic DNA was fragmented by sonication to 100-300bp with Bioruptor (Diagenode Sparta, NJ). The fragmented DNA was end-repaired and ligated to methylated sequencing adapters provided by Illumina according to manufacturer's instructions (Illumina, San Diego, CA). Bisulphite conversion was carried out with ZYMO EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA) as described in manufacturer's instructions. Following desalting, size selection, PCR amplification, and a second size selection step, the library was sequenced using the Illumina Genome Analyzer (GA) according to manufacturer's instructions. Raw GA sequencing data were processed by Illumina base-calling pipeline.

Initial processing and mapping of bisulphite sequencing reads

Prior to mapping, short reads from Illumina sequencing were subjected to filtering: i) adapter sequences were trimmed out of reads; ii) reads with N bases over 10% of reads were filtered out; iii) if there are reads with over 10% of bases of read whose qualities are less than 20, these bases were trimmed off from reads.

Due to the nature of bisulphite sequencing, both reads and references genome should be transformed for alignment and mapping as follows: observed cytosines on forward read were *in-silico* replaced by thymines and observed guanines on reverse reads were *in-silico* replaced by adenosines. *M. oryzae* reference genome sequencesthe were downloaded from Magnaporthe comparative database (*Magnaporthe comparative* Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/)) and computationally

converted to alignment target sequences: every cytosine in the plus strand was converted to thymine and every guanine in the minus strand was converted to adenosine. Alignment of reads to reference genomes was carried out using SOAP aligner, allowing two mismatches (Li et al., 2009). Hits with a single placement and a clear strand assignment were defined as unambiguous alignment (that is, uniquely mapped reads) and used to compute the coverage and methylation levels of the local region.

Bisulphite-PCR validation for target regions

At least two microgram of genomic DNA was fragmented by sonication to 100bp-500b with Bioruptor (Diagenode Sparta, NJ) and bisulphite-converted manually as previously described (Clark et al., 2006). Primers were designed to amplify target regions of the bisulphite-converted DNA for validation of high-throughput bisulphite-sequencing results. Amplified products were cloned into pGEM-T Easy vector (Promega, Madison, WI) and at least ten independent clones were Sanger-sequenced for each target region.

RNA-seq

RNA-seq was carried out with the Illumina platform according to manufacturer's instructions for total RNAs extracted from mycelia tissues of wild-type and DMTase mutants. Resulting paired-end sequencing reads were aligned and quantified using TopHat(Trapnell et al., 2009) and Cufflinks(Trapnell et al., 2012) with default parameter values. Sequencing reads were first mapped to annotated transcripts in *M. oryzae* genome using TopHat (Trapnell et al., 2009) and fed to Cufflinks (Trapnell et al., 2012) to estimate transcript

abundance and test for differential expression. Gene expression was calculated as FPKM (fragments per kilobase of transcript per million mapped fragments).

RESULTS

Preliminary analysis of DNA methylation using 5-azacytidine

5-azacytidine (5-aza) is a nucleoside analogue of cytidine and has been shown to inhibit DNA methyltransferase activities by incorporating into DNA (Jones et al., 1983). There have been several studies where 5-aza was used to infer the importance of DNA methylation in fungal development and toxin production (Tamame et al., 1983a; Kritsky et al., 2001; Wilkinson et al., 2011). As an attempt to probe the global impact of DNA methylation on *M. oryzae*, we used 5-azacytidine to induce DNA demethylation and observe phenotypic changes. To distinguish the effect of demethylation from that of concentration, cytidine was included as a control. The effectiveness of 5-aza in inducing demethylation was assayed via Southern blot analysis of MAGGY methylation using two pairs of methylation insensitive and sensitive isoschizomers (*MspI*-*Hpa*II and *Dpn*I-*Sau3A*I), each of which recognize CCGG and GATC, respectively (Figure 1). Southern blot analysis showed that 5-aza is very effective in demethylation as indicated by disappearance of band patterns in samples treated with 5-aza (arrows). This also shows that in *M. oryzae*, there seems to be no adenosine methylation and that cytosine methylation may occur in non-CG context (*Dpn*I-*Sau3A*I) as well.

Due to cytotoxicity and mutagenic activity of 5-aza, a range of concentrations were tested to find the optimal concentration, at which growth was not inhibited. In minimal media, I found that 10μM or higher concentration of 5-aza significantly inhibited growth of the fungus, while the equal concentration of cytidine did not (Figure 2). Thus, I determined to

use 1 and 5μM in follow-on experiments using solid media. On complete media, minimal media, and V8 media containing 1 or 5 μM of 5-aza, the fungus showed aberrant morphologies (Figure 3) and these were reproducible when the fungus was grown after being transferred onto media without 5-aza, indicating that morphological changes observed were heritable. However, these morphological changes were not common among 6 the replicates in the experiment. This fact probably suggests that these changes might be attributed to DNA mutations independently induced in some replicates by 5-aza during culture. From these data, I concluded that 5-aza is an effective DNA demethylating agent but is cytotoxic and mutagenic for *M. oryzae*, limiting its use in further studies for DNA methylation in this fungus. To further explore DNA methylation in *M. oryzae*, I turned to gene(s) encoding DNA methyltransferase.

Identification and phylogenetic analysis of DNA methyltransferase genes in *M. oryzae*

In *Neurospora crassa*, DNA methylation has been intensively studied, revealing genetic components required for DNA methylation. *DIM-2*, a DNA methyltransferase is one such component and is responsible for most cytosine methylation in *N. crassa* (Kouzminova and Selker, 2001). Although *N. crassa* possesses another protein, RID having a DNA methyltransferase domain, this protein is known to have no DNA methylation activity and is involved primarily in the repeat-induced point mutation (RIP) process (Freitag et al., 2002). Using amino acid sequences of DIM-2 and RID as queries, I carried out BLAST search against the proteome of *M. oryzae* and found two genes (MGG_00889 and MGG_02795) having homology to DIM-2 and RID, respectively. These genes encoding putative DNA

methyltransferases were designated as *MoDMT1* and *MoDMT2*. When amino acid sequences of MoDMT1 and MoDMT2 were aligned with known DNA methyltransferases, major motifs such as catalytic PC doublet (motif IV) in C-terminal DNA methyltransferase domain were well conserved, suggesting that both MoDMT1 and MoDMT2 may have *in-vivo* DNA methyltransferase activity (Figure 4).

Eukaryotic DNA methyltransferases are grouped into 6 subfamilies, based on sequence similarity: DNMT1/MET1 (DNMT1), CMT (DNMT1), DIM-2/MASC2 (DNMT1), DNMT2/PMT1 (DNMT2), DNMT3/DRM (DNMT3), and RID/MASC1 (DNMT4) (Ponger and Li, 2005). Phylogenetic analysis showed that MoDMT1 and MoDMT2 belong to DIM-2/MASC2 subfamily presumably having maintenance DNA methylation activity and RID/MASC2 subfamily, respectively (Figure 5). This indicates that *M. oryzae* has a potential maintenance DNA methyltransferase and there is no specialized *de-novo* DNA methyltransferase in this fungus. In many eukaryotic DNA methyltransferases, catalytic domain is associated with variable length of N-terminal extensions harboring conserved domains, which are involved in functional specializations (Chen and Li, 2004). In line with this, domain architecture analysis with InterPro scan showed that MoDMT1 shares N-terminal extension containing BAH (bromo-adjacent homology) domain (IPR001025) with DIM-2 long but does not have a PWWP (proline-tryptophan-tryptophan-proline) domain (IPR000313) commonly found in DNMT3 (Figure 6).

Distribution of DNA methyltransferases across fungal phylums

Despite the enormous diversity of life styles and abundance of species within the fungal kingdom, DNA methylation machineries have been identified and studied only in a small number of model systems such as *N. crassa*. Recently, efforts were made to address this gap and reconstruct phylogeny of fungal methylomes using methylC-seq (Zemach et al., 2010). However, the fungal species used in this work included only five species from three phylums (ascomycota, basidiomycota, and zygomycota) and none of them were pathogens.

To consider DNA methyltransferases of fungi in a broader perspective, distribution of six subfamilies of DNA methyltransferases were mapped onto a cladogram of 25 fungal species and 8 species from other kingdoms including plants and animals (Figure 7). These 33 species were chosen to represent each phylum, based on availability of genome sequence information and with a priority on pathogens. These mapping makes possible three generalizations about distribution of DNA methyltransferases across the fungal kingdom: 1. Distribution of each subfamily is highly patchy even within the phylum. 2. Fungi in pezizomycotina, however, tend to have two DNA methyltransferases: DIM-2 type and RID type (DNMT4). 3. Presence of DNA methyltransferase does not guarantee genome methylation and vice versa. For example, *S. pombe* has DNA methyltransferase that belongs to DNMT2 subfamily but its genome is free of DNA methylation, while the genome of *C. albicans* is methylated even though it has no orthologs for any known DNA methyltransferase to date. The patchy distribution may be the reflection of diverse life styles and complex evolutionary history of fungal kingdom. However, the possibility that unknown evolutionary factor has driven loss or retention of DNA methylation machinery in the fungal kingdom cannot be ruled out as

demonstrated in the case of compatibility of RNAi-deficient fungi with killer (Drinnenberg et al., 2011).

Expression profile of *MoDMT1* and *MoDMT2*

Transcriptional expression of *MoDMT1* and *MoDMT2* during vegetative growth and fungal development was analyzed by RT-PCR (Figure 8). Expression level of genes was normalized using *β-tubulin* as a reference gene and compared to expression level in conidia. In this analysis, expression of *MoDMT1* and *MoDMT2* showed a high degree of correlation across all the conditions used for analysis (correlation coefficient = 0.988), suggesting that transcription of the two genes may be under the same or overlapping regulatory mechanisms. Compared to their expression in conidia, expression of the two genes was down-regulated during germination and appressorium formation, while it was up-regulated during invasive growth within host plants and vegetative growth on artificial media. This pattern of gene expression suggests that DNA methylation may vary, depending on which developmental program the fungus is undergoing.

Targeted deletion of *MoDMT1* and *MoDMT2*

To genetically examine the activity of MoDMT1 and MoDMT2 as DNA methyltransferase and to characterize the roles of the two genes, targeted gene deletion was carried out. A knockout construct was generated by modified double-joint PCR (Yu et al., 2004), through which approximately 1kb of flanking sequences of target genes were amplified and fused to hygromycin cassette by PCR. The resulting PCR product was directly used to transform

protoplasts of wild-type strain KJ201 (Figure 9a and c). The hygromycin-resistant colonies produced by protoplast transformation were screened for the integration of knockout construct by PCR. Correct replacement of target gene by homologous recombination was confirmed in seven for *MoDMT1* and two mutants for *MoDMT2* by Southern blot analysis (Figure 9b and d). Gene deletion was further confirmed by RT-PCR, wherein no transcripts of *MoDMT1* and *MoDMT2* were detected in the mutants (data not shown).

Since mating is not feasible in wild-type strain KJ201, protoplasts of *Δdmt1* (hygromycin resistance) were transformed with knockout construct for *MoDMT2* (geneticin resistance) to generate double deletion mutants. Mutants resistant to both hygromycin and geneticin were selected following protoplast transformation and deletion of both genes was confirmed by PCR and Southern blot analysis from two mutants (Figure 10).

MoDMT1 is a major DNA methyltransferase

To test if *MoDMT1* and *MoDMT2* are required for DNA methylation, we first took advantage of Southern blot analysis of DNA methylation in a retrotransposon, MAGGY. Using a pair of isoschizomers (*Msp*I and *Hpa*II), methylation of CG dinucleotide in CCGG context within MAGGY (~5.6kb) was examined with genomic DNAs isolated from mycelia (Figure 11a) Methylated MAGGY sequences produce no typical hybridization band when cut with methylation-insensitive restriction enzyme *Msp*I unlike when cut with methylation-sensitive *Hpa*II. In the wild-type strain KJ201, MAGGY is methylated as expected but in the *Δdmt1*, the expected band pattern disappeared, indicating that there is no or little methylation in MAGGY within *Δdmt1* genome. In contrast to *Δdmt1*, considerable, but lower than wild-

type, level of methylation was detected in *Δdmt2*. Two independent experiments using approximately equal amount of DNA (~4μg) consistently showed weaker band pattern in *Δdmt2*, indicating that difference in band intensity cannot be attributed to difference in amount of DNA used. The *Δdmt1Δdmt2* showed the same result as *Δdmt1*. When the *MoDMT1* gene under native promoter was introduced into *Δdmt1*, MAGGY methylation was restored, demonstrating that *MoDMT1* is required for DNA methylation in *M. oryzae* (Figure 10b).

To provide a more quantitative view on roles of MoDMT1 and MoDMT2 in DNA methylation of *M. oryzae*, bisulphite-sequencing was carried out on a selected region (209bp) of MAGGY (Figure 11). This region was PCR-amplifed from bisulphite-treated genomic DNA of KJ201, *Δdmt1* and *Δdmt2*, and TA-cloned for Sanger sequencing. A part of *MoDMT1* exon sequences (214bp) that are free of cytosine methylation was included in the experiment in order to estimate conversion rate during bisulphite treatment. At least ten clones were sequenced for each strain. Of 43 cytosines present in the partial exon of *MoDMT1*, every cytosine (C) except one was converted to thymine (T), indicating 98% or higher conversion rate during bisulphite treatment, given that the remaining cytosine may be false-positive (data not shown). Bisulphite-sequencing of MAGGY sequences revealed that of 71 Cs present in 209bp region, 16 C positions showed 10 to 20% methylation (methylated C reads/total reads per site) in the wild-type. Most of methylation, meanwhile, disappeared in *Δdmt1* and 4 C positions remained methylated in *Δdmt2*. In addition, sequence context doesn't seem to be critical in selecting methylation site by two enzymes. This result is in accordance with the result from the previous Southern blot analysis of MAGGY methylation.

From these data, I conclude that MoDMT1 is the major DNA methyltransferase in *M. oryzae* and that MoDMT2 plays some role in DNA methylation but to limited extent. However, the possibility that MoDMT1 and MoDMT2 catalyze transfer of methyl group to target cytosines in a cooperative manner cannot be ruled out as suggested by the fact that a small number of methylated C positions remain in *Δdmt2*.

MoDMT1 and *MoDMT2* are generally down-regulated under stress conditions

To have an idea about roles of *MoDMT1* and *MoDMT2*, I examined the transcription of *MoDMT1* and *MoDMT2* under stress conditions. Real-time PCR (RT-PCR) analysis was carried out to check the relative expression of the two genes under stress conditions, including oxidative stresses, ionic stresses, UV, heat shock, and cell wall stresses, to expression without stress (CM) (Figure 13). Unlike the expression profile during fungal development, expressions of two genes were not correlated (correlation coefficient =-0.11). Nevertheless, under most conditions except UV stresses requiring DNA repair processes, the two DNA methyltransferase genes were in general, down-regulated, compared to the CM condition. This result suggests that under stress conditions, genomic DNA methylation should be suppressed.

MoDMT1 and *MoDMT2* play roles during vegetative growth

To investigate the function of DNA methylation in fungal biology, I searched for phenotypic changes in the mutants during the fungal life cycle. First, the mutants were grown in complete media (CM) and minimal media (MM) for 9 days and growth rate was measured

as diameter of fungal colonies. Growth rate of mutants was comparable to that of wild-type strain KJ201 (Figure 14a). However, the mutants exhibited different colony morphologies, compared to the wild-type (Figure 14b). Colonies of mutants seemed less fluffy at the growing margin than that of wild-type on CM. Autolysis of aerial hyphae on MM was consistently observed one or two days earlier in the mutants than in the wild-type. Contrary to expectation, this aberration was not exacerbated in double deletion mutants. Morphological aberrations manifested in colonies of the mutants suggest that *MoDMT1*- and *MoDMT2*-mediated DNA methylation may have roles in coordinating and developing structures above the substratum.

The observation of aberrant colony morphology led us to test if the mutants were compromised in the ability to produce biomass. Quantifying amount of hyphae embedded in agar media is technically challenging. Therefore I measured dry-weight of the fungus grown in complete media broth for 4 days. Measurement of dry-weight showed that the mutants were not able to produce as much biomass as wild-type (Figure 15).

Conidiation and appressorium formation are influenced by deletion of DNA methyltransferases

In both *Δdmt1* and *Δdmt2,* the ability to produce conidia was considerably impaired (Figure 14a). Conidiation of the mutants was only approximately 54%, 64% and 45% of that of the wild-type for *Δdmt1, Δdmt2* and *Δdmt1Δdmt2,* respectively. However, the shape and size of conidia produced in the mutants were indistinguishable from those in the wild-type. To explain the observed difference in conidiation, I compared conidiophore development

between wild-type and the mutants (Figure 16b). The mutants, in general, produced fewer number of conidiophores than wild-type. Mutant conidiophores were capable of bearing three or more conidia sympodially but production was retarded compared to wild-type. Defects in conidiation of mutants were restored to wild-type level by introduction of wild-type version of genes into the mutants, suggesting that *MoDMT1* and *MoDMT2* are required for efficient conidiogenesis (Figure 16a).

When conidial suspensions from the mutants were placed on plastic coverslip, on which appressorium formation can be induced, conidia of mutants were able to germinate as well those of wild-type (Figure 17). In contrast to germination rate, I observed that appressorium formation in *Δdmt1* was delayed during time course experiment (Figure 17 and 18). At 4 hours post induction (hpi) when ~30% of germinating conidia developed appressoria in the wild-type and *Δdmt2*, only about 10% of *Δdmt1* conidia were able to develop appressoria. At 8 hpi, appressorium formation of *Δdmt1* remained delayed but was able to catch up at 12 hpi. *Δdmt1Δdmt2* showed appressorium formation rate similar to *Δdmt1*. These differences observed in conidiation and appressorium formation between wild-type and mutants suggest that *MoDMT1* and *MoDMT2* is involved in fine-tuning rather than directing developmental processes.

***MoDMT1* and *MoDMT2* are dispensable for fungal pathogenicity**

To investigate the role of *MoDMT1* and *MoDMT2* in pathogenesis, conidial suspensions from *Δdmt1, Δdmt2,* and *Δdmt1Δdmt2* were sprayed onto rice seedlings of susceptible cv. Nagdongbye. This pathogenicity assay showed that all three mutants were capable of causing

diseases on rice plants (Figure 19). This indicates that the mutants have no defects in appressorium-mediated penetration into and invasive growth within host plants, despite delayed appressorium formation. I concluded that *MoDMT1* and *MoDMT2* are not required for fungal pathogenicity.

Bisulphite sequencing of *M. oryzae* genomes

To broaden and deepen the understanding of DNA methylation in *M. oryzae*, I used bisulphite-converted genomic DNA together with Illumina sequencing technology (methylC-seq) to generate a DNA methylation map across the fungal genome. I included in methylC-seq experiment genomic DNAs isolated from wild-type mycelia, conidia, appressoria, and mycelia of *Δdmt1* and *Δdmt2* mutants to monitor changes in methylome during the fungal development and in the absence of individual DNA methyltransferases. methylC-seq of these samples yielded about 19 to 21 million reads corresponding to sequence output ranging from 1.73 to 1.89 giga bases per sample after filtering as described (Materials and Methods) (Table 2). Lambda DNA, which doesn't have methylated cytosines, was included in bisulphite treatment of samples to calculate the conversion rate. Conversion rate ranged from 99.33% to 99.51% with minimal template degradation. To minimize the bias that may result from unequal number of reads per sample, ~18.9 million reads (1.7Gbp) were chosen at random from each sample. Among these reads, 72.65 to 81.57% (except for *Δdmt2* sample showing 62.75% mapping rate*)* of reads were successfully aligned to the converted reference genome, resulting in 30.1 to 33.8x coverage of the 41Mb genome (Table 3 and Figure 20). In this data, over 97% of total cytosines present in *M. oryzae* genome were covered by at least one

sequencing read and more than 88% of cytosines in *M. oryzae* genome were covered by at least four sequencing reads, allowing the methylation level of individual sites to be determined with reasonable confidence. Through this analysis, I identified varying number of mC sites from the nuclear genomes of the different samples. The samples that had the largest and smallest number of mCs were wild-type mycelia (46,124 mCs) and *Δdmt1* (4,563 mCs), respectively. The number of mCs identified in nuclear genome of mycelia accounts for ~0.22% of all genomic cytosines, or ~0.25% of the cytosines for which sufficient depth (more than four reads) were generated.

Validation of methylC-seq via bisulphite-PCR

To validate mC prediction made in methylC-seq, I performed bisulphite sequencing. Nine loci within *M. oryzae* genome that were predicted to contain mCs were chosen and subjected to PCR amplification following bisulphite treatment of genomic DNAs isolated from mycelia, conidia, appressoria, *ΔModmt1*, *ΔModmt2*, and *ΔModmt1ΔModmt2*. To control for conversion rate, one locus predicted to have no mCs was also included. PCR products were cloned and at least 10 independent inserts per each region in 6 types of genomic DNAs were sequenced. Comparison of sequencing results to mCs identified in methylC-seq revealed that over 90% of mCs identified in sequencing data overlapped with mCs identified in methylC-seq data, indicating that identification of mCs using methylC-seq is highly reliable (Table 4, Figure 21 and Figure 22). Some mC sites identified in methylC-seq data was not found to be methylated in bisulphite sequencing. This discrepancy was, however, probably due to low methylation level (10-30%) in combination with lack of power to detect methylated cytosines

using only ~10 sequencing reads, demonstrating the greater sensitivity of methylC-seq approach in detecting mCs.

