Comparative Genome Analysis of Indian Isolates of *Magnaporthe*

SYNOPSIS

submitted to

The Maharaja Sayajirao University of Baroda

For the award of the Degree of

DOCTOR OF PHILOSOPHY

By

Mehta Khyati Natvarlal

Department of Microbiology and Biotechnology Centre

Faculty of Science

The Maharaja Sayajirao University of Baroda

Vadodara – 390 002 Gujarat

INDIA

October 2021

Comparative Genome Analysis of Indian Isolates of *Magnaporthe*

1. Introduction

1.1 Fungal pathogens and plant diseases

Fungal kingdom is very large comprising of about 1.5 million species. Around 15000 of these are known to cause plant diseases and mainly belong to Ascomycetes and Basidiomycetes divisions. Plant pathogenic fungi are classified into biotrophic, necrotrophic and hemibiotrophic, based on their pathogenic lifestyles. Biotrophic fungi derive nutrients from their host plant cells, while keeping them alive. Whereas necrotrophs kill their host plant cells to feed on them. Interestingly, hemibiotrophs use both the modes sequentially. Fungal pathogens attacking plants are considered as a major threat to global food security and modern agricultural ecosystem, since more than 125 million tons of staple food crops get destroyed by various fungi every year (Fisher *et al*, 2012); (Risks to Plant Health - Launch of EASAC Report 24 | EASAC - Science Advice for the Benefit of Europe). Top 10 fungal pathogens list, based on the scientific and agro-economic importance, include *Magnaporthe oryzae* that causes blast disease on major cereal crops, *Botrytis cinerea* that has a wider host range and ability to damage produce pre- and post- harvest, *Puccinia* spp. that causes rust disease on wheat, and *Fusarium* Spp. with a wide host range (Dean *et al*, 2012). Knowledge regarding pathogen's species, lineages, populations, extent of genetic variability and reproductive mode is important in order to understand the mechanism underlying disease emergence and spread (Taylor & Fisher, 2003)

1.2 Cereal blast disease caused by *Magnaporthe oryzae* - aka Cereal Killer

In this study, we have used a model ascomycete filamentous fungus *Magnaporthe oryzae* (syn *Pyricularia oryzae*) - the causal agent of the most devastating blast disease. It can infect virtually all the aerial part of the cereal host plant. Typical annual yield losses range from 10-30%, which can go up to 100% when the environmental conditions are conducive. Due to various reasons, *M. oryzae* is considered as a principal model system to study the molecular basis of host (plant)-pathogen interactions. While certain lineages of the blast fungus have been found to be adapted to a specific and narrow host range, the pathogen

can infect more than 50 different plant species and is capable of adapting to a new host with its tremendous ability to break the natural resistance therein. Thus, a new field strain can emerge with more potent virulence and multiple host-adapted lineages within species (Gladieux *et al*, 2018). This speciation process is largely influenced by the interaction between pathogen and its host-plants, and the emergence of a new strain poses a serious threat to crop production. Particularly, such a threat has been exemplified by recent emergence of wheat blast as a result of host-shift in the fungal pathogen *M. oryzae*. Many fungal effectors are believed to be involved in effective pathogenesis; however molecular signatures leading to emergence of a new within-species-lineage have not been studied in cereal-blast patho-biology.

1.3 Role of Effectors in virulence variability of the blast fungus

The ability of the blast pathogen to infect different host largely depends on the effector molecules, which help in manipulating the host defense to the pathogen's advantage, and may act as major determinants in specialization on a host (Sánchez-Vallet *et al*, 2018). However, in the due course of evolution, host cells may begin to recognise such a effector molecule and make the fungal virulence strategy ineffective. For example, some small secreted proteins (SSP) from pathogens act as avirulence (AVR) factors, as their recognition by host resistance (R) proteins makes host-pathogen interactions incompatible. Further, the AVR genes, either through a gain or loss of function, seem to play important roles in the pathotype evolution. Indeed, *Eleusine* (finger millet)-specific group likely evolved through a gain of *PWL1* gene, which led to loss of pathogenesis on weeping lovegrass (Asuke *et al*, 2020). Similarly, the *Triticum* pathotype likely emerged from the *Lolium* isolates via loss-of-function of avirulence gene *PWT3* upon wide-spread cultivation of wheat cultivar harboring the corresponding *RWT3* resistance gene (Inoue *et al*, 2017). Several fungal secondary metabolites (SMs) have been reported to be produced at the plant-fungus interface and a few of these non-proteinaceous molecules have been shown to play effector-like functions during host-pathogen interactions (Collemare *et al*, 2019). For example, the Ace1-derived SM product is likely required for full virulence on most cultivars of the host species except for those which carry the resistance gene *Pi33* (Böhnert *et al*, 2004; Collemare *et al*, 2008). Various fungal pathogens, including *M. oryzae*, are known to produce analogs of phytohormones such as