Dynamic changes in DNA methylome during fungal development

Genome-wide DNA methylation profiles revealed that number of mCs decreased during fungal development. A total of 46,124 mCs was identified in mycelia, while 14,530 mCs and 11,585 mCs were identified in conidia and appressoria, respectively (Figure 23a). This observation is consistent with real-time PCR data (Figure 8) that showed less abundance of DNA methyltransferases transcripts in conidiation and appressorium formation than in mycelia, suggesting that reduction in DNA methylation was not the result of stochastic losses of DNA methylation but a process that is regulated at the transcriptional level. In mycelia of *M. oryzae*, cytosine methylation was present in all sequence contexts. However, there seemed to be a preference for CG and CHH context: about half of mCs occurred in the CG context and the remaining half in the CHH context, while only about 8% of mCs was found in CHG context. However, mCs in each sequence context didn't decrease in equal proportion during development. Although the absolute number of mC sites in all sequence contexts decreased, the mCs in CG context showed larger reduction in conidia and appressoria, than mCs in other sequence contexts (Figure 23b).

In an attempt to understand the dynamics of DNA methylation underlying the reduction in mCs sites during conidiation and appressorium formation, I compared mC sites between samples (Table 5 and 6). Through comparison, I noticed that overall reduction in mC sites during development is not simply the result of loss of methylation in existing mC sites. For

example, in comparison of mC sites between mycelia and conidia, I observed that considerable number of mC sites was created anew in conidia, although a large proportion of mC sites identified in conidia was the same as those methylated in mycelia. Interestingly, new mC sites were enriched in exon regions rather than in other genomic features such as promoters and introns.

MethylC-seq data for *ΔModmt1* and *ΔModmt2* showed that total number of mCs was reduced dramatically with preference to mCs in the CG context in *ΔModmt1* (to less than 10% of wild-type) but reduced to 80% of wild-type in *ΔModmt2*. MoDMT2 is an ortholog of RID that has no DNA methyltransferase activity in *N. crassa*. From Southern blot analysis of MAGGY methylation, I hypothesized that MoDMT1 is a major player of DNA methylation in *M. oryzae* and MoDMT2 has some but limited roles in DNA methylation. This data provide evidence supporting this conclusion at genomic scale and suggest that unlike its ortholog RID, MoDMT2 may possess DNA methyltransferase activity. Comparison of mC sites between wild-type and *ΔModmt1* showed that few mC sites overlap between two samples probably due to loss of major maintenance mechanism of DNA methylation (Table 7). This is in contrast to comparison between mycelia and conidia, or conidia and appressoria (Table 5 and 6), supporting that reduction in mC sites during development is a controlled process. In all samples of which DNA methylome was examined in this work, however, the percentage average methylation level (defined as (number of mC reads / total number of reads per each cytosine site)x100) peaked around 30%, regardless of sequence contexts (Figure 24). Dramatic changes in average methylation levels were not observed except for *ΔModmt1*. Absence of *MoDMT1* shifted the peak to lower level.

Mosaic pattern of DNA methylation in the fungal genome

In order to examine the genome-wide pattern of distribution of mCs in *M. oryzae*, the density of mCs (number of mC sites/length of bin) was plotted across the fungal chromosomes (Figure 25). Density of mCs fluctuated drastically across chromosomes, indicating that relatively dense methylated domains are interspersed with regions that are not methylated or methylated at low density. As expected, densely methylated domains in mycelia shrink in conidia and appressoria, and nearly disappear in *ΔModmt1*, while they are maintained in *ΔModmt2*.Such clustered pattern of DNA methylation observed in mycelia positively correlates with the position of transposable elements and negatively correlates with gene density. Nevertheless, significant proportion of mC sites was found in and around annotated genes (2,868 genes in mycelia) (Figure 26). The vast majority of genes associated with DNA methylation were either predicted or hypothetical proteins, making it difficult to link DNA methylation and gene ontology analysis.

DNA methylation in genic regions

To explore the distribution of DNA methylation in genomic features, I analyzed the methylation profiles for the whole genome, genes (exons and introns), and upstream (1.5kb) and downstream (1kb) of genes (Figure 27). Transposable elements (TEs) were not included in this analysis since they are poorly annotated in the reference genome sequence of *M. oryzae*. Compared to the whole genome, density of mCs in genic regions (defined as promoter + exon + intron + downstream) was low, indicating relatively intensive methylation in intergenic regions of the genome. Notably, I found that methylation is higher in upstreams

and downstreams of genes than in exons and introns. However, the average methylation level did not change significantly among features (Figure 28). To scrutinize the effect of DNA methylation in the fungus, I plotted density and methylation level of mC sites across genic regions and normalizaed over genes associated with DNA methylation (Figure 29 and 30). DNA methylation density peaks at upstream and downstream sequences of genes. These peaks showed a sharp drop at the boundaries of coding sequences. Gene body methylations occurred mostly at around the start and end of coding sequences and were depleted at the center of coding region. The hypermethylation in upstream and downstream regions of genes almost disappeared in conidia, appressoria, and *ΔModmt1*, suggesting the regulatory role of DNA methylation and demethylation during asexual development and infection-specific development. In accordance with the whole genome, average methylation level did not change dramatically. It was, however, noteworthy that the average methylation level was higher in upstream and downstream of gene than in coding sequences (Figure 30).

As mentioned above, most of genes associated with DNA methylation encode either predicted or hypothetical proteins to which biochemical functions cannot be assigned. I compared length of methylation-associated genes with that of genes having no methylation at all. This comparison revealed that methylation-associated genes are, on average, longer than genes with no methylation (Figure 31). The mean length of the methylation-associated genes was 2.78kb, which is longer by 30% than the mean length of genes with no methylation (1.89kb). The difference in average length was significant at $p<2.2x10^{-16}$, based on permutation test. This observation suggests the possibility of DNA methylation being involved in enhancing transcriptional accuracy or splicing efficiency.

DNA methylation in transposable elements of *M. oryzae*

Interestingly, I observed enrichment of mC sites in chromosome 1, 3, and, 7 of the nuclear genome in mycelia (Figure 32). These three chromosomes have higher density of mC sites than the remaining chromosomes (expected density = 10.9 mC sites per 10kb). The portion of the genome unassigned to chromosomes showed the highest density, probably due to highly repetitive nature of these sequences. Density of mC sites in three chromosomes had little correlation with gene density but considerable degree of correlation (r = 0.79) with density of TEs in each chromosome. This data supports that TEs and their neighboring sequences are one of the major targets for DNA methylation in the fungus.

I next checked if different types of TEs would be methylated at similar density and level. For eleven types of TEs whose defining sequences are available, density and methylation level of mCs sites in each type of TE were calculated and compared. This analysis revealed that TEs were differentially methylated, depending on length and type (Figure 33). In general, the density of mC sites negatively correlated with the length of TEs. Especially, two undefined TEs, MGR619 and MGR608 were heavily methylated, compared to other types of TEs. These TEs are short, dispersed, highly polymorphic sequences that frequently appear together as contiguous sequences (Khang et al., 2008). It is known that many degenerate LTR-retrotransposons are sandwiched between copies of MGR619/MGR608 (205bp). However, it is not clear why these TEs are more heavily methylated than others.

DNA methylation and transcriptional activity of genome

To examine the effect of DNA methylation in transcriptional activity of genome, I employed RNA-seq technology (Table 8). In *M. oryzae*, TEs are poorly annotated, so I did separate analysis on TEs from gene expression analysis. In comparison of resulting transcriptomes, I found that 1.2% and 2.4% of annotated genes have altered transcript abundance (≥ 2-fold difference) in the mycelia of mutants, compared to those of wild-type (Figure 34). Inverse correlation between DNA methylation and transcript abundance was not obvious from these data. Nonetheless, genes having methylation in their upstream and downstream regions were expressed at a lower level than genes having methylation in coding sequences or no methylation, supporting roles of DNA methylation in transcription of those genes (Figure 35). In this study, transcriptomes of conidia and appressoria were not investigated since the effect of DNA methylation on transcriptions of these samples can be masked by genetic factors involved in development. Therefore, it cannot be rejected that development-specific factors can translate changes in DNA methylation landscapes during development into changes in transcriptional outputs.

TEs are thought to be suppressed by DNA methylation. However, generality of this idea hasn't been tested at the genomic scale. In analysis of RNA-seq data on TEs, surprisingly, changes in transcript abundance of TEs displayed diverse pattern, depending on not DNA methylation but types and locations (Figure 36). These data suggest that DNA methylation is not the general mechanism to suppress transcriptional activity of TEs in *M. oryzae*.

DISCUSSION

DNA methylation is a conserved mechanism of chromatin modification in eukaryotes. In plants and mammals, it has been established that a large fraction of cytosine residues is methylated and proper methylation is required for normal development. In fungi, most of our understanding on DNA methylation came from studies of *N. crassa*. These studies suggested that a much smaller fraction of cytosines is methylated and its function is thought to be limited to suppressing transcriptional activity of transposable elements (Rountree and Selker, 1997; Kouzminova and Selker, 2001; Selker et al., 2003). The work using genetic and genomic technologies revealed new aspects DNA methylation in the fungal kingdom, challenging the conventional view of fungal DNA methylation as a stable mark for silencing of transposable elements.

The most notable finding is that DNA methylation is reprogrammed during fungal development in a way that massive demethylation of pre-existing mC sites is accompanied by *de-novo* methylation of new cytosine sites. Such dynamic nature of DNA methylation has not been reported in fungi to date and has intriguing parallels to mammalian DNA methylation during stem cell differentiation. In mammals, embryonic stem cells undergo global demethylation during pre-implantation development, followed by lineage-dependent DNA methylation in lineage-committed cells (Reik et al., 2001; Santos et al., 2002). This global reprogramming of DNA methylation ensures differential transcriptional program in each cell type. In *M. oryzae*, as development progresses from mycelia through conidia to appressoria, developmental potency of cells becomes restricted: appressoria have no option

but penetration for host infection. In contrast to mammalian system, DNA methylation in *M. oryzae* is reduced as developmental potency decreases. This work also suggests that DNA methylation may not be the primary mechanism of suppressing transcriptional activity of TEs. Together with the developmental defects observed in DNA methyltransferase mutants, our findings suggest that fungal DNA methylation may have assumed newer roles than controlling proliferation of invasive elements such as transposable elements.

Nakayashiki *et al*. reported that DNA methylation is not the main force suppressing TE activity (Nakayashiki et al., 2001), which is consistent with our conclusion on TEs and DNA methylation. One key difference is that they showed transcriptional up-regulation of MAGGY when it is demethylated by 5-azacytidine treatment. RNA-seq analysis in *Δdmt1* showed that most MAGGYs in the genome are not up-regulated by genetically-forced demethylation. It is generally accepted that diverse environmental stresses can induce transcriptional activation of transposable elements (Ikeda et al., 2001; Ogasawara et al., 2009). Therefore it is possible that up-regulation of MAGGY by 5-azacytidine treatment may be a stress response of the fungus caused by 5-azacytidine itself.

In our study, it doesn't appear that DNA methylation has a primary role in directing development, since the methylation-deficient mutants were able to produce conidia and appressoria in a reduced and delayed manner. Analysis of expression by RNA-seq showed that loss of DNA methylation has significant, though not dramatic, effects on transcriptional activity of fungal genome. These data suggest that DNA methylation in *M. oryzae* may contribute to fungal development by fine-tuning transcriptional activity of the genome. The role of DNA methylation as a fine-tuner is supported by differences in transcript abundance

of genes having methylation in upstream, gene body, and downstream regions of genes (Figure 35). It is not clear yet how the observed pattern of DNA methylation in and around genes is translated into differences in transcript abundance.

One interesting conjecture that arises from these data and the literature is that DNA methylation in fungi is relevant to polarity of growth. For example, it was shown in *C. albicans* that DNA methylation of some genes changes during switching between hyphal and yeast form (Mishra et al., 2011). In addition, it was reported that chemically-induced demethylation can lead to reduction in production of asexual spores, which are less polar, compared to mycelia (Tamame et al., 1983b). Since polarity of the growth form is associated with pathogenesis in many plant and animal pathogens, this conjecture deserves more attention in future research.

Considering the lack of DNA methylation in a model fungus, *Saccharomyces cerevisiae*, this study provides a compelling model system in which epigenetic components can be investigated in conjunction with development of microbial eukaryotes. I anticipate that my work would give new insights into evolution of DNA methylation in microbial eukaryotes and provide the basis to expand our perspective on fungal development.

REFERENCES

Autran, D., Huanca-Mamani, W., and Vielle-Calzada, J.P. 2005. Genomic imprinting in plants: the epigenetic version of an Oedipus complex. Curr Opin Plant Biol 8:19-25.

Charalambous, M., da Rocha, S.T., and Ferguson-Smith, A.C. 2007. Genomic imprinting, growth control and the allocation of nutritional resources: consequences for postnatal life. Curr Opin Endocrinol Diabetes Obes 14:3-12.

Chen, T., and Li, E. 2004. Structure and function of eukaryotic DNA methyltransferases. Curr Top Dev Biol 60:55-89.

Chi, M.-H. 2009. A quick and safe method for fungal DNA extraction. The Plant Pathology Journal 25:108-111.

Clark, S.J., Statham, A., Stirzaker, C., Molloy, P.L., and Frommer, M. 2006. DNA methylation: bisulphite modification and analysis. Nat Protoc 1:2353-2364.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452:215-219.

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.

Drinnenberg, I.A., Fink, G.R., and Bartel, D.P. 2011. Compatibility with killer explains the rise of RNAi-deficient fungi. Science 333:1592.

Ferguson-Smith, A.C. 2011. Genomic imprinting: the emergence of an epigenetic paradigm. Nat Rev Genet 12:565-575.

Freitag, M., Williams, R.L., Kothe, G.O., and Selker, E.U. 2002. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. Proc Natl Acad Sci U S A 99:8802-8807.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A 89:1827-1831.

Hellman, A., and Chess, A. 2007. Gene body-specific methylation on the active X chromosome. Science 315:1141-1143.

Ikeda, K., Nakayashiki, H., Takagi, M., Tosa, Y., and Mayama, S. 2001. Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea*. Mol Genet Genomics 266:318-325.

Illingworth, R., Kerr, A., Desousa, D., Jorgensen, H., Ellis, P., Stalker, J., Jackson, D., Clee, C., Plumb, R., Rogers, J., Humphray, S., Cox, T., Langford, C., and Bird, A. 2008. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. PLoS Biol 6:e22.

Jones, D.T., Taylor, W.R., and Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275-282.

Jones, P.A., Taylor, S.M., and Wilson, V.L. 1983. Inhibition of DNA methylation by 5-azacytidine. Recent Results Cancer Res 84:202-211.

Khang, C.H., Park, S.Y., Lee, Y.H., Valent, B., and Kang, S. 2008. Genome organization and evolution of the AVR-Pita avirulence gene family in the *Magnaporthe grisea* species complex. Mol Plant Microbe Interact 21:658-670.

Kim, S., Ahn, I.P., Rho, H.S., and Lee, Y.H. 2005. *MHP1*, a *Magnaporthe grisea* hydrophobin gene, is required for fungal development and plant colonization. Mol Microbiol 57:1224-1237.

Kouzminova, E., and Selker, E.U. 2001. *dim-2* encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. EMBO J 20:4309-4323.

Kritsky, M.S., Filippovich, S., Afanasieva, T.P., Bachurina, G.P., and Russo, V.E. 2001. The involvement of photoregulated DNA methylation into a choice between sexual and asexual developmental pathways in *Neurospora crassa*. Dokl Biochem Biophys 378:203-205.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., and Wei, C.L. 2010. Dynamic changes in the human methylome during differentiation. Genome Res 20:320-331.

Lee, Y.H., and Dean, R.A. 1993. cAMP Regulates Infection Structure Formation in the Plant Pathogenic Fungus *Magnaporthe grisea*. Plant Cell 5:693-700.

Li, E., Bestor, T.H., and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69:915-926.

Li, R., Yu, C., Li, Y., Lam, T.-W., Yiu, S.-M., Kristiansen, K., and Wang, J. 2009. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25:1966-1967.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. 2008. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell 133:523-536.

Mishra, P.K., Baum, M., and Carbon, J. 2011. DNA methylation regulates phenotype-dependent transcriptional activity in *Candida albicans*. Proc Natl Acad Sci U S A 108:11965-11970.

Nakayashiki, H., Ikeda, K., Hashimoto, Y., Tosa, Y., and Mayama, S. 2001. Methylation is not the main force repressing the retrotransposon MAGGY in *Magnaporthe grisea*. Nucleic Acids Res 29:1278-1284.

Ogasawara, H., Obata, H., Hata, Y., Takahashi, S., and Gomi, K. 2009. Crawler, a novel Tc1/mariner-type transposable element in *Aspergillus oryzae* transposes under stress conditions. Fungal Genet Biol 46:441-449.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. Cell 99:247-257.

Ponger, L., and Li, W.H. 2005. Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. Mol Biol Evol 22:1119-1128.

Reik, W., Dean, W., and Walter, J. 2001. Epigenetic reprogramming in mammalian development. Science 293:1089-1093.

Rountree, M.R., and Selker, E.U. 1997. DNA methylation inhibits elongation but not initiation of transcription in Neurospora crassa. Genes Dev 11:2383-2395.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406-425.

Sambrook, J., and Russel, D.W. 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

Santos, F., Hendrich, B., Reik, W., and Dean, W. 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol 241:172-182.

Schumacher, A., Kapranov, P., Kaminsky, Z., Flanagan, J., Assadzadeh, A., Yau, P., Virtanen, C., Winegarden, N., Cheng, J., Gingeras, T., and Petronis, A. 2006. Microarray-based DNA methylation profiling: technology and applications. Nucleic Acids Res 34:528-542.

Selker, E.U., Tountas, N.A., Cross, S.H., Margolin, B.S., Murphy, J.G., Bird, A.P., and Freitag, M. 2003. The methylated component of the *Neurospora crassa* genome. Nature 422:893-897.

Slotkin, R.K., and Martienssen, R. 2007. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8:272-285.

Suzuki, M.M., and Bird, A. 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 9:465-476.

Tamame, M., Antequera, F., Villanueva, J.R., and Santos, T. 1983a. High-frequency conversion to a "fluffy" developmental phenotype in *Aspergillus* spp. by 5-azacytidine treatment: evidence for involvement of a single nuclear gene. Mol Cell Biol 3:2287-2297.

Tamame, M., Antequera, F., Villanueva, J.R., and Santos, T. 1983b. 5-Azacytidine Induces Heritable Biochemical and Developmental-Changes in the Fungus *Aspergillus-Niger*. Journal of General Microbiology 129:2585-2594.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596-1599.

Trapnell, C., Pachter, L., and Salzberg, S.L. 2009. TopHat: discovering splice junctions with RNA-seq. Bioinformatics 25:1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562-578.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schubeler, D. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet 37:853-862.

Wilkinson, J.R., Kale, S.P., Bhatnagar, D., Yu, J., and Ehrlich, K.C. 2011. Expression Profiling of Non-Aflatoxigenic *Aspergillus parasiticus* Mutants Obtained by 5-Azacytosine Treatment or Serial Mycelial Transfer. Toxins (Basel) 3:932-948.

Yu, J.H., Hamari, Z., Han, K.H., Seo, J.A., Reyes-Dominguez, Y., and Scazzocchio, C. 2004. Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi. Fungal Genet Biol 41:973-981.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328:916-919.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet 39:61-69.

Table 1. List of primers used in this study

Name	Sequences (5′→3′)	Use
MGG00889 UF	GTGCAGGTTCGTTCCTACTT	KO
MGG00889 UR	GCACAGGTACACTTGTTTAGAGATCTGATGGTCAAGGTGAGAA	KO
MGG00889 DF	CCTTCAATATCATCTTCTGTCGACATCATTGTCCTGGAAAACA	KO
MGG00889 DR	GGTTGGCTGGCTAAACTAGA	KO
MGG00889 nested F	AATGGATCATCAGCGAAGG	KO
MGG00889 nested R	TGATGGGTTATTCGTAATGGC	KO
MGG00889 RT F	AAGGTAAATCCGACTGAACA	RT
MGG00889 RT R	AGGCTTGCGAGAAATGAC	RT
MGG02795 UF	GCAGGACCGCATCTCGCTCA	KO
MGG02795 UR	GCACAGGTACACTTGTTTAGAGACCCTCGCGGCATCCCTCAG	KO
>MGG02795 DF	CCTTCAATATCATCTTCTGTCGAGTACCATGCGCCTTCCACA	KO
MGG02795 DR	AAGCAGGCTCCGAAGAAGAA	KO
MGG02795 nested F	CATCGAGCGGCGCAAGAAG	KO
MGG02795 nested R	GCCTTTTCGCCACCATTCT	KO
MoDMT2 RT F	CAGCTCTATTGTCATTTGTGGC	RT
MoDMT2 RT R	TCCGAGTCTGCAAAGTTGTC	RT
MGG_889 BS F	TTAAAAAAGGTAAATTTGATTGAAT	BS-PCR
MGG_889 BS R	CAAACCACCTTATATTTCCATA	BS-PCR
MAGGY BS F	GAATATAATTAATGATTTGATTTGA	BS-PCR
MAGGY BS R	TCATACTACTTTAACCAAAACATT	BS-PCR
MGG_10510 P_F	AAAAAATAGGATTATATAAGGAA	BS-PCR
MGG_10510 P_R	TTAAAATATCAAATAATTTATTTTT	BS-PCR
MGG_15334 P_F	AAGGTTGAGTAATTATTGGTAA	BS-PCR
MGG_15334 P_R	CCACAACTTTATAATAATTTTTC	BS-PCR
MGG_16073 ORF F	ATGTTAATAATAATGGTGTTAATAG	BS-PCR
MGG_16073 ORF R	CTACCCTATACCAAATTTAATAC	BS-PCR
MGG_16885 ORF F	TATTAAATTTGGTATAGGGTAGT	BS-PCR
MGG_16885 ORF R	TTATTTTACATATAATAAACCCAC	BS-PCR
Chr1_Ig F	GTTATTATGTTTGATAGTGATTATT	BS-PCR
Chr1_Ig R	TACATACTCACAATTTACAAAAA	BS-PCR

Table 2. Statistics of methylC-seq data after filtering

Sample	Library	Insert size (bp)	Conversion rate (%)	Read length (bp)	Clean reads	Clean data (Gbp)
KJ201 mycelia	MAGfflHAQDEAAPEMI-17	338	99.51	90	18,888,890	1.70
KJ201 conidia	MAGcfkHABDEAAPEMI-2	348	99.38	90	18,888,890	1.70
KJ201 appressoria	MAGcfkHACDEAAPEMI-3	357	99.43	90	18,888,890	1.70
KJ201 *Δdmt1*	MAGcfkHADDEAAPEMI-4	337	99.36	90	18,888,890	1.70
KJ201 *Δdmt2*	MAGdadHACDEAAPEMI-57	365	99.33	90	18,888,890	1.70

Table 3. Data statistics after reads alignment

Sample	Raw reads (M)	Raw data (Gb)	Mapped reads (M)	Mapped data (Gb)	Average map rate (%)	Whole genome average coverage depth (X)
Mycelia	18.89	1.70	14.58	1.31	77.18	31.98
Conidia	18.89	1.70	15.41	1.39	81.57	33.80
Appressoria	18.89	1.70	14.12	1.27	74.76	30.98
ΔModmt1 mycelia	18.89	1.70	13.72	1.24	72.65	30.10
ΔModmt2 mycelia	18.89	1.70	11.85	1.07	62.75	26.00

Table 4. List of loci that are validated for DNA methylation in conidia-derived mycelia

Locus		Feature	B[b]	M[c]	B∩M[d]
Chr.[a]	position				
1	4653636-4653982	upstream of MGG_10510	9	9	8
2	5278750-5279211	upstream of MGG_15334	15	18	14
1	821210-821420	Intergenic region	15	16	13
3	5011776-5011934	ORF (MGG16885)	13	15	11
5	1232184-1232244	ORF (MGG00889)[e]	0	0	0
7	Multiple positions including Supercontig 7: 3282757-3282812	a retrotransposon, MAGGY[e]	10	14	9

[a]Chromosome

[b]Number of mC sites in BS-PCR experiment

[c]Number of mC sites in methylC-seq

[d]Number of mC sites that overlap between BS-PCR and methylC-seq

[e]Sequences that are predicted to have no methylation or methylation by Southern blot analysis and used as negative and positive control of the experiment

*Note: BS-PCR and Sanger sequencing for loci listed above were also carried out for conidia, appressoria, and *ΔModmt1*, confirming little or absence of mC sites predicted in methylC-seq.

Table 5. Genome-wide comparison of mC sites between mycelia and conidia

Genome features	Context	Mycelia only	Mycelia ∩ Conidia	Conidia only	Total
UPSTREAM (1.5kb)	CG	4487	1018	290	5795
	CHG	811	108	227	1146
	CHH	4043	839	968	5850
	Total	9341	1965	1485	12791
EXON	CG	919	137	561	1617
	CHG	451	22	557	1030
	CHH	1548	122	1918	3588
	Total	2918	281	3036	6235
INTRON	CG	257	60	48	365
	CHG	74	9	62	145
	CHH	413	75	246	734
	Total	744	144	356	1244
DOWNSTREAM (1.0kb)	CG	3439	700	206	4345
	CHG	554	67	113	734
	CHH	2909	516	637	4062
	Total	6902	1283	956	9141
INTERGENIC	CG	10769	2191	660	13620
	CHG	1591	263	181	2035
	CHH	7507	1600	815	9922
	Total	19867	4054	1656	25577

Table 6. Genome-wide comparison of mC sites between conidia and appressoria

Genome features	Context	Conidia only	Conidia ∩ Appressoria	Appressoria only	Total
UPSTREAM (1.5kb)	CG	925	383	525	1833
	CHG	301	34	251	586
	CHH	1549	258	1295	3102
	Total	2775	675	2071	5521
EXON	CG	662	36	594	1292
	CHG	574	5	614	1193
	CHH	1995	45	2244	4284
	Total	3231	86	3452	6769
INTRON	CG	88	20	63	171
	CHG	70	1	57	128
	CHH	300	21	311	632
	Total	458	42	431	931
DOWNSTREAM (1.0kb)	CG	635	271	381	1287
	CHG	158	22	166	346
	CHH	969	184	843	1996
	Total	1762	477	1390	3629
INTERGENIC	CG	2090	761	1053	3904
	CHG	338	106	182	626
	CHH	1806	609	948	3363
	Total	4234	1476	2183	7893

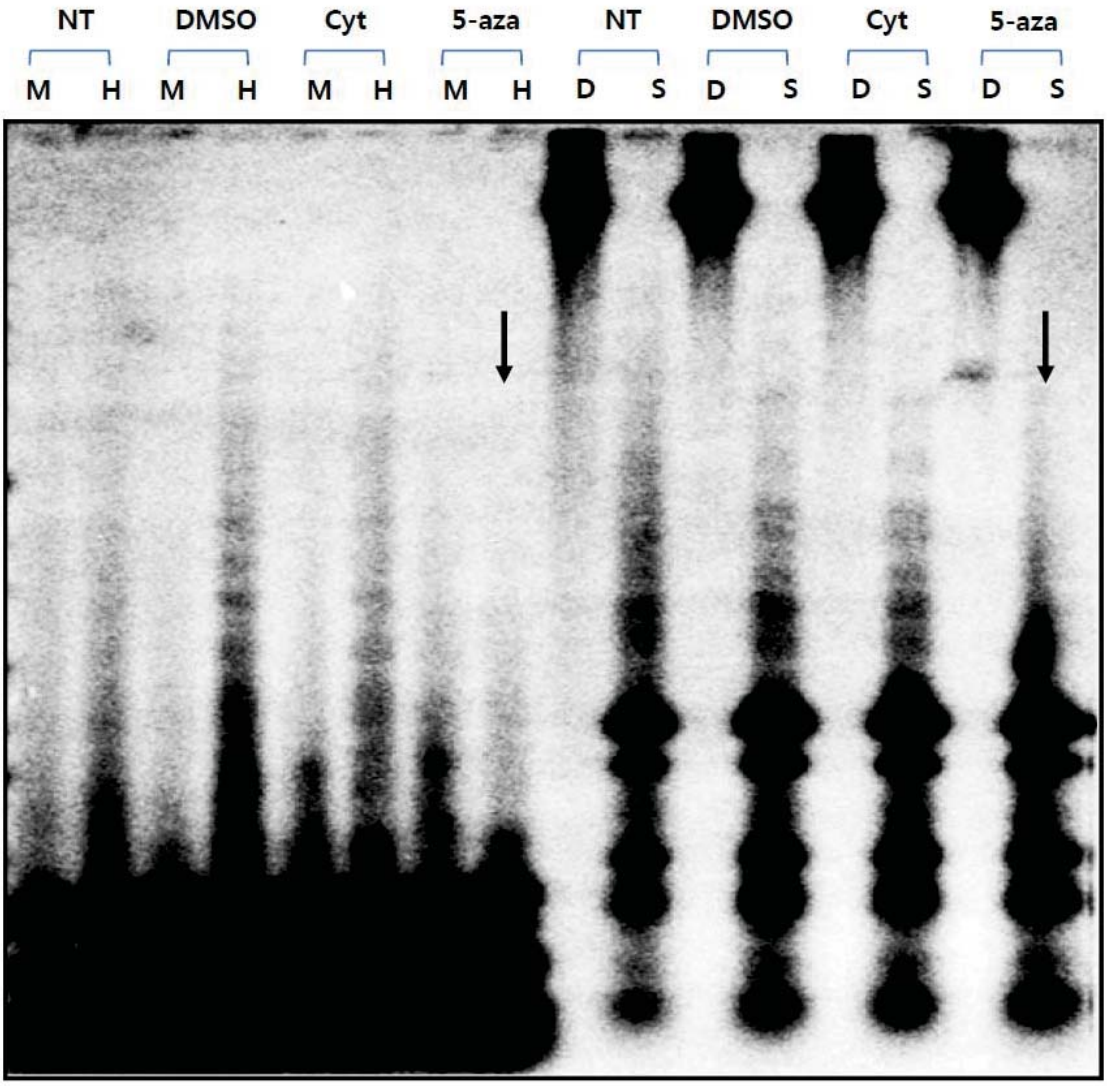
Table 7. Genome-wide comparison of mC sites between wild-type and *Δdmt1*

Genome features	Context	Mycelia	Mycelia ∩ *Δdmt1*	*Δdmt1*	Total
UPSTREAM (1.5kb)	CG	5502	3	118	5623
	CHG	916	3	145	1064
	CHH	4878	4	594	5476
	Total	11296	10	857	12163
EXON	CG	1050	6	575	1631
	CHG	473	0	570	1043
	CHH	1668	2	1820	3490
	Total	3191	8	2965	6164
INTRON	CG	317	0	37	354
	CHG	83	0	45	128
	CHH	488	0	179	667
	Total	888	0	261	1149
DOWNSTREAM (1.0kb)	CG	4137	2	67	4206
	CHG	619	2	83	704
	CHH	3420	5	349	3774
	Total	8176	9	499	8684
INTERGENIC	CG	12958	2	34	12994
	CHG	1852	2	28	1882
	CHH	9101	6	165	9272
	Total	23911	10	227	24148

Table 8. Statistics of RNA-seq experiment

	KJ201		***Δdmt1***		***Δdmt2***	
	1	**2**	**1**	**2**	**1**	**2**
Read No	64,370,558	64,370,558	52,563,922	52,563,922	53,862,448	53,862,448
Length	6,501,426,358	6,501,426,358	5,308,956,122	5,308,956,122	5,440,107,248	5,440,107,248
total mapped read	112,223,694		92,153,620		94,980,311	
mapped	57,162,678	55,061,016	46,823,080	45,330,540	48,175,787	46,804,524
properly paired	104,844,824		86,393,318		89,123,394	
mate mapped	105,101,488		86,591,990		89,333,332	
singletons	7,122,206		5,561,630		5,646,979	

Figure 1. Southern blot analysis of DNA methylation in MAGGY. Two pairs of isoschizomers were used: *Msp*I (cytosine methylation-insensitive)-*Hpa*II (methylation sensitive) that recognizes CCGG and *Dpn*I (adenosine methylation-dependent)-*Sau3A*I (adenosine methylation-independent but cytosine methylation-sensitive) that recognizes GATC. Cytidine and 5-aza were added to the final concentration of 20μM. Genomic DNAs were isolated from mycelia grown in complete media broth containing each chemical. NT, no treatment; DMSO, dimethylsulfoxide; Cyt, cytidine; Aza, 5-azacytidine; M, *Msp*I; H, *Hpa*II; D, *Dpn*I; S, *Sau3A*I.


NT
DMSO
Cyt
5-aza
NT
DMSO
Cyt
5-aza
M H M H M H M H D S D S D S D S

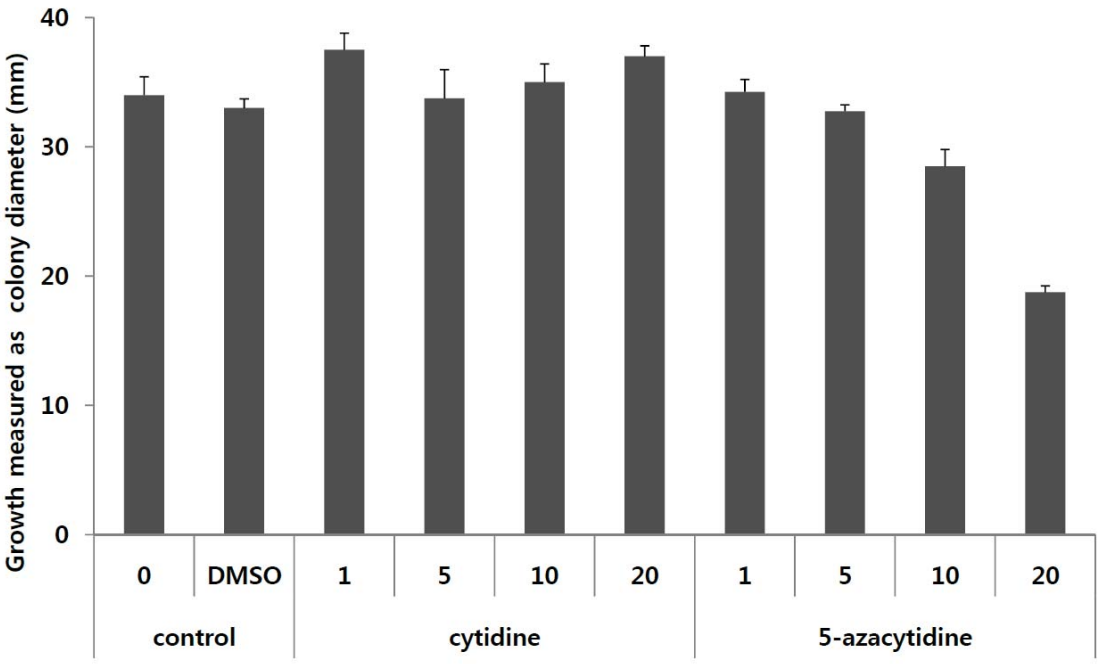


Figure 2. Growth of fungus on minimal media containing chemicals with varying concentration (μM). Growth was measured as colony diameter 7 days after inoculation with four replicates for each treatment. Diameter of initial inoculum is 5mm, which is the diameter of cork borer (No. 1).

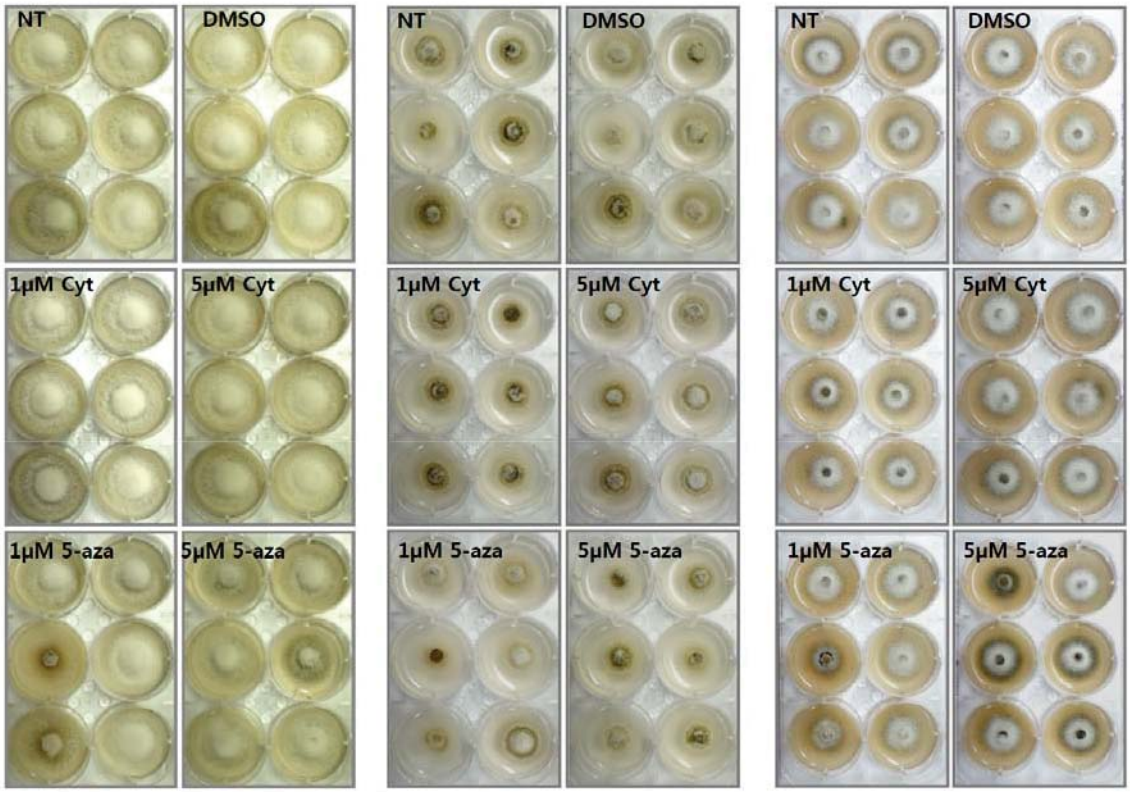
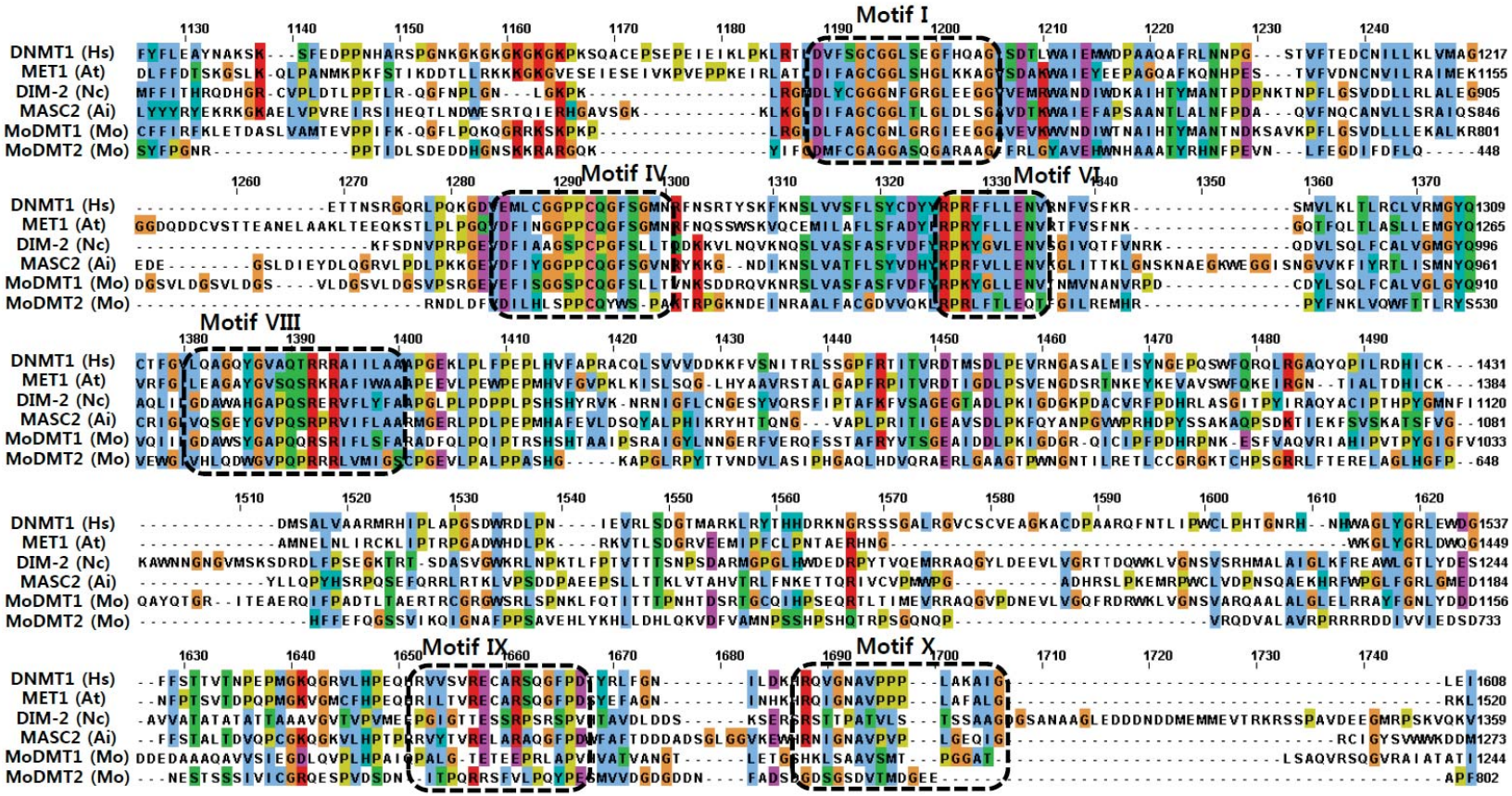


Figure 3. Morphological changes of the fungus on complete media, minimal media, and V8 media (from left to right) containing chemicals. Wild-type fungus was grown in the media for 7 days. NT, no treatment; DMSO, dimethylsulfoxide; Cyt, cytidine; 5-aza, 5-azacytidine.

Figure 4. Sequence alignment of MoDMT1 and MoDMT2 with known DNA methyltransferases. Multiple sequence alignment was carried out using ClustalW2 and a part of alignment was taken for presentation of motifs in C terminus. Hs, *Homo sapiens*; At, *Arabidopsis thaliana*; Nc, *Neurospora crassa*; Ai, *Ascobulus immersus*; Mo, *Magnaporthe oryzae*.