jasmonic acid, gibberellin, cytokinins and ethylene. Fungal analogs of JA and cytokinins are known to manipulate plant hormone-based defense signaling and subvert plant growth, respectively, during tissue colonization (Patkar & Naqvi, 2017; Shen *et al*, 2018). Interestingly, fungal JA derivative produced by *M. oryzae* is found to play a key role only during rice-blast fungus interaction (Patkar *et al*, 2015).

1.4 Genomics, a powerful tool, to study pathobiology of the blast fungus

Due to ever-increasing fungal genomes and advent in computational biology, it is nowpossible to study population structure at a large scale. The gaps between molecular biology, evolutionary genetics and epidemiology of a given disease have been bridged by revolution in the genomics-related tools. Major advances in comparative and population genomics have made possible to better understand the effector biology and its evolution (Plissonneau *et al*, 2017). Genomics approaches are also emerging as a gold standard to infer the genetic basis of adaptive divergence or to identify genomic determinants of host adaptation or host range expansion. The first reference genome draft of *M. oryzae* strain 70-15 was made available in public domain, through the Broad Institute, using sanger sequencing (Dean *et al*, 2005). Later, with the broad accessibility of next generation sequencing (NGS) technologies, whole-genome sequencing has become quite affordable for larger number of research groups. Currently, approximately 300 genomes of *M. oryzae* are available in public domain.

South-East Asia is believed to be the centre of origin of blast fungus and most migrations of *M. oryzae* populations is believed to have occurred from Asia to the other continents (Saleh *et al*, 2014). *M. oryzae* population structure is comprised of multiple diverse lineages in Asia, majorly in India, China and Thailand, as compared to simple clonal populations in the USA, Europe, Japan and Philippines (Skamnioti & Gurr, 2009). Genetic variations between the lab strain and field isolates of *M. oryzae* and among the rice and non-rice isolates have been reported based on comparative genomics studies (Dean *et al*, 2005; Xue *et al*, 2012; Chiapello *et al*, 2015). More genome analyses revealed that the gene gains/losses and/or chromosomal structural changes could be the major factors for changes in host specificity, and that these changes are largely caused by the activity of transposable elements (Yoshida *et al*, 2016; Peng *et al*, 2018; Bao *et al*, 2017). Based on

world-wide isolates of *M. oryzae* infecting distinct hosts, it was found that *M. oryzae* diversified into multiple host-specific lineages, and inter-lineage gene flow keeps them as a single species (Gladieux *et al*, 2018). Due to the fluidic nature of the fungal genome, such speciation is often derived as a result of significant changes in the pathogen's genome so that the new lineage could get adapted to a new host environment (Raffaele & Kamoun, 2012). A need to get adapted to a host seems to be a driving force behind change in the genome architecture giving rise to a population or species differentiation. In *M. oryzae* isolates, adaptation to rice host is likely associated with a small number of lineage-specific gene families, which include genes involved in effectors functions - secreted proteins of unknown functions and biosynthesis of SMs (Chiapello *et al*, 2015). A large amount of genomics information for the blast fungus has been accumulated in the past decade.

Despite the knowledge about various protein effectors involved in host adaptation, the role of secondary metabolites as effectors, likely to represent the metabolic strategies used by the pathogen to get adapted to a different host, is largely unexplored. Thus, in my doctoral studies, we used comparative genomics approach to study Indian field strains of *M. oryzae* isolated from different host plants. Here, we focused mainly on the biosynthetic diversity, especially pertaining to SM gene clusters, in Indian populations of *M. oryzae* in order to understand the genetic and molecular basis of host adaptability.

2. Objectives

1. Collection and Isolation of *Magnaporthe* from rice and non-rice hosts from various geographical regions across India
2. Diversity assessment of isolates using phenotypic and molecular markers
3. Whole genome sequencing and comparative genome analysis of the selected isolates

3. Results

3.1 Collection and Isolation of *Magnaporthe* from rice and non-rice hosts from various geographical regions across India

Blast disease is favoured by long periods of free moisture and high humidity caused by frequent and prolonged rain showers, and cool temperatures in the daytime that supports optimal sporulation of the blast fungus. Initial symptoms on leaf appear as white to grey-green lesions or spots with dark green borders, which eventually develop into spindle-shaped necrotic lesions with whitish to grey centre and red to brownish borders.

Infected tissues with blast disease were collected from various hosts like rice (*Oryza sativa*), finger millet (*Eleusine coracana*) and foxtail millet (*Setaria italica*). *M. oryzae* infects almost all the aerial parts of the plants, leading to leaf blast, neck and panicle rot, collar rot and node blast. Samples were collected from different infected plant parts such as leaf, node, neck, panicles (in case of rice) and fingers (in case of millets). These blast-infected tissues were collected from 15 different geographical locations, during rainy seasons, across India. Sampling sites included some of the blast hotspots such as Sikkim, Karnataka and the Dangs region of Gujarat. A pure fungal culture is fundamental to the identification of taxa as well as phenotypic characterisation such as culture morphology, mating type, degree of virulence and host specificity. Therefore, all the collected samples were subjected to single-spore isolation technique to obtain the pure fungal cultures. Total number of *M. oryzae* isolates thus collected from rice and non-rice hosts across India is eighty. A fungal culture repository was established by making filter paper stocks for all the isolates. All the *M. oryzae* isolates were grown for ~10 days on oat meal agar plates containing filter paper discs. The filter discs covered with fungal biomass were collected and appropriately desiccated before placing in sterile small envelopes. These envelopes were stored in polybags at -20 or -80 °C for short- or long-term storage, respectively.

3.2 Diversity assessment of isolates using phenotypic and molecular markers

3.2.1 Molecular characterization of the isolates

To determine the genetic structure of fungal populations, we used Pot2 and Mg-SINE Rep-PCR fingerprinting approach. There are various repetitive elements present in *M. oryzae* genome; however Pot2 and Mg-SINE are found to be in higher (~100 copies per haploid genome) in both rice and non-rice isolates (Kachroo *et al*, 1995). Further, Pot2 Rep-PCR fingerprinting has been found to be reliable, reproducible, and highly discriminatory for assessing diversity in large collections of *M. oryzae* strains (Kachroo *et al*, 1994). A similar

pattern of Pot2 and Mg-SINE in a given population is a strong indication of common ancestry of isolates. Total DNA was extracted from our isolated s*M. oryzae* strains using the Dellaporta method with minor changes (Dellaporta *et al*, 1983). After spectrophotometric quantification of genomic DNA, it was used for the Pot2 and Mg-SINE Rep-PCR. The Rep-PCR results showed different Pot2 and Mg-SINE fingerprinting patterns indicating a genetic variability among various isolates of the blast fungus.