Motif I
Motif IV
Motif VI
Motif VIII
Motif IX
Motif X
DNMT1 (Hs)
MET1 (At)
DIM-2 (Nc)
MASC2 (Ai)
MoDMT1 (Mo)
MoDMT2 (Mo)

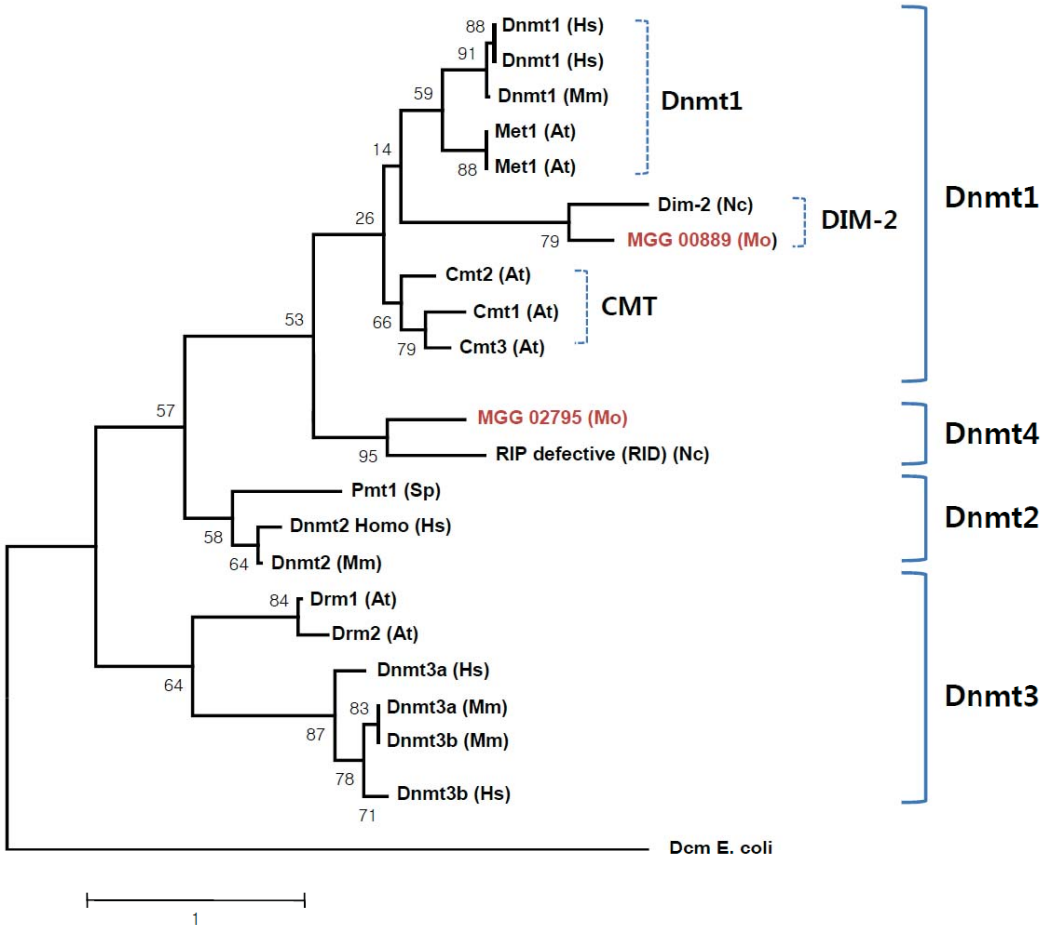


Figure 5. Phylogeny of DNA methyltransferases including two from *M. oryzae*. Phylogenetic tree was constructed using neighbor-joining method with Dam sequences from *E. coli* as an outgroup. Hs, *Homo sapiens*; Mm, *Mus musculus*; At, *Arabidopsis thaliana*; Nc, *Neurospora crassa*; Mo, *Magnaporthe oryzae*; Sp, *Schizosaccharomyces pombe*.

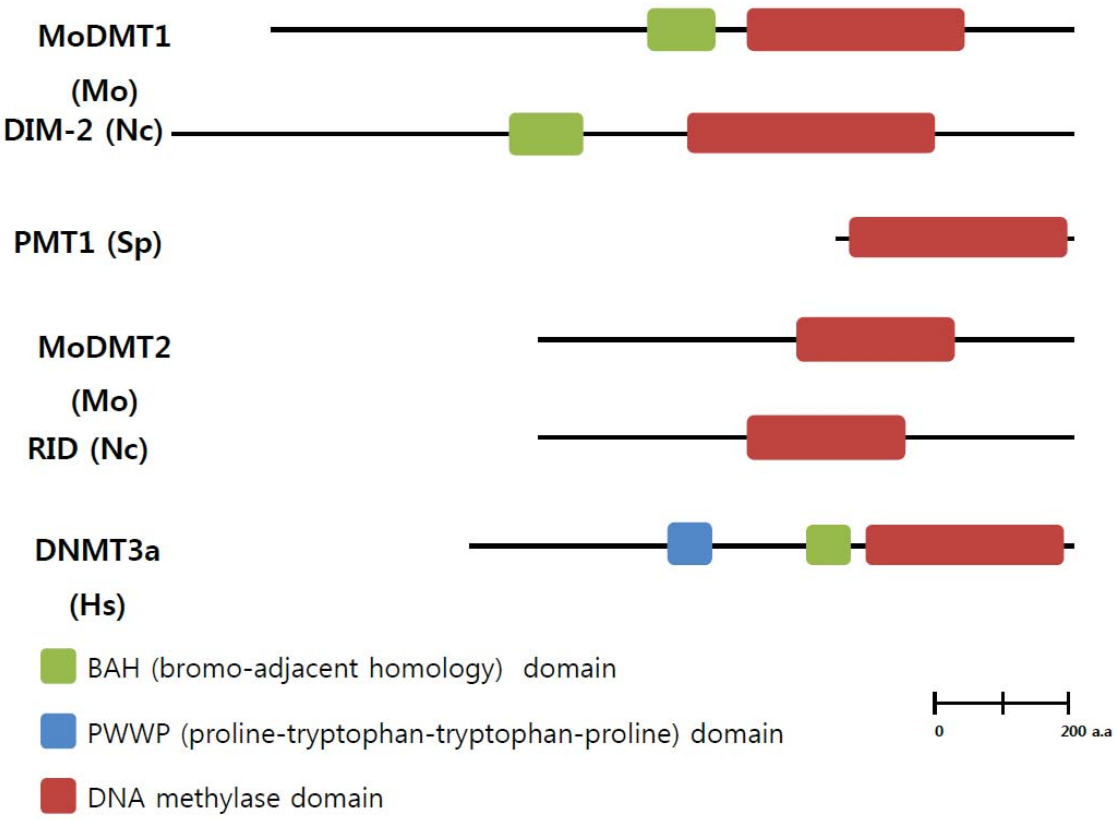
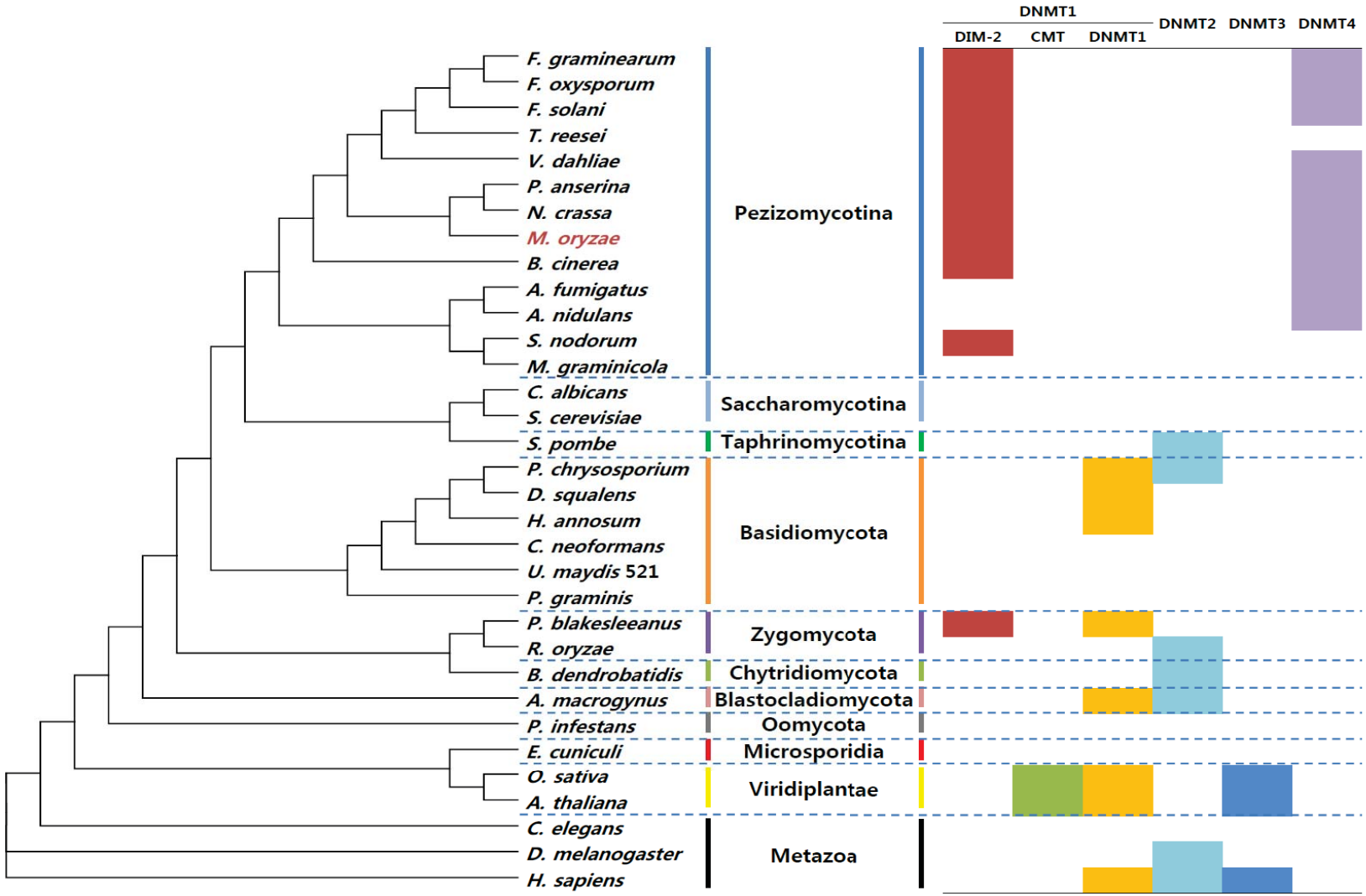


Figure 6. Domain architecture of DNA methyltransferases. Representative DNA methyltransferase(s) for each family is presented for comparison with those from *M. oryzae*. Mo, *Magnaporthe oryzae*; Nc, *Neurospora crassa*; Sp, *Schizosaccharomyces pombe*; Hs, *Homo sapiens*.

Figure 7. Kingdom-wide distribution of DNA methyltransferases. Cladogram was constructed for species that are representatives of each phylum by CVtree with the following options: K-tuple length=7 and sequence type= amino acid (http://tlife.fudan.edu.cn/cvtree/). Presence (colored box) and absence (box with no color) of each family of DNA methyltransferases in different taxa was determined using BLAST matrix function embedded in CFGP (http://cfgp.riceblast.snu.ac.kr/main.php). To remove spurious hits, e-value cutoff of $1e\text{-}10^{-5}$ and score cutoff of 100 were initially used and remaining hits were manually examined.


DNMT1
DIM-2
CMT
DNMT1
DNMT2
DNMT3
DNMT4
F. graminearum
F. oxysporum
F. solani
T. reesei
V. dahliae
P. anserina
N. crassa
M. oryzae
B. cinerea
A. fumigatus
A. nidulans
S. nodorum
M. graminicola
C. albicans
S. cerevisiae
S. pombe
P. chrysosporium
D. squalens
H. annosum
C. neoformans
U. maydis 521
P. graminis
P. blakesleeanus
R. oryzae
B. dendrobatidis
A. macrogynus
P. infestans
E. cuniculi
O. sativa
A. thaliana
C. elegans
D. melanogaster
H. sapiens
Pezizomycotina
Saccharomycotina
Taphrinomycotina
Basidiomycota
Zygomycota
Chytridiomycota
Blastocladiomycota
Oomycota
Microsporidia
Viridiplantae
Metazoa

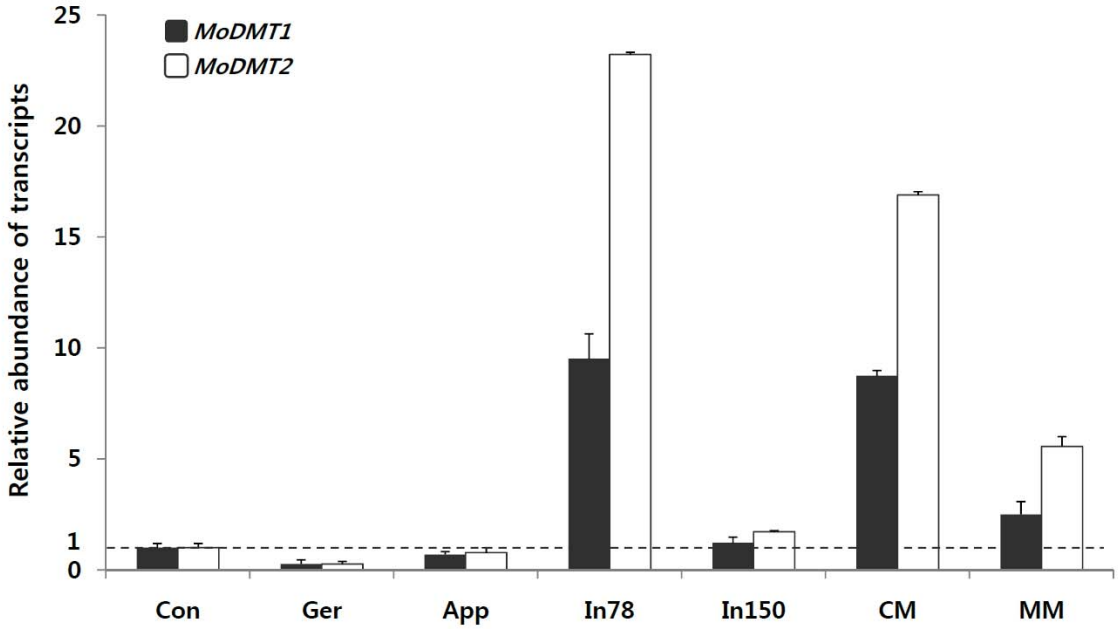


Figure 8. Expression profile of *MoDMT1* and *MoDMT2* during fungal development and vegetative growth. Expression level in each development and condition was calculated using *β-tubulin* as reference gene. The relative abundance of transcripts was measured relative to expression level in conidia. Con, conidia; Ger, germinating conidia; App, appressoria; In78, 78 hours post inoculation; In150, 150 hours post inoculation; CM, complete media; MM, minimal media.

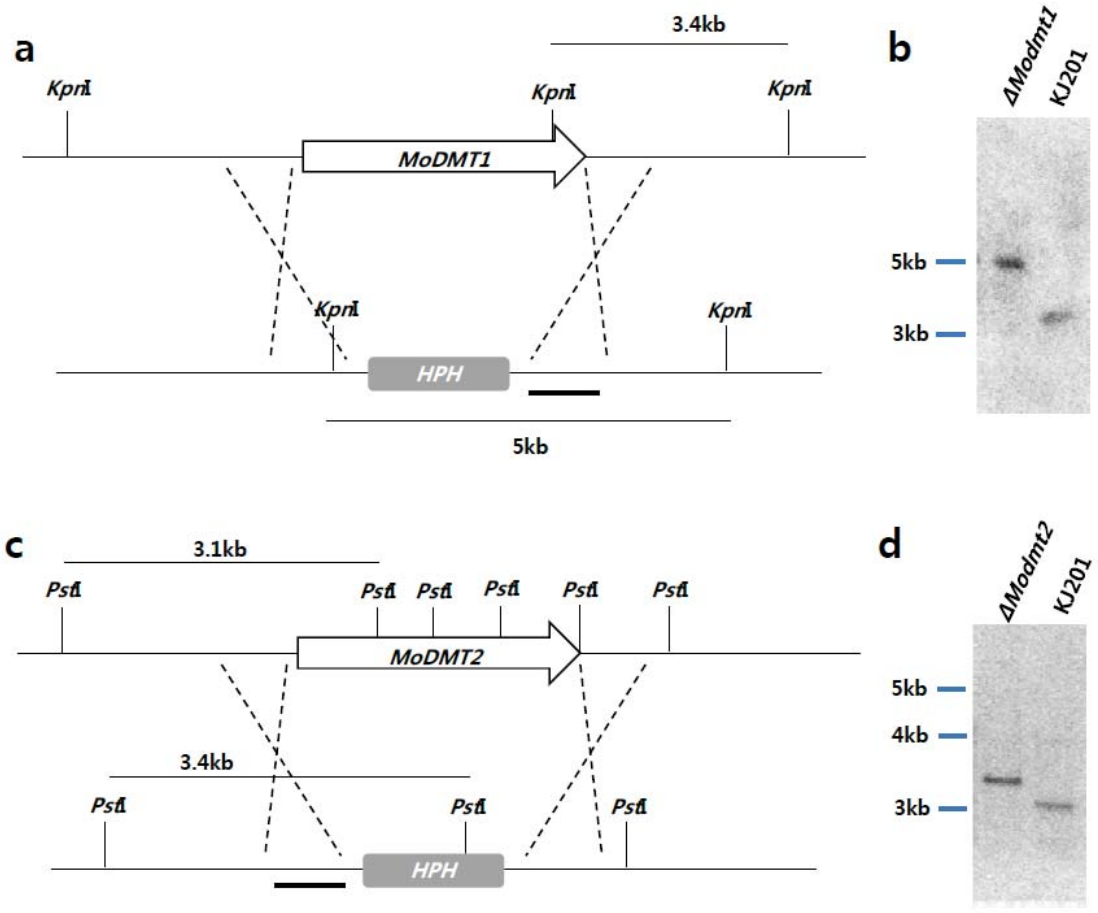


Figure 9. Targeted deletion of *MoDMT1* and *MoDMT2* in *M. oryzae*. **a** and **c,** schematic diagram of gene deletion. By homologous recombination, target genes were replaced by *hph* cassette . A bold line below the disruption construct represents the sequences used as a probe in Southern blot analysis. **b** and **d**, Southern blot analysis for deletion of*MoDMT1* and *MoDMT2*. A representative result is shown.

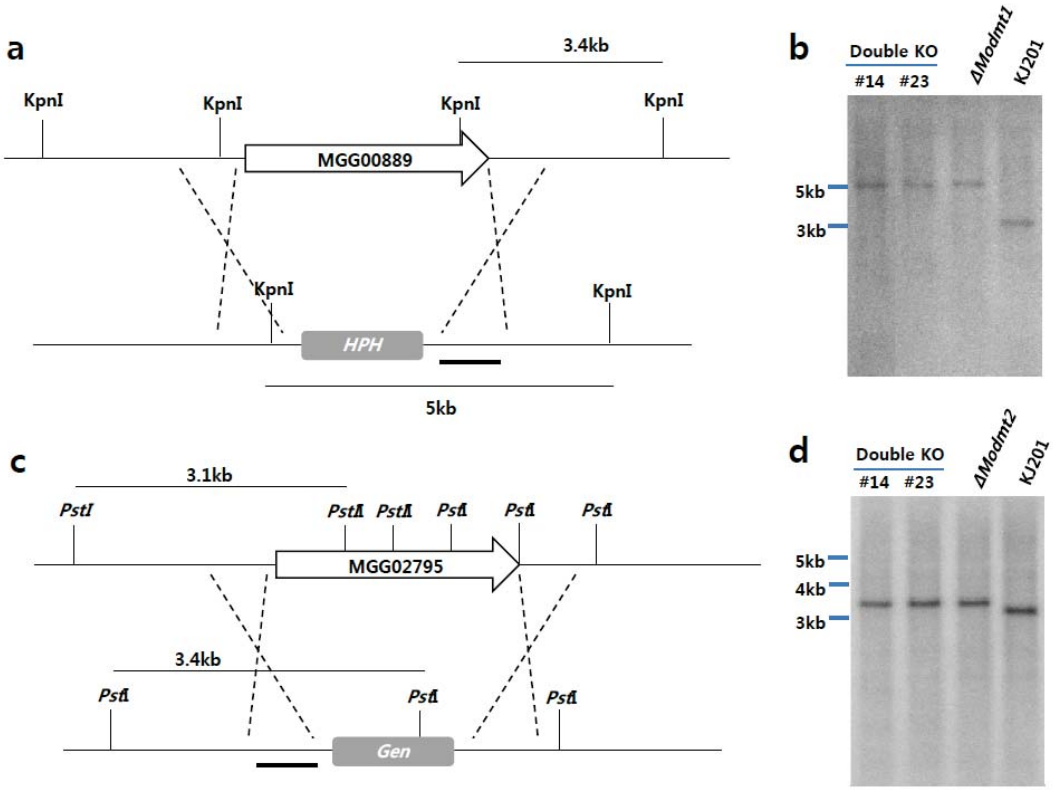


Figure 10. Generation of double deletion mutants of *MoDMT1* and *MoDMT2* in *M. oryzae*. **a** and **c,** schematic diagram of gene deletion. By homologous recombination, target genes were replaced by *hph* cassette. A bold line below the disruption construct represents the sequences used as a probe in Southern blot analysis. **b** and **d**, Southern blot analysis of deletion mutants for *MoDMT1* and *MoDMT2*. A representative result is shown.

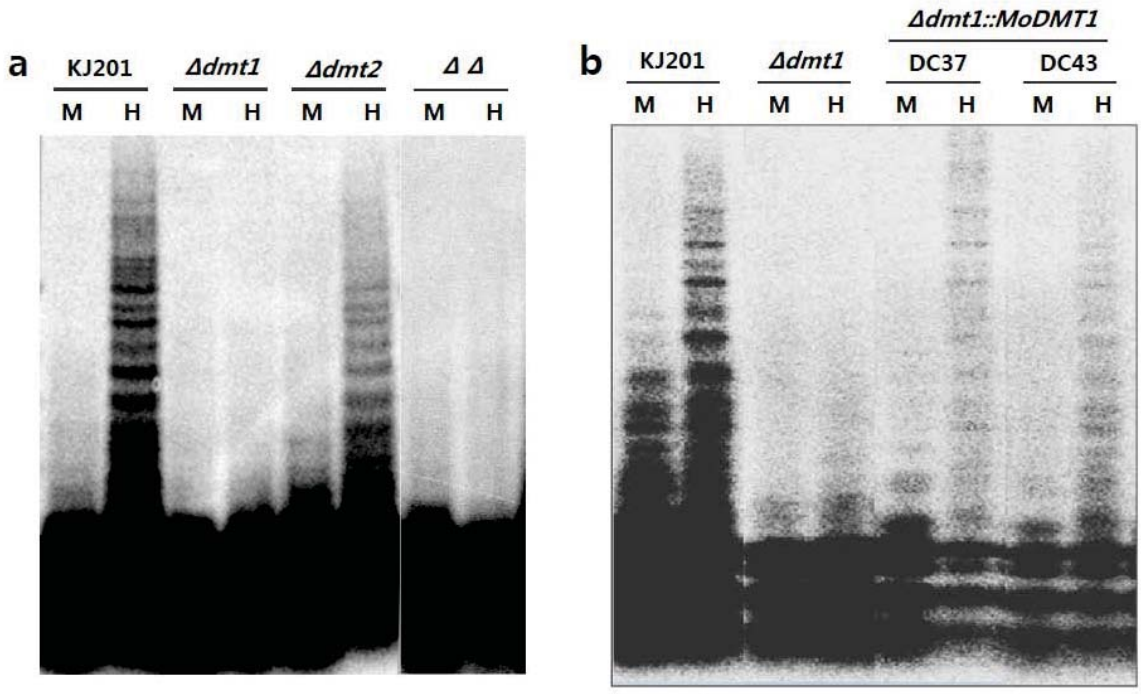


Figure 11. Southern hybridization analysis of DNA methylation in MAGGY. To assess the methylation level of MAGGY in KJ201 and mutants, isoschizomers, *Msp*I (methylation-insensitive) and *Hpa*II (methylation-sensitive) were used to digest genomic DNAs. **a**, MAGGY methylation in *Δdmt1* and *Δdmt2* in comparison to KJ201. **b**, Complementation of methylation defect in *Δdmt1* by introduction of native *MoDMT1* copy. Note that the amount of DNA used in each lane is not equal in b.