Avirulence genes in *M. oryzae* are recognised specifically by the corresponding resistance (*R*) genes of its host species. So far, about fourteen *AVR* genes have been identified from the blast fungal pathogen. Given that the *AVR* genes are diverse and that the effector repertoire of pathogen is associated with its virulence, it is worth studying diversity of the *AVR* genes in various blast fungal isolates sampled from different host species. We used a total of six *AVR* genes to check their presence/absence polymorphisms using PCR with gene-specific primers in the current study. These include *ACE1, PWL2, AVR-Pita, AVR1-CO39, Avr-Pizt* and *Avr-Pik* genes. The *AVR-Pita* gene, located in the telomeric region, encodes a putative zinc metallopeptidase. Its corresponding *R* gene is *Pi-ta*. This gene *AVR-Pita* is highly diverse and confers variability in the virulence via rapid changes such as deletions, transposon insertions, and point mutations. *AVR-Pita* has been found in various *M. oryzae* isolates from wheat, millets and other related species besides rice. In our study, we found that *AVR-Pita* gene was present in all the foxtail millet isolates and a few of rice isolates, but absent from finger millet isolates. The *Magnaporthe* isolates carrying the Avirulence Conferring Enzyme1 (*ACE1*) gene are specifically recognised by the rice cultivars carrying the *R* gene *Pi33*. *ACE1* encodes a putative hybrid polyketide synthase non-ribosomal peptide synthetase, which is involved in secondary metabolism in microbes. It is expressed at the onset of host penetration and is reported to be localised in the cytoplasm of appressorium. Interestingly, the recognition by Pi33 does not involve the Ace1 protein itself but the secondary metabolite synthesized by it. We found that the *ACE1* gene is present in all the isolates in the current study, suggesting a role in pathogen's overall fitness and virulence. The *AVR1-CO39* gene confers avirulence towards rice cultivar carrying *Pi-CO39* resistance gene. This gene has come from a *Magnaporthe grisea* isolate infecting weeping love grass, and is not present in the rice-infecting *M. oryzae* isolates collected from Brazil, China, Japan, India, Indonesia, Mali and Philippines (Farman, 2002). We found similar results, where *AVR1-CO39* was present in all the non-rice isolates and

absent in almost all of the rice isolates. Intriguingly, *AVR1-CO39* was present in isolates OSBC1 and OS-KK that were isolated from rice plants, which were cultivated next to millet fields. Similarly interesting is the case of the *PWL* gene family that includes genes such as *PWL1*, *PWL2*, *PWL3* and *PWL4*. One or more *PWL* genes are present in rice isolates, and while they do not function as avirulence factors toward any of the rice cultivars known, they are recognized by weeping lovegrass (*Eragrostis curvula*). Interestingly, rice isolates, upon loss of the *PWL* genes, have gained the ability to infect weeping lovegrass successfully. Thus, the *PWL* genes appear to act as avirulence factors, mainly at the host species level but not at the cultivar level (Ebbole, 2007). The *PWL2* gene was not found in any of the non-rice isolates used in our study and this is consistent with earlier findings.

The *AVR-Pizt* avirulence gene can induce immune response if recognised by the resistance gene *Piz-t*. We found that *AVR-Pizt* was present in most of the isolates and did not show any particular pattern with respect to host specificity. The *AVR-Pik* gene was identified from field isolate Ina168 after genome sequencing, along with two other novel *AVR* genes - *AVR-Pia* and *AVR-Pii* (Yoshida *et al*, 2009). The corresponding *R* gene in rice is the *Pik* gene. In our study, *AVR-Pik* was found in most of the rice isolates, and absent from all the non-rice isolates. It appears that the selection pressure from different R genes carried by the host plants led to variability in terms of presence/absence of AVR genes among different isolates.

M. oryzae is a heterothallic fungus, with two distinct mating types (MAT1-1 and MAT1-2) and sexual reproduction being possible only between the two opposite mating types. This fungus shows a high degree of pathogenic variation (pathotypes) in the field conditions. Sexual reproduction can enhance and influence the genetic variability of *M. oryzae* populations, wherein the recombinant progenies may have been conferred with new capabilities to infect different host species and/or cultivars. Both MAT 1-1 and MAT 1-2 are idiomorphs of *MAT* – a single mating type locus. Assessment of mating-type alleles has been used as a marker to measure population diversity in this pathogen. We found both kind of mating types in our isolates and it needs to be further studied if the mating type or sexual reproduction, if any, has a role in host adaptability.

Based on the molecular characterisation (presence/absence of *AVR* genes and mating type loci) of different isolates, a dendrogram was generated using a distance-based method –

UPGMA (Unweighted Pair Group Method with Arithmetic mean). There was a high degree of variability among both rice and non-rice isolates, which associated with the genetic background of the host. Based on this, we were able to group the isolates based on the genetic diversity assessed using these molecular markers.