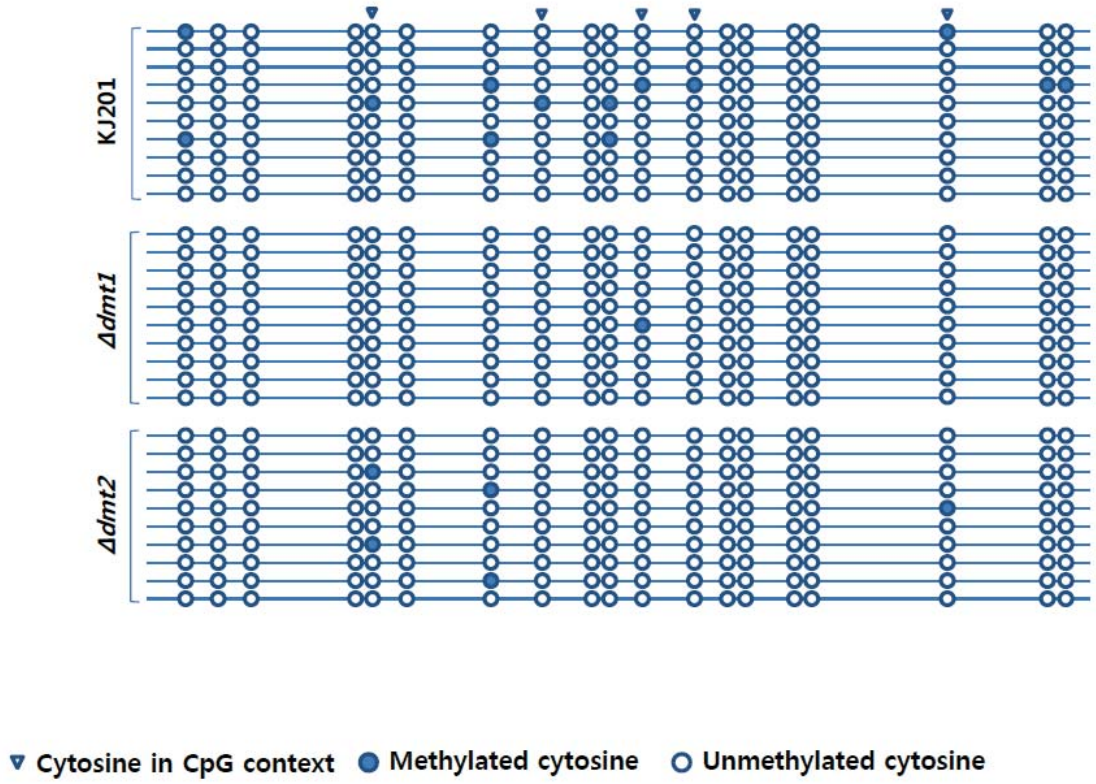


Figure 12. Methylation map of a part of MAGGY sequences determined by bisulphite-sequencing. A part of MAGGY sequences (209bp) were amplified from bisulphite-treated genomic DNAs and TA-cloned. Ten clones per each sample were picked randomly for their inserts to be sequenced by Sanger sequencing. Each horizontal line represents a part of amplicon (61bp) and circles on the line represent positions of cytosines.

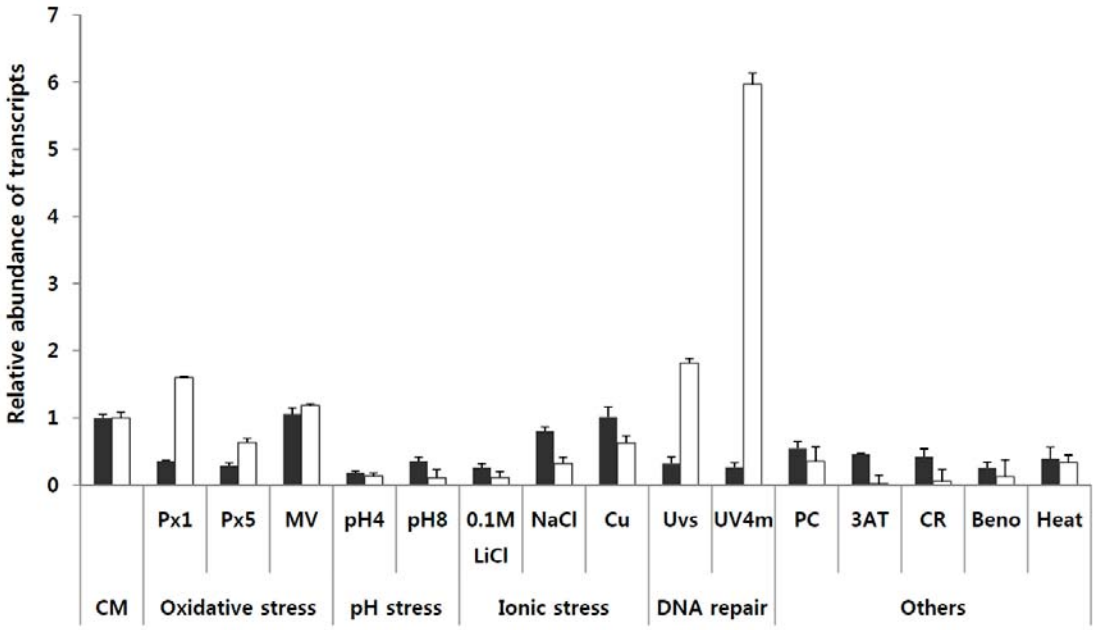


Figure 13. Expression profile of *MoDMT1* and *MoDMT2* under various stress conditions. Expression level in each condition was calculated using *β-tubulin* as reference gene. The fold change was calculated relative to expression level in complete media condition without stress.

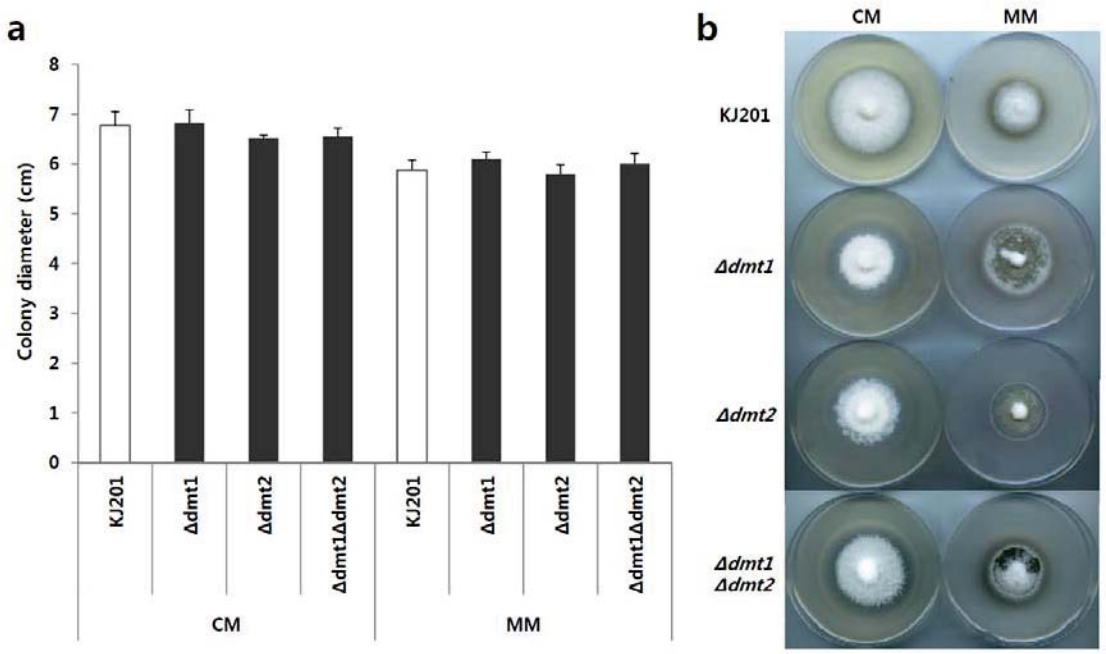


Figure 14. Growth and colony morphology of *Δdmt1, Δdmt2,* and *Δdmt1Δdmt2* on complete media (CM) and minimal media (MM). CM and MM were inoculated with the fungus and incubated at 25℃ for 9 days. a, Growth was measured as diameter of colony with three replicates. b, colony morpholoy of *Δdmt1*, *Δdmt2* and *Δdmt1Δdmt2* on CM and MM.

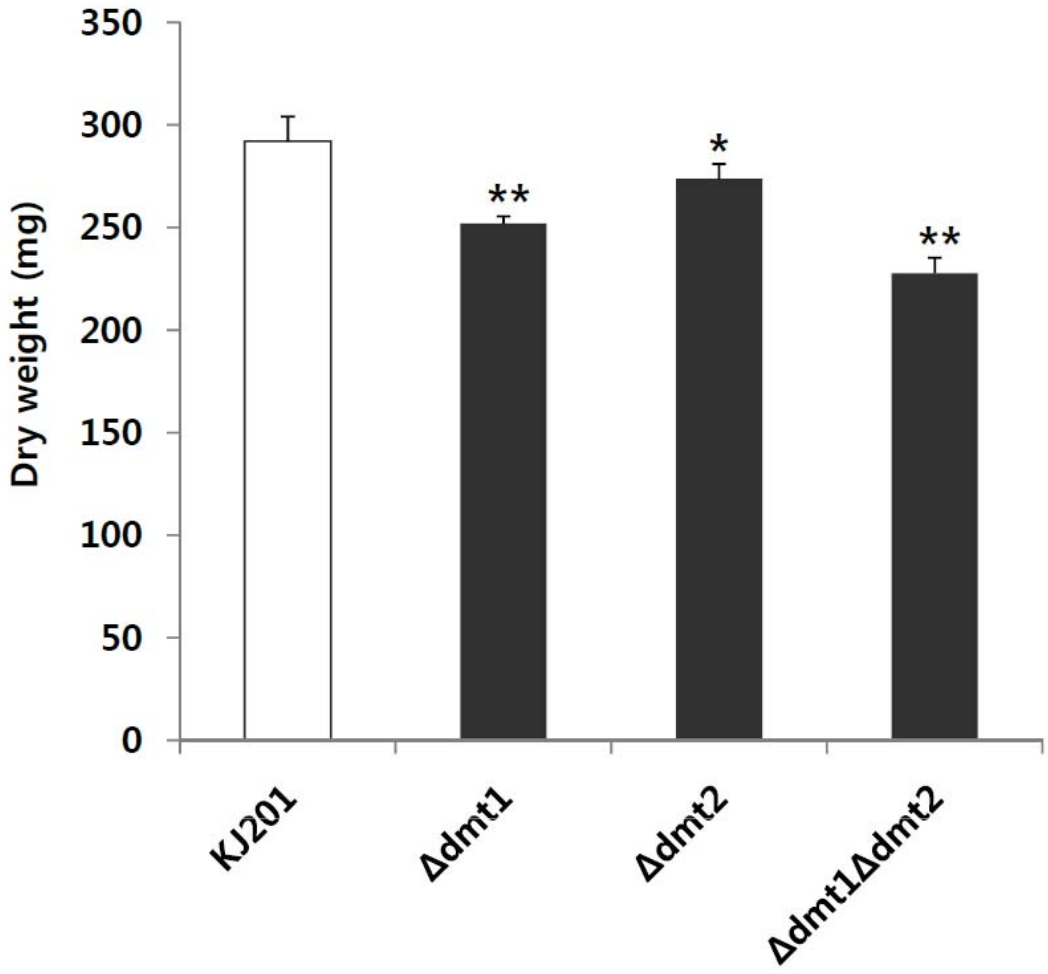


Figure 15. Dry weight measurement of DNA methyltransferase mutants grown in complete media (CM). CM were inoculated with spore suspensions of equal concentration (~1x10^5/ml) and incubated at 25℃ for 4 days. **p<0.001 and *p<0.05.

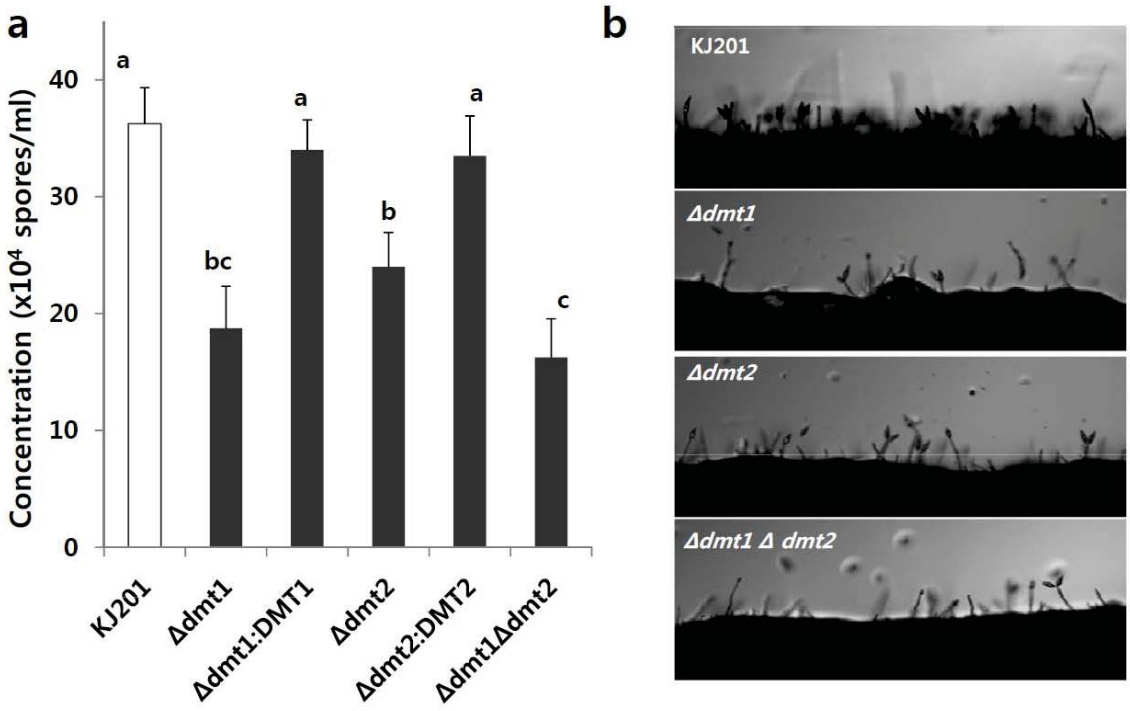
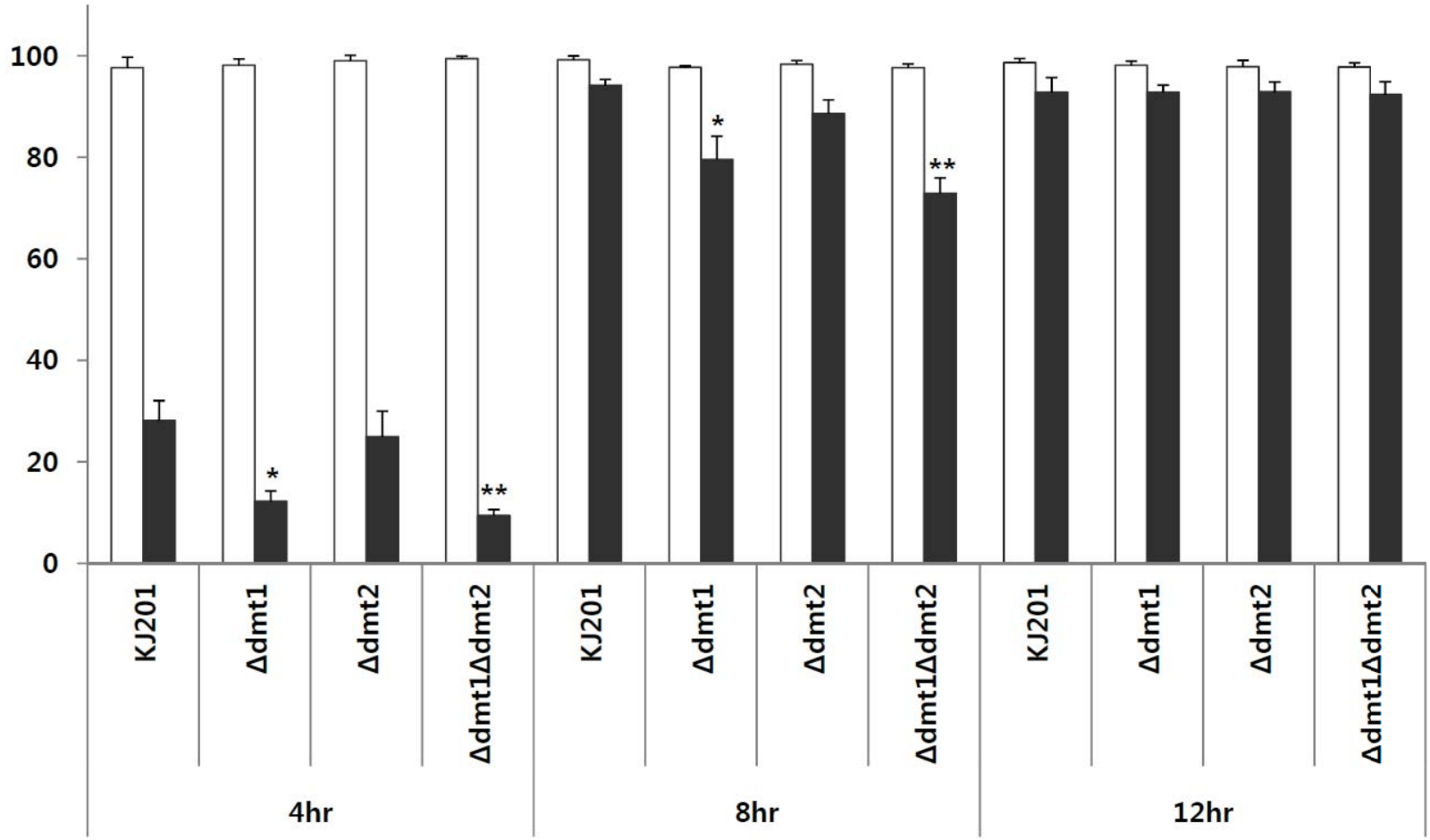


Figure 16. Conidiation of DNA methyltransferase mutants. **a,** Conidiation measurement**.** conidia were harvested using 5ml of water from 8 days old oatmeal agar in 6-well plates with 2 days of aeration. The number of conidia was counted using hemacytometer in 4 replicates per strain. Tukey-Kramer method was used to test statistical significance of difference. **b**, conidiophore development at 8hpi on oatmeal agar block.

Figure 17. Germination and appressorium formation of *Δdmt1* and *Δdmt2* on hydrophobic surface. Conidia were harvested from 8 day-old oatmeal agar plates. A 40μl of conidial suspension was pipetted onto plastic coverslips to induce appressorium formation following adjustment of concentration to 2~4x10^4/ml. Germination (# of germinating conidia/# of total conidia) and appressorium formation rate (# of conidia developing appressoria/# of germinating conidia) was measured at 4, 8, and 12hrs post induction. White bar represents germination ratio and black bar represents appressorium formation rate in percentage. **$p<0.001$ and *$p<0.01$.