3.2.2 Phenotypic characterization of isolates

Phenotypic characterization of the isolates included the culture morphology, pigmentation, sporulation and ability to infect different host plants. As expected, the colonies of different isolates obtained upon single spore isolation were different in terms of mycelial growth, mycelial morphology and pigmentation. Interestingly, in some instances, colonies obtained from the same lesion showed distinct characteristics. All the isolates were able to produce conidia at varying capacities.

Representative isolates, from the aforementioned groups, were subjected to cross-infectivity assays, to check their virulence on different plant species other than the host of origin. Here, leaf sheath assay was carried out using four different hosts – rice, finger millet, wheat and barley. Invasive hyphae of *M. oryzae*, without any sign of host defense response, in the epidermal layer of the sheath tissue was considered as an indication of ability to successfully infect the host. We found that not all the isolates were able to invade all four hosts. Further, we did not find any obvious pattern among the isolates with respect to the ability to infect different hosts. Thus, it was decided to study cross-infectivity using whole-planta infection assay.

In order to study the reproduction mode of field isolates of *M. oryzae* in this work, we determined the mating type of these isolates, and carried out mating assays with a few of them. Crosses were made by pairing strains of opposite mating types with high-fertile standard tester strains. The fertility of *M. oryzae* isolates was assessed by their capacity to form perithecia, asci and ascospores in each cross, and isolates unable to produce perithecia were classified as sterile. Six different isolates from western India were crossed with opposite-mating-type tester strains so far. Most of the strains that we tested for fertility were highly male-fertile and were able to produce perithecia; however, we did not find any female-fertile strain. The future experiments include testing more isolates to find any female-fertile isolates and to assess whether they can mate with the male-fertile field isolates.

3.3 Whole-genome sequencing and comparative analyses of the selected isolates

The fungal strains have now been classified into distinct groups on the basis of DNA fingerprinting, their geographic locations and the source host plant species. Based on the classification, we identified 15 representative strains collected across India (Fig. 1A) and sequenced them for whole genome, using the Illumina HiSeq2500 platform. The *De novo* sequence assemblies were then constructed using CLC Genomics Workbench 11.0 and default parameters, after preprocessing raw reads using Trimmomatic v0.38 at a threshold for the minimum read length of 80 bp for paired-end reads (Bolger *et al*, 2014). The genomic regions with repeat elements were masked using the RepeatMasker version open-4.0 (Smit *et al*) and "*Magnaporthe oryzae*" as the species model. We used MAKER2 pipeline (Campbell *et al*, 2014; Holt & Yandell, 2011), where evidence-based and *ab initio* (Augustus and SNAP) predictions were combined to predict genes. We used PyParanoid pipeline (Melnyk *et al*, 2019) to infer the orthologs, and identified a total of 9169 gene families shared by all the 15 isolates from three different host plant species. Whereas, 257, 374 and 994 gene families were found to be unique to the isolates from rice, finger millet and foxtail millet hosts, respectively (Fig. 1B). These unique genes might serve as diagnostic markers for the blast populations.

We mapped the raw reads to the reference genome of *M. oryzae* 70-15 and used SAMtools (Li *et al*, 2009) for variant calling analysis and to identify SNPs therein. Our Principal Component Analysis (PCA) based on 544,642 SNPs obtained from all sequenced isolates, revealed three major clusters of the blast fungus corresponding to the aforementioned three different source host plants (Fig. 1C). To gain insight into the alteration in the fungal genomic architecture resulting from host specialization, we constructed a phylogenetic tree based on whole-genome SNPs dataset using MEGA X (Fig. 1D; Kumar *et al*, 2018). The phylogenomic analysis thus far is consistent with the finding that the selection pressure from the host likely plays a major role towards host-specificity of the pathogen (Zhong *et al*, 2016).