100
80
60
40
20
0
*
**
*
**
KJ201
Δdmt1
Δdmt2
Δdmt1Δdmt2
KJ201
Δdmt1
Δdmt2
Δdmt1Δdmt2
KJ201
Δdmt1
Δdmt2
Δdmt1Δdmt2
4hr
8hr
12hr

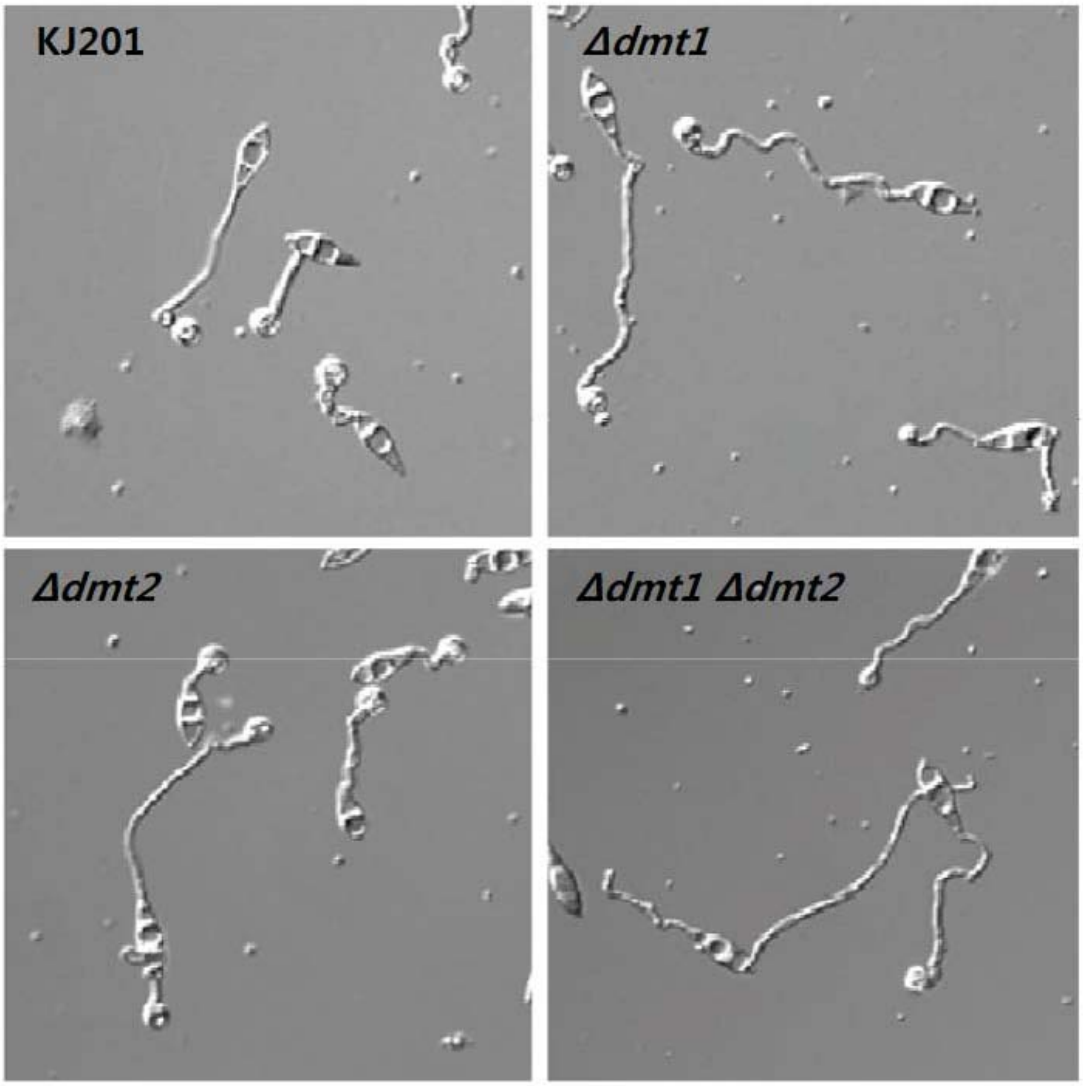


Figure 18. Appressorium formation of DNA methyltransferase mutants. Spores were harvested using watef from 8 days old cultures (oatmeal agar plate) and 40µl was dropped on plastic cover slips to induce appressorium formation with three replicates. Images were taken 12hrs post induction.

Figure 19. Pathogenicity of *Δdmt1, Δdmt2* and *Δdmt1Δdmt2* on rice plants. Conidia were harvested from 8 day-old oatmeal agar plates and suspended in 10ml of sterile distilled water. Following adjustment of concentration to ~1x10^5/ml, condia suspensions were sprayed onto susceptible rice cultivar, Nagdong. Disease symptoms were assessed 7 day post inoculation.

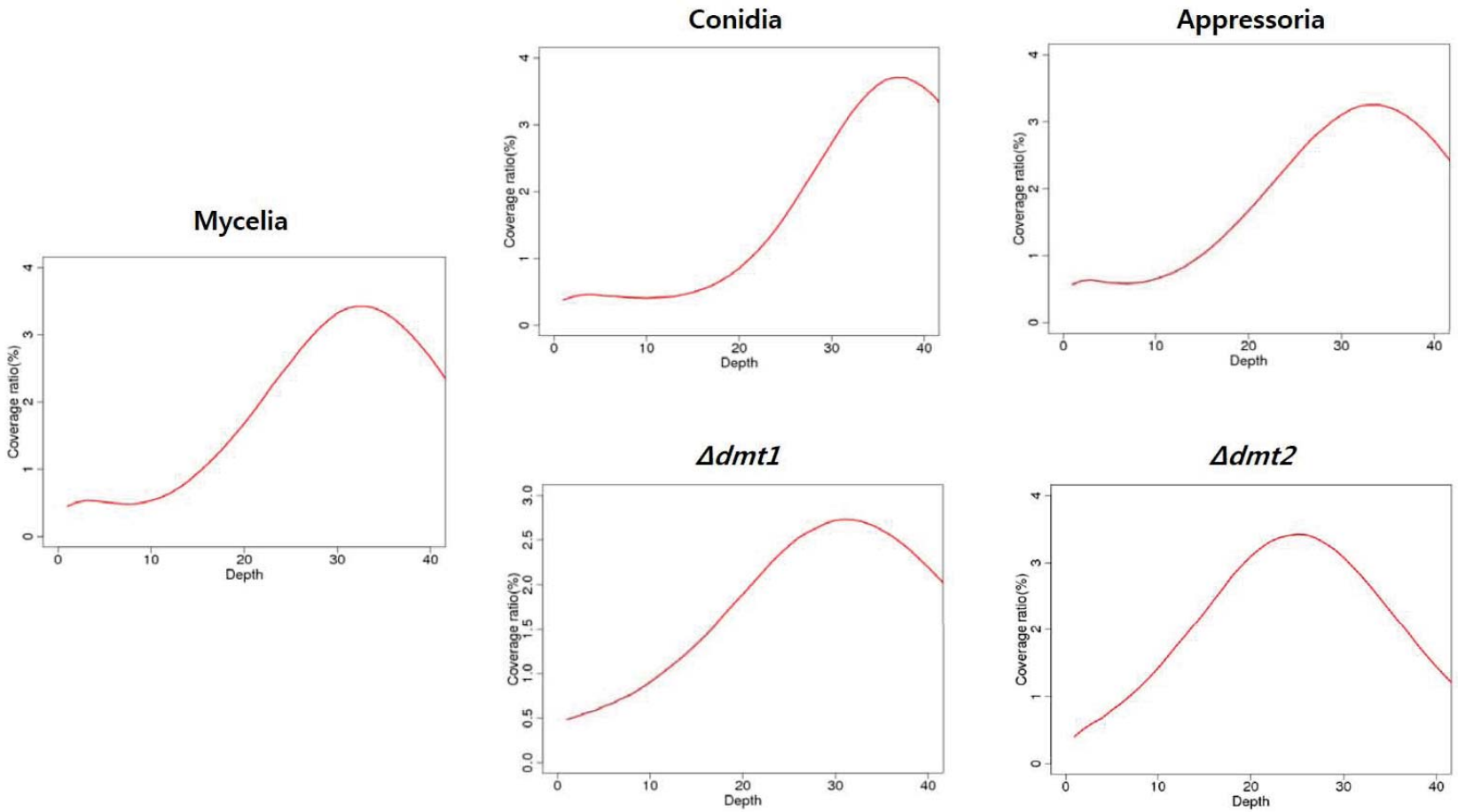


Figure 20. Distribution of read depth. Percentage of nucleotide sites was plotted against read depth that covers the sites.

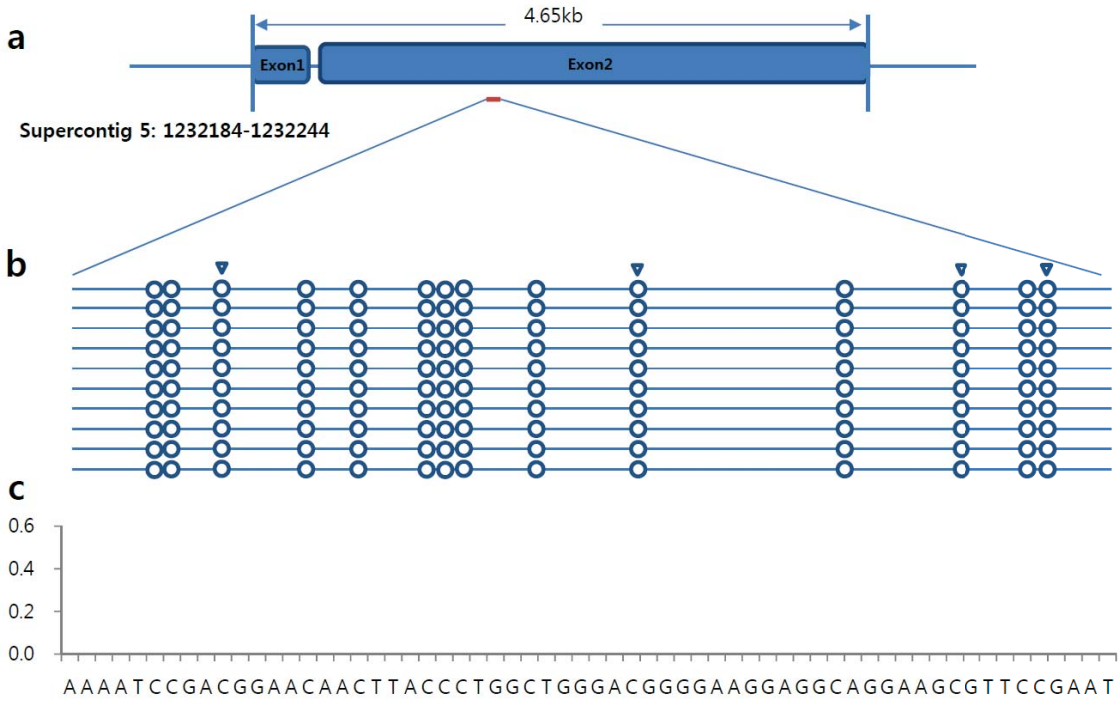


Figure 21. Comparison of bisulphite sequencing to methylC-seq data on an unmethylated region. **a**, schematic diagram of ORF, from which amplicon is derived. **b**, schematic summary of sequencing results. Each horizontal line represents independent sequencing read of a part of amplicon (red bar below ORF). Empty circle represents unmethylated cytosine and filled circle represents methylated cytosine within the amplicon. Triangle indicates position of cytosines in CG context. **c**, summary of methylC-seq data on the corresponding region. X-axis represents nucleotides on that region and y-axis represents methylation level of cytosine. All the data shown here is for mycelia of wild-type.

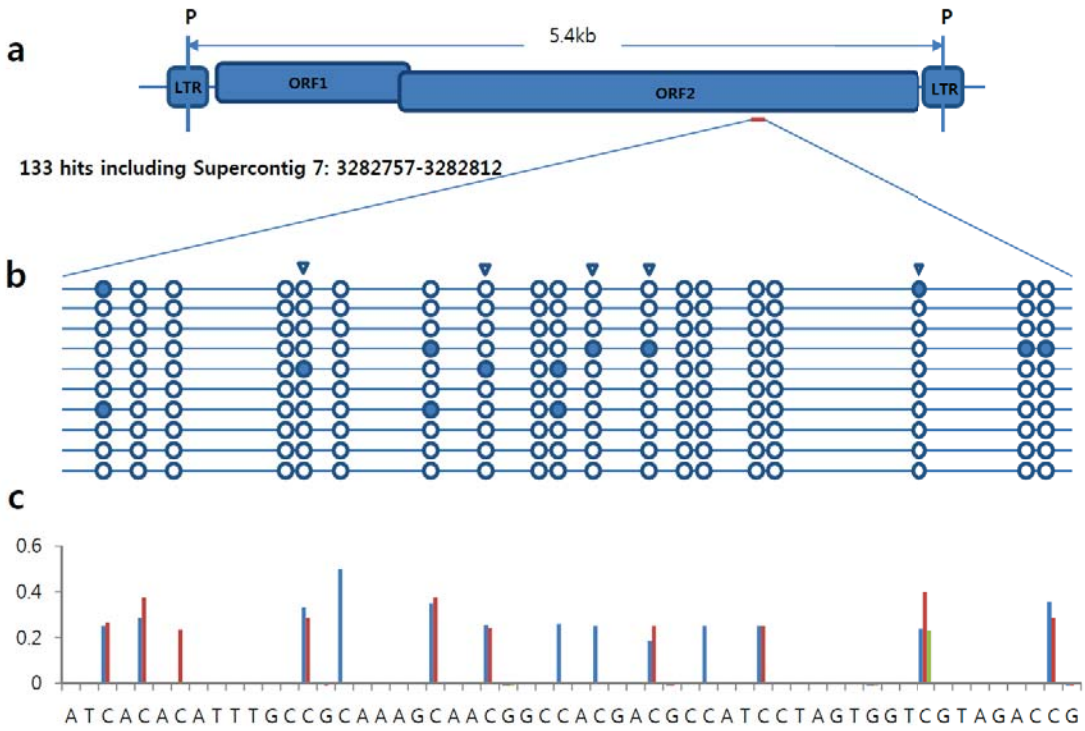


Figure 22. Comparison of bisulphite sequencing to methylC-seq data on a methylated region. **a**, schematic diagram of ORF, from which amplicon is derived. **b**, schematic summary of sequencing results. Each horizontal line represents independent sequencing read of a part of amplicon (red bar below ORF). Empty circle represents unmethylated cytosine and filled circle represents methylated cytosine within the amplicon. Triangle indicates position of cytosines in CG context. **c**, summary of methylC-seq data on the corresponding region. X-axis represents nucleotides on that region and y-axis represents methylation level of cytosine. All the data shown here is from wild-type mycelia.

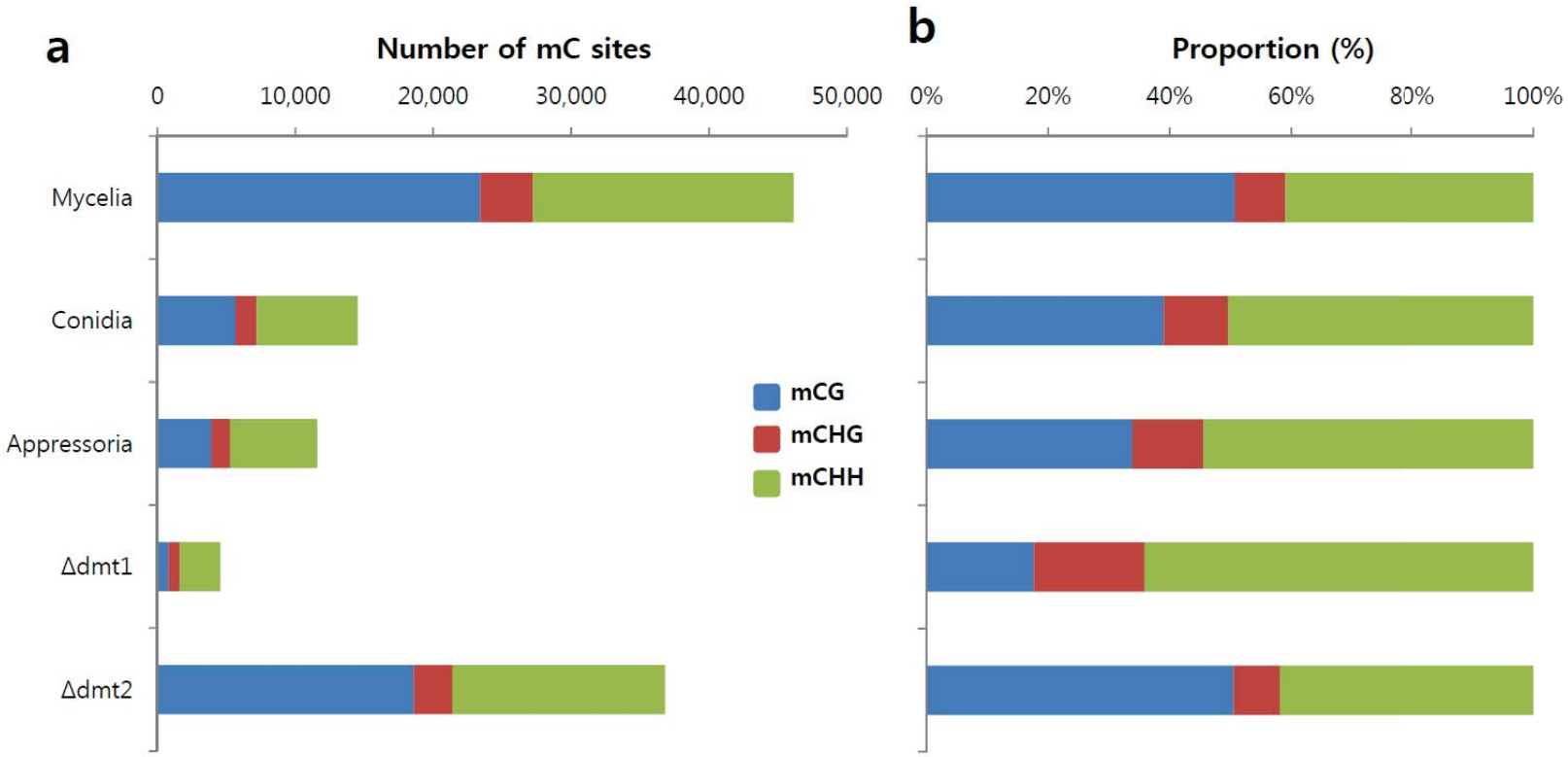


Figure 23. Amount and proportion of methylated cytosines. Each horizontal bar represents cumulative numbers (**a**) and proportions (**b**) of mCs in three different sequence contexts. H represent non-G base (H=A, T, or C).

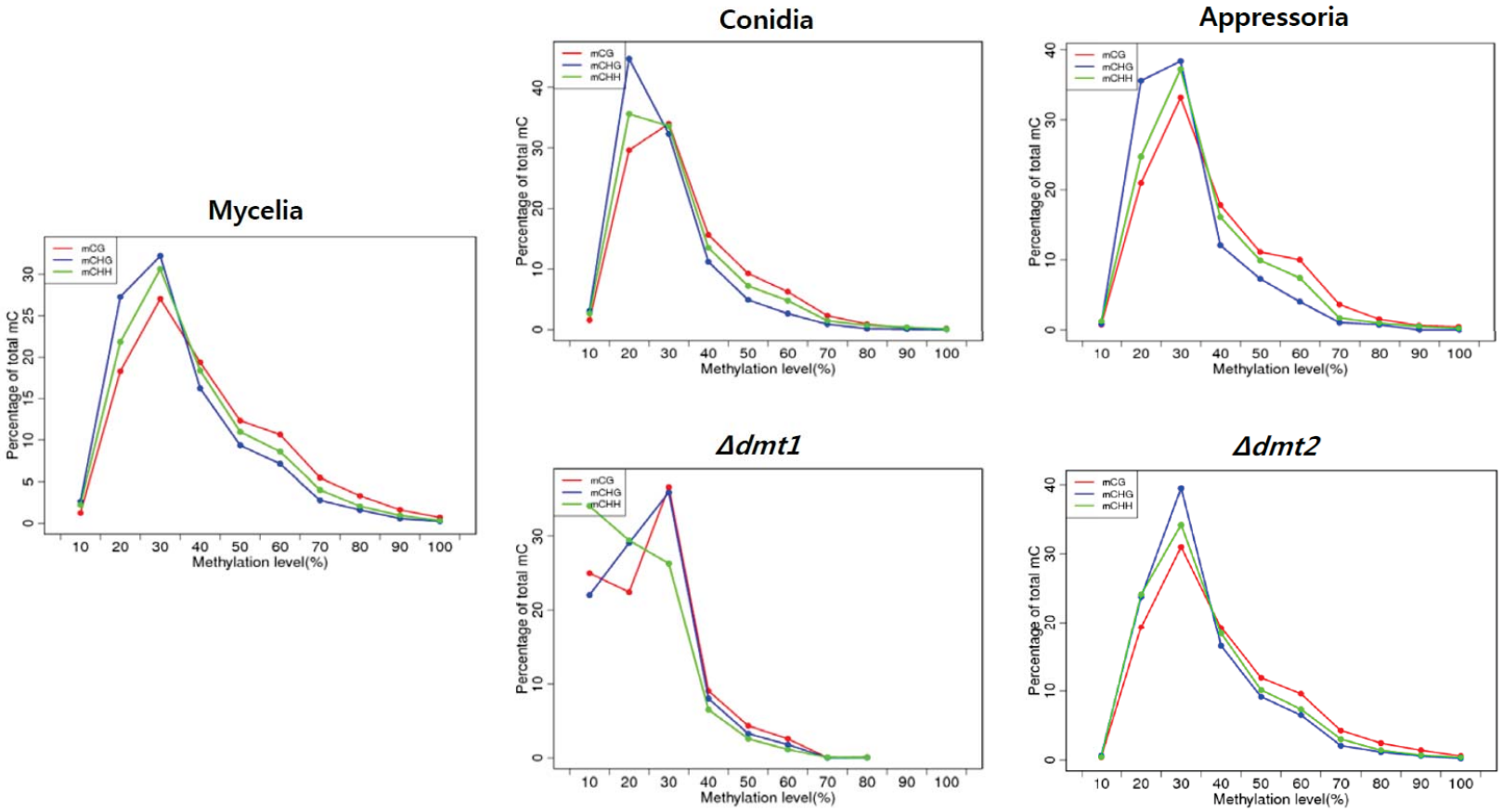
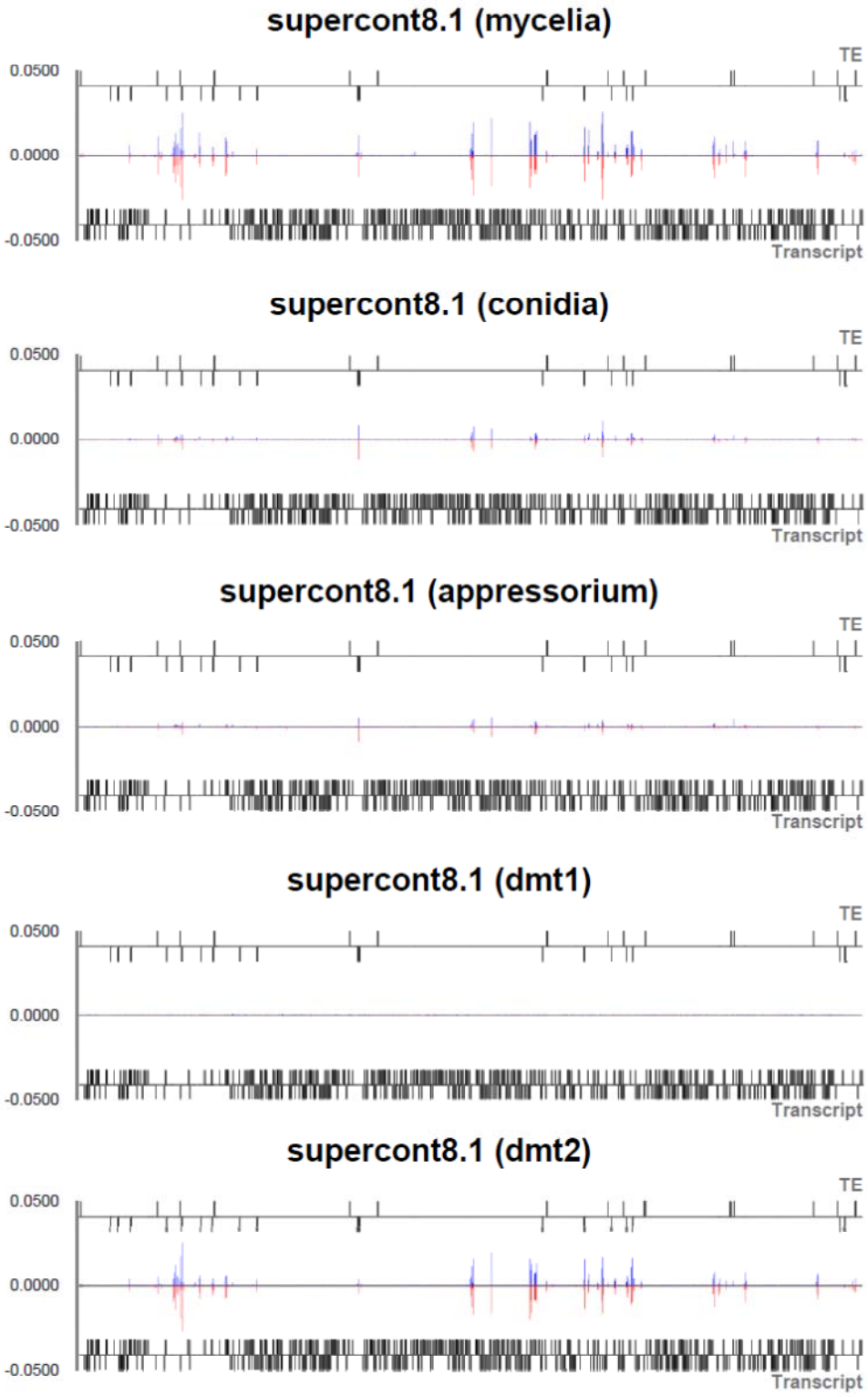


Figure 24. Distribution of methylation level for methylated cytosines in the context of CG, CHG, and CHH. X-axis represents methylation level of mC divided into 10 intervals. Y-axis represents percentage of mCs that fall into certain methylation level.

Figure 25. Chromosomal distribution of DNA methylation in *M. oryzae.* X-axis represents chromosome 1 of *M. oryzae* and y-axis represents density of mCs. Cell types and strains for which methylome was profiled are denoted in parenthesis.


supercont8.1 (mycelia)
TE
0.0500
0.0000
-0.0500
Transcript
supercont8.1 (conidia)
TE
0.0500
0.0000
-0.0500
Transcript
supercont8.1 (appressorium)
TE
0.0500
0.0000
-0.0500
Transcript
supercont8.1 (dmt1)
TE
0.0500
0.0000
-0.0500
Transcript
supercont8.1 (dmt2)
TE
0.0500
0.0000
-0.0500
Transcript

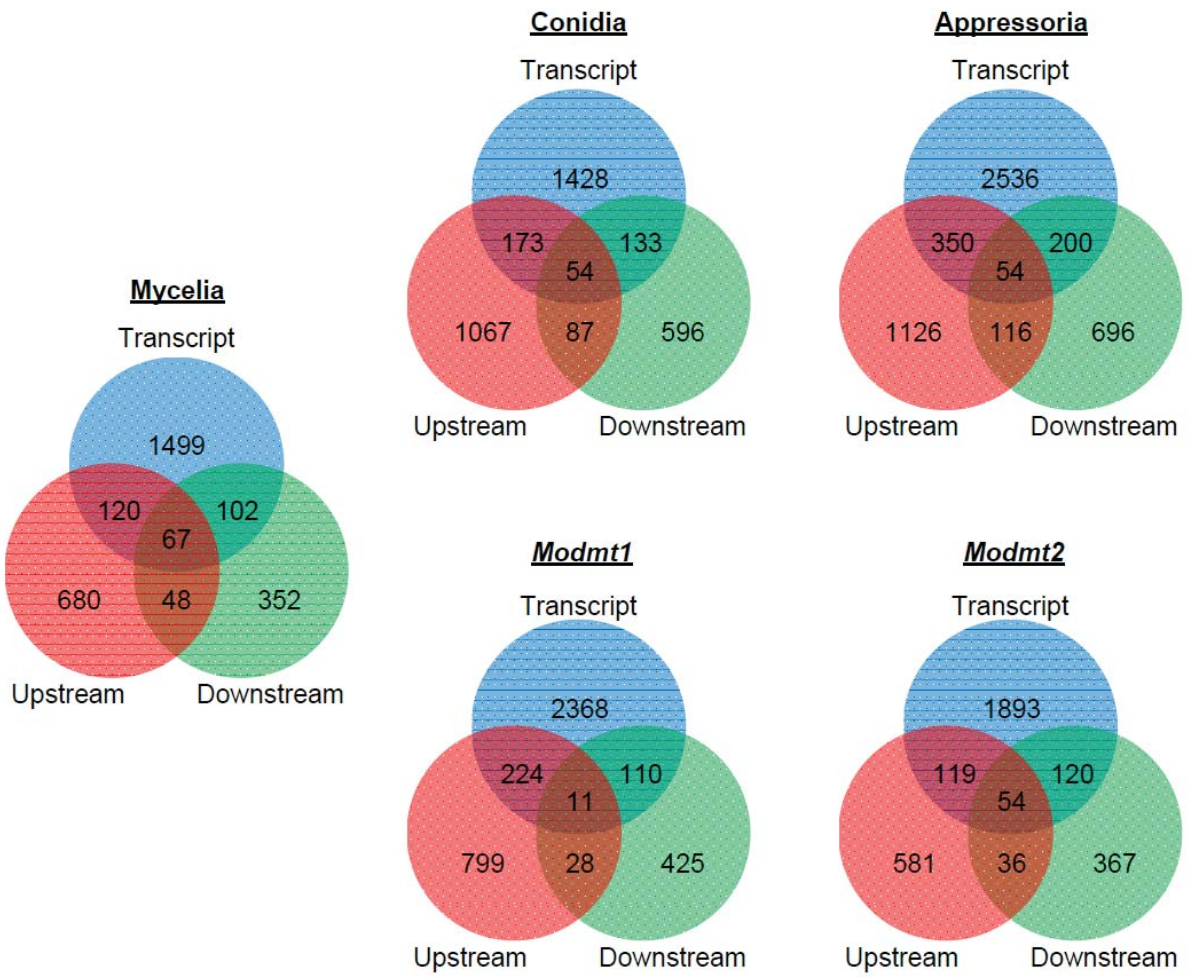
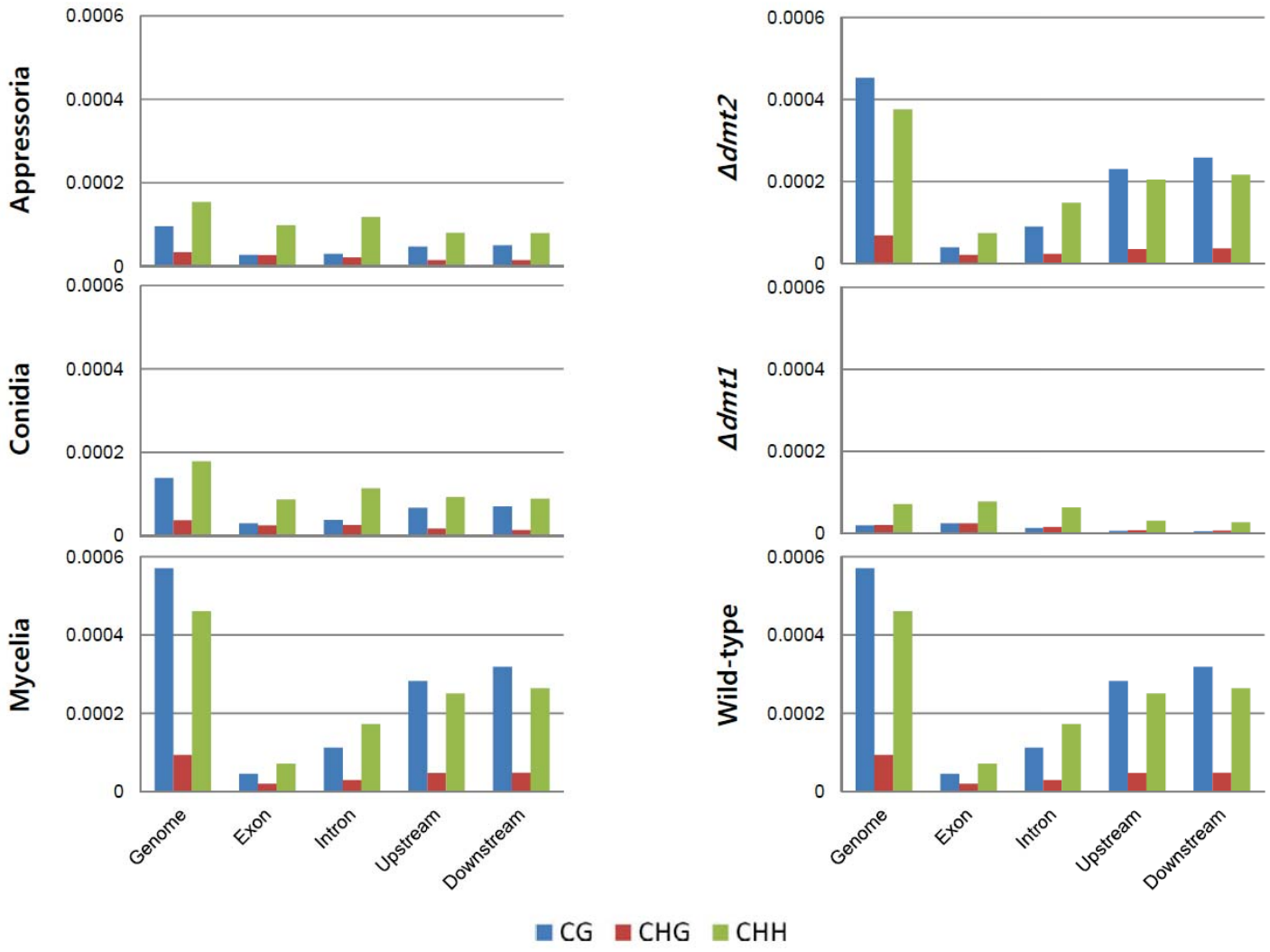


Figure 26. Comparison of number of mC sites within genic region in different samples.

Figure 27. Density of mC sites in different genome features. Density is calculated as number of mC sites divided by total length of each feature in the genome.


Appressoria
Conidia
Mycelia
Δdmt2
Δdmt1
Wild-type
0.0006
0.0004
0.0002
0
Genome
Exon
Intron
Upstream
Downstream
CG
CHG
CHH

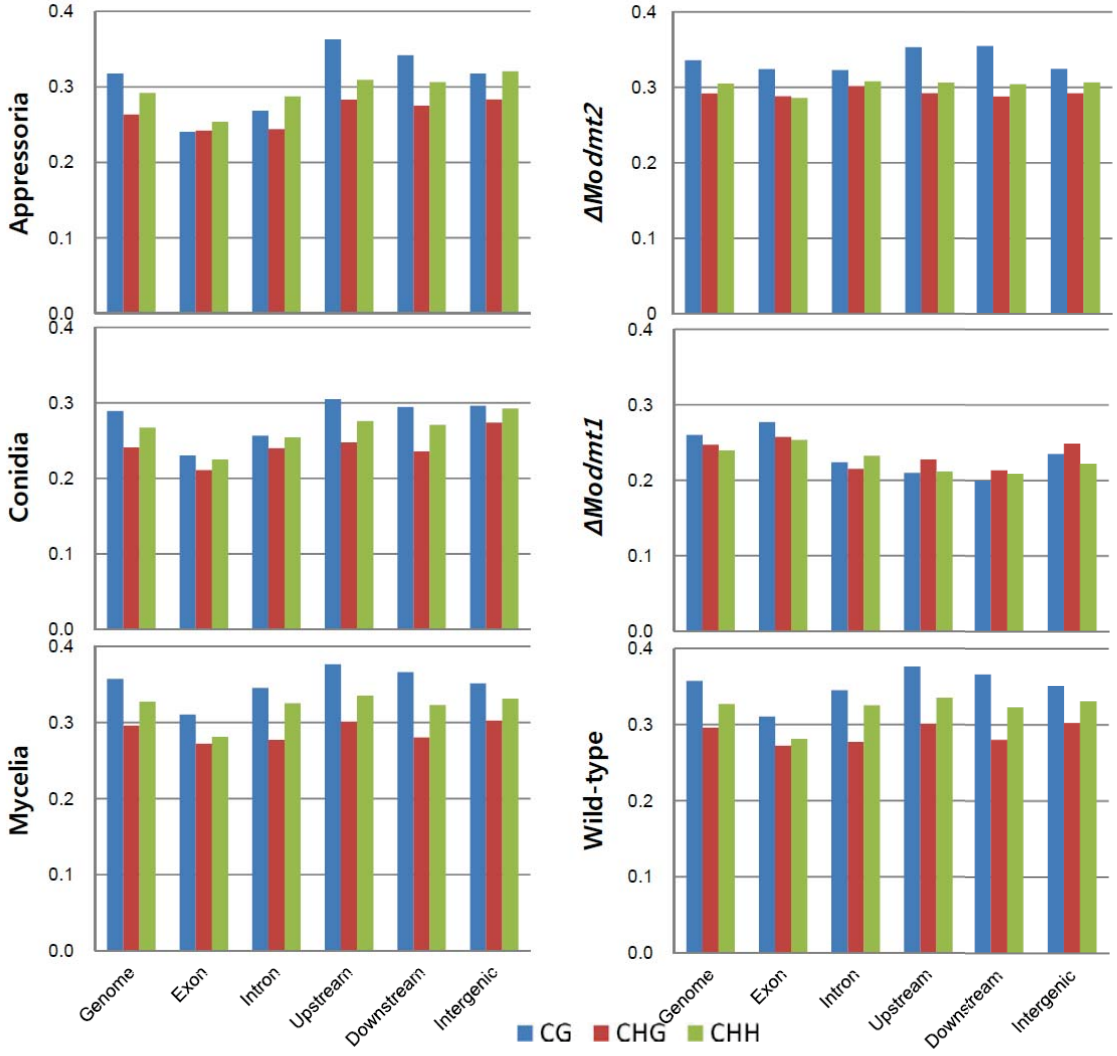
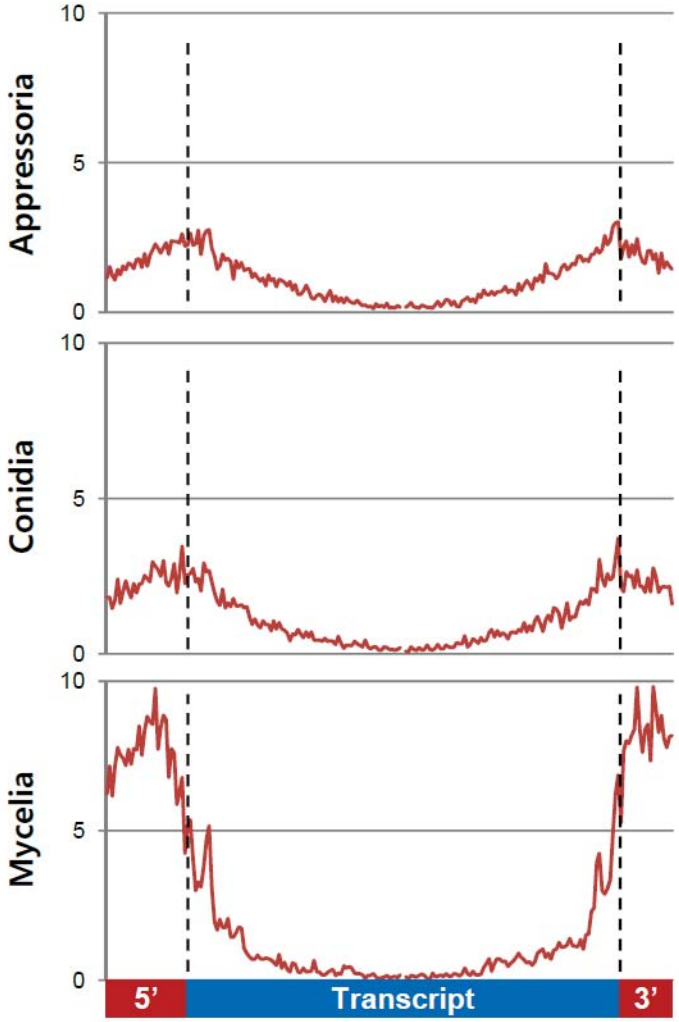


Figure 28. Average methylation level of mC sites in different genome features. Average methylation level is calculated as number of mC reads divided by total number of reads covering the cytosine site in each feature within the genome.

Figure 29. Density of mC sites within genic region. Genes were aligned at the 5′ end (left dashed line) or the 3′ end (right dashed line) and number of mCs in each 50bp interval was plotted across genic region.


Appressoria
Conidia
Mycelia
10
5
0
5'
Transcript
3'

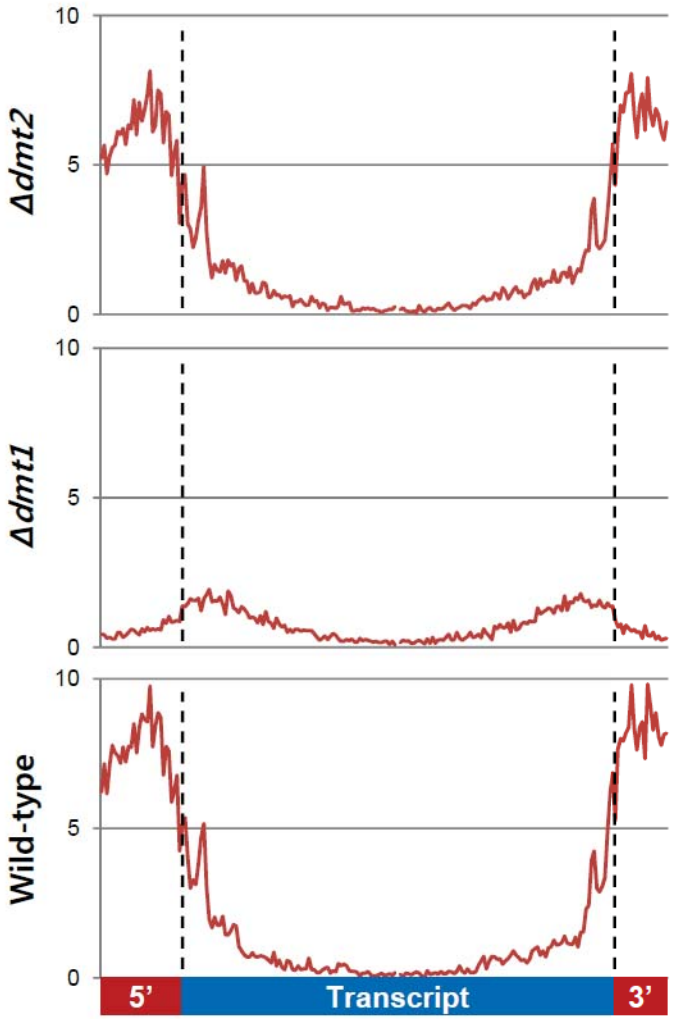

Δdmt2
Δdmt1
Wild-type
10
5
0
5'
Transcript
3'

Figure 30. Average methylation level of mC sites within genic region. Genes were aligned at the 5′ end (left dashed line) or the 3′ end (right dashed line) and mean of average methylation levels in each 50bp interval was plotted across genic region.

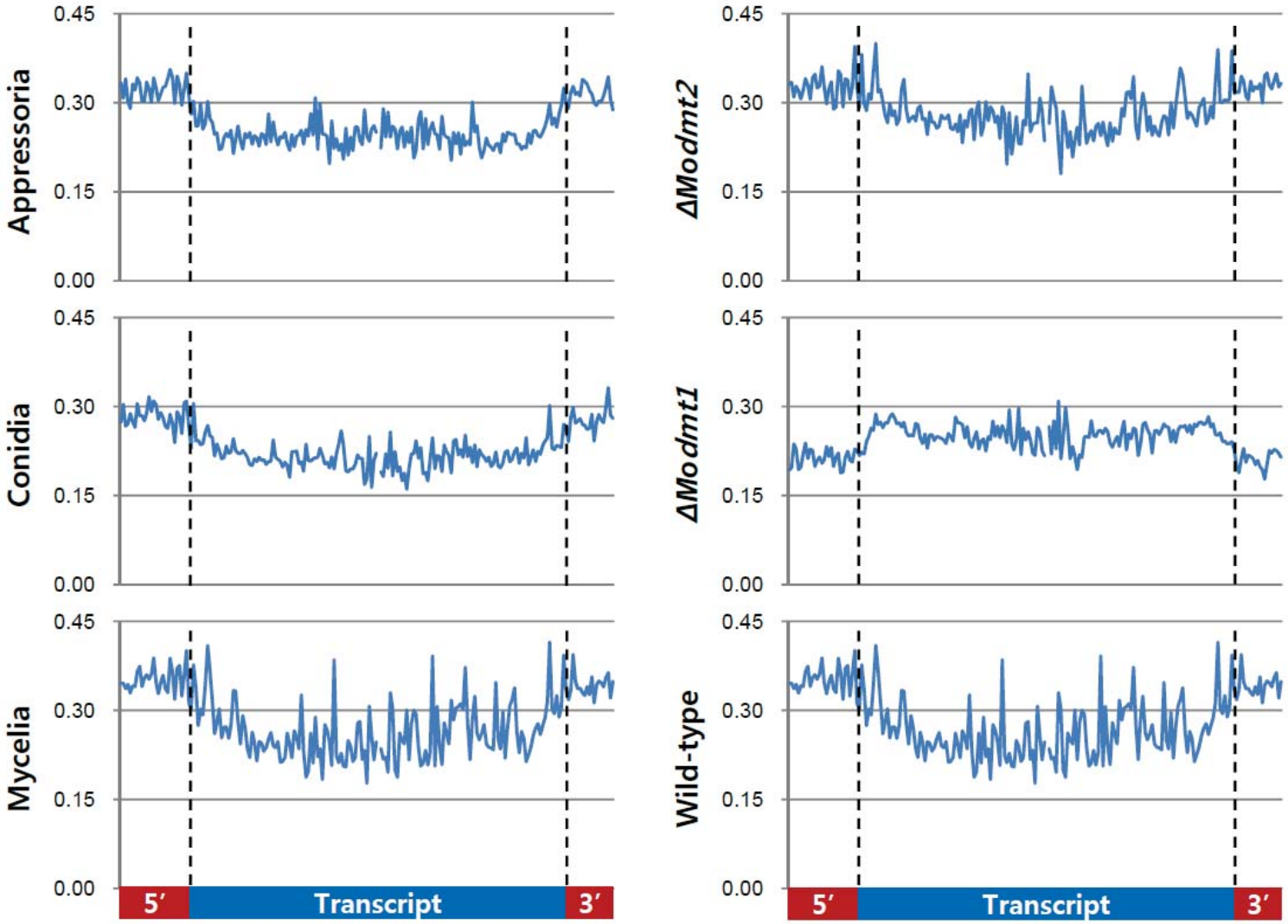
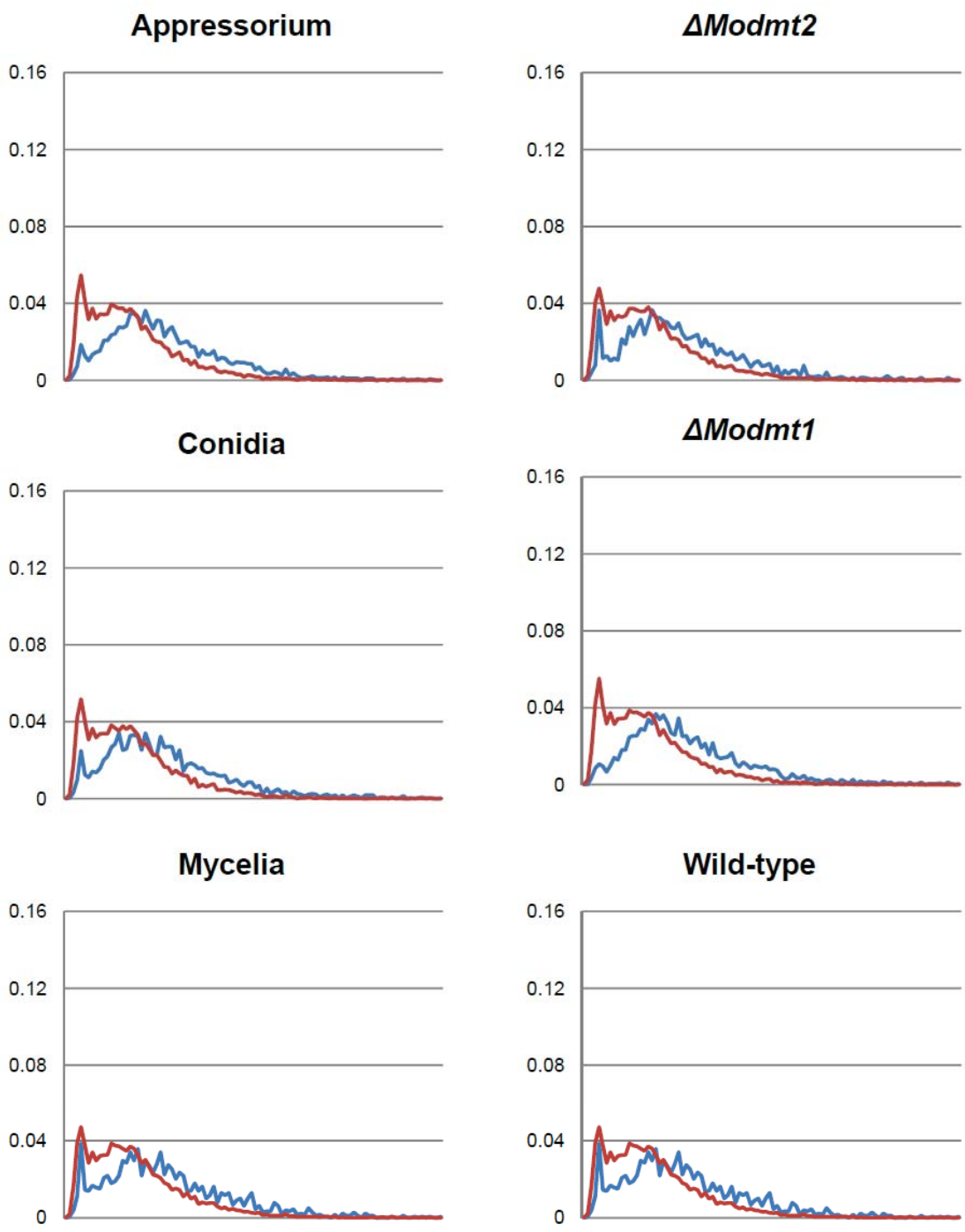

Appressoria
Conidia
Mycelia
ΔModmt2
ΔModmt1
Wild-type
0.45
0.30
0.15
0.00
5′
Transcript
3′

Figure 31. Length distribution of genes with (blue) and without (red) body methylation. X-axis and y-axis represent length of genes (up to 10kb) and fraction of genes haveing corresponding length.


Appressorium
ΔModmt2
Conidia
ΔModmt1
Mycelia
Wild-type
0.16
0.12
0.08
0.04
0

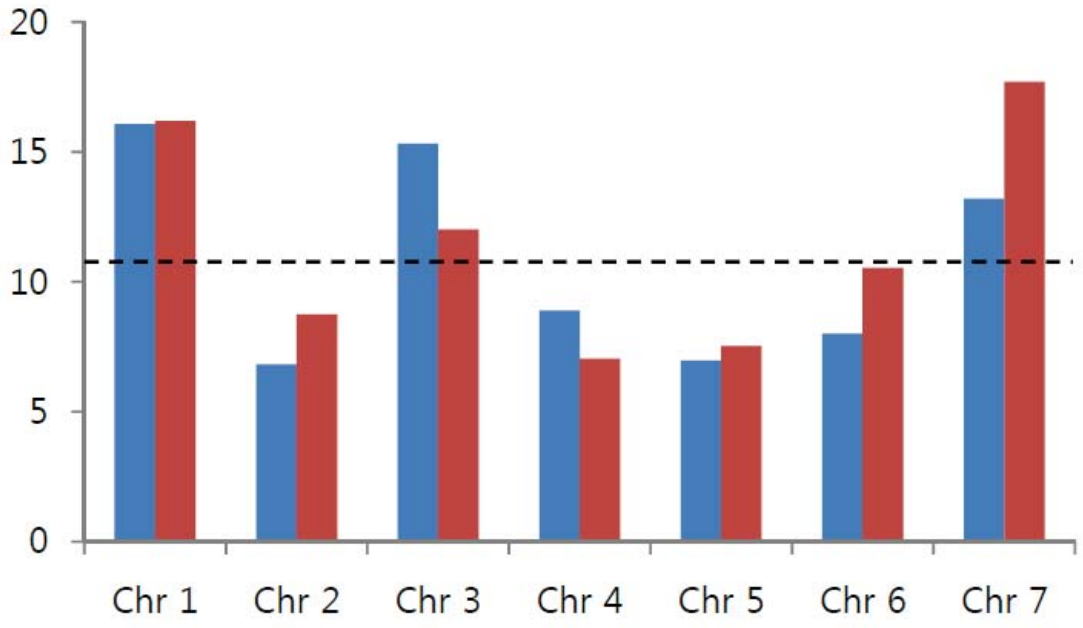


Figure 32. Density of mC sites and TEs over chromosomes. Blue bar represents # of mC sites per 10kb in each chromosome, and red bar represents density of TEs in each chromosome ((total length of TEs/length of corresponding chromosome)*1000). Dotted line indicates expected mC density across chromosomes. Unassigned genome sequences is not included.

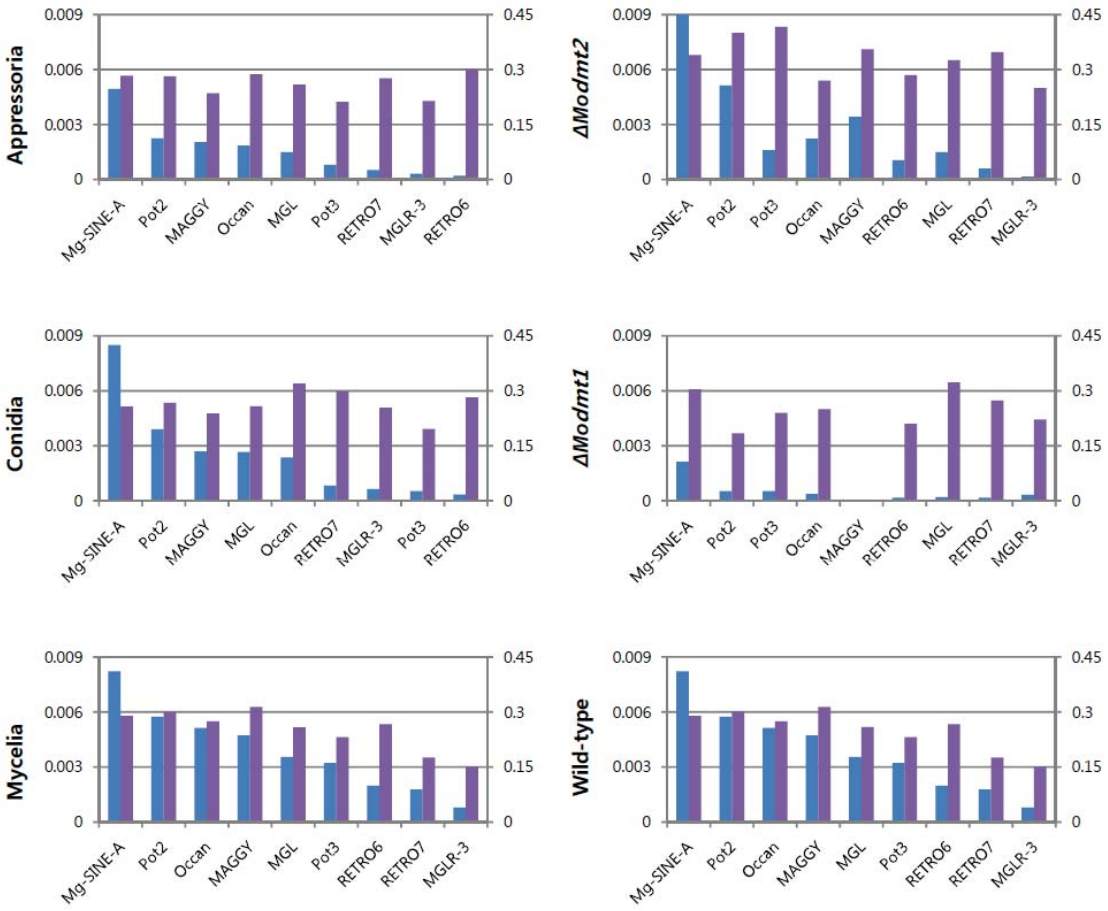


Figure 33. Density and average methylation level of mC sites in different types of TEs. TEs are arranged in ascending order of length.

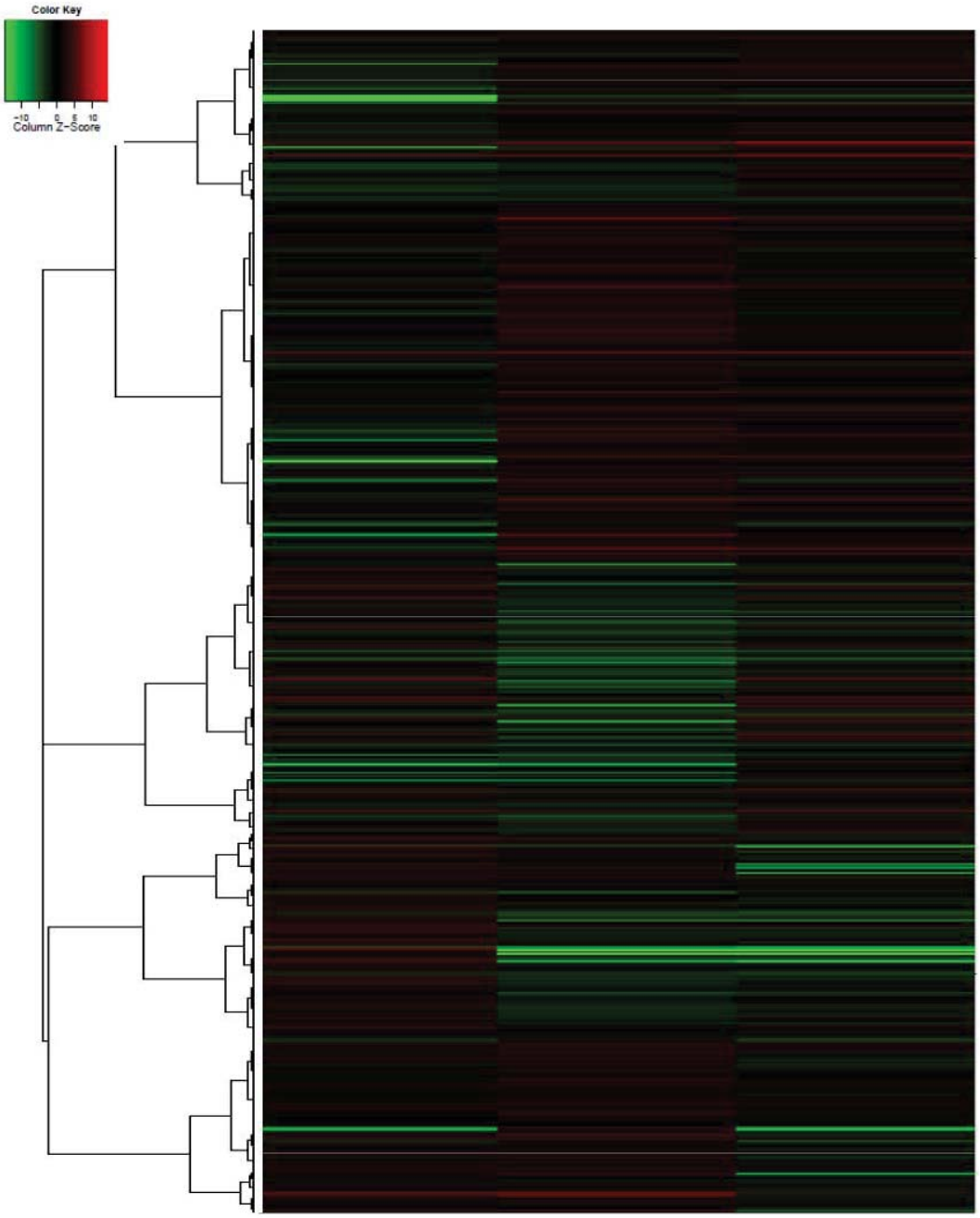


Figure 34. Heatmap of RNA-seq results. Transcript abundance of *Δdmt1*, *Δdmt2*, and *Δdmt1Δdmt2 (from left to right column of heatmeaps)* was shown relative to wild-type.

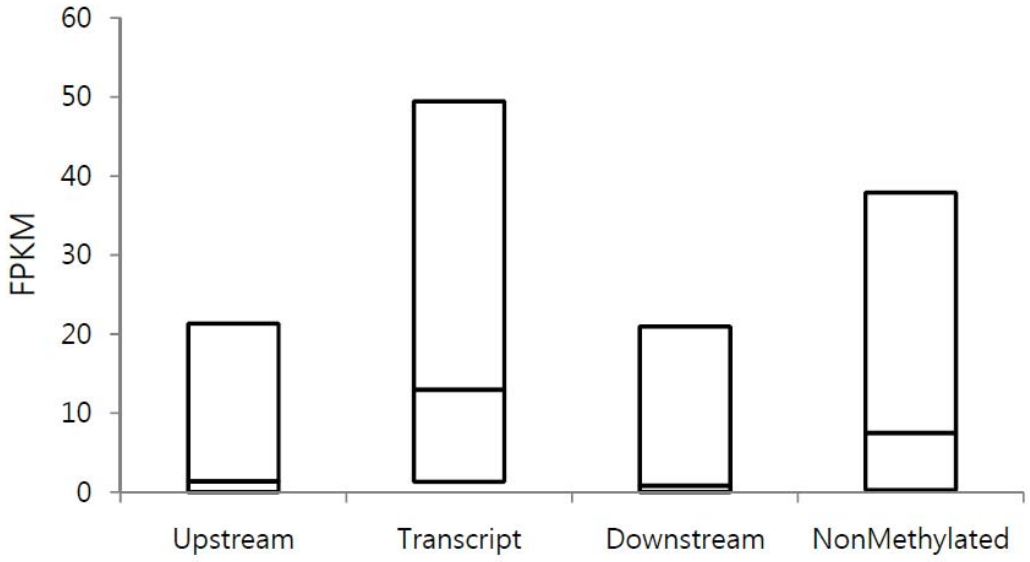


Figure 35. Boxplots showing expression of genes having DNA methylation in upstream, coding sequences, and downstream of genes. Expression of non-,methylated genes is shown for comparison. Expression was calcuated as FPKM (fragments per kilobase of transcript per million mapped fragments) (Y-axis). Only the lower quartile, median, and upper quartile were shown.

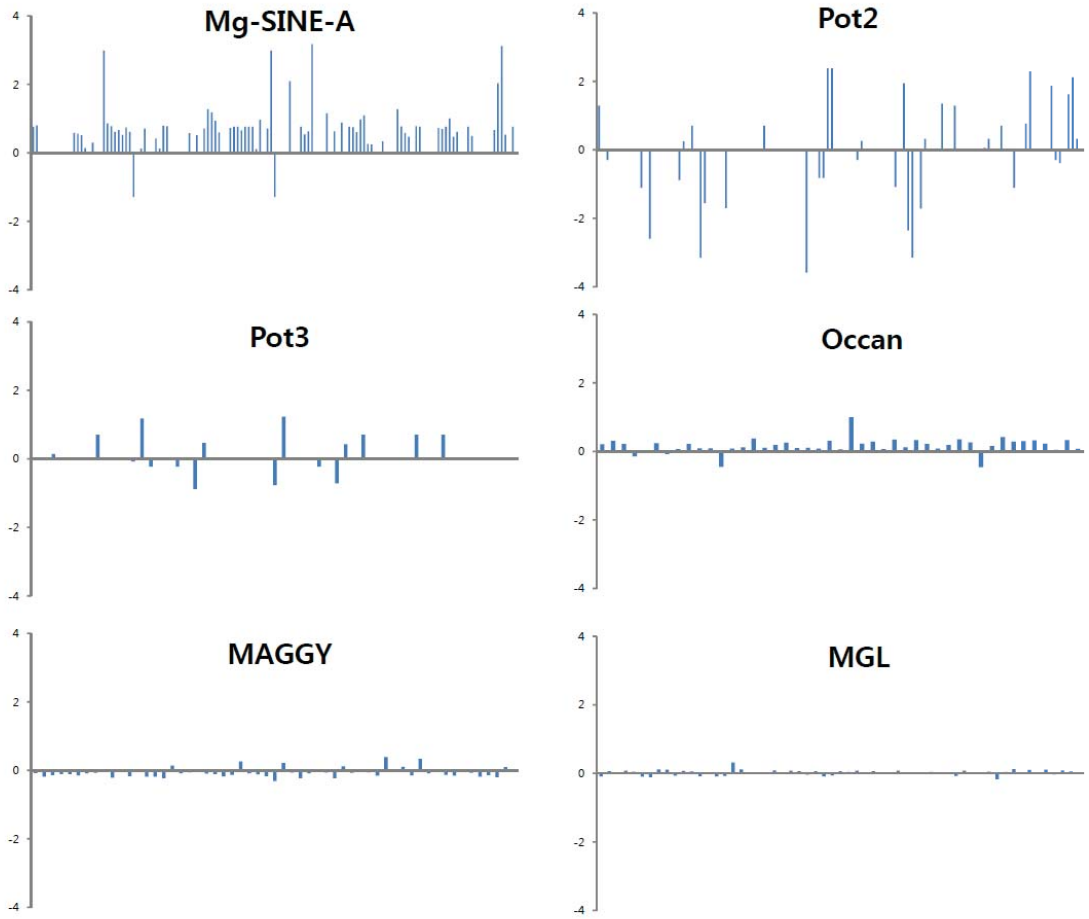


Figure 36. Relative expression of each type of TEs in *Δdmt1* to wild-type ($\log_2$(*Δdmt1*/wild-type)). Each bar represent individual transposable elements of the kind, located in different genomic positions.