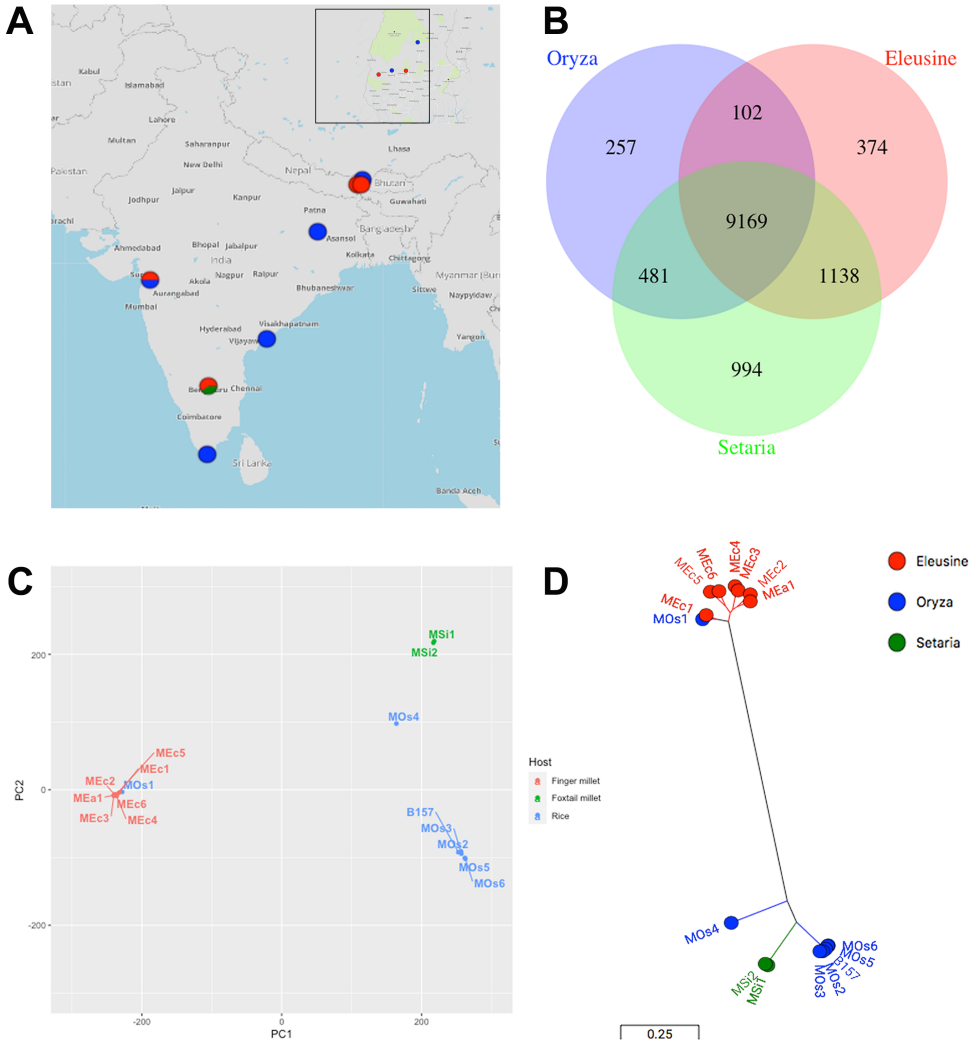

Figure 1. Comparative genomic analyses of *M. oryzae* isolates collected from different host plants. (A) Geographic region from where blast lesion samples were collected to isolate various *Magnaporthe* field strains. (B) Venn diagram showing number of orthologous genes in 15 representative *Magnaporthe* isolates. Os, *Oryza sativa*; Ec, *Eleusine coracana*; Ea, *Eleusine africana*; Si, *Setaria italica*. (C) Principal Component Analysis based on whole-genome SNP datasets of the aforementioned 15 representative isolates. (D) Phylogenomic tree based on whole genome datasets of Oryza, Eleusine and Setaria isolates of *Magnaporthe*.

Virulence of different host-specific population of *M. oryzae* largely depends on its effector repertoire, which includes small secretory proteins and/or chemical effectors such as secondary metabolites. We used SignalP and TMHMM on proteomes to identify the secretory protein effector-coding genes in the genomes. Homologs of known blast effectors, including *AVR* genes, were identified using BlastP on custom database

comprising of secretory proteins. Heatmap was generated depicting the presence and/or absence variations of known blast effectors in a given set of genomes. Some of the effectors like BAS2, BAS4, ACE1, AVR-Pi9 and AVR-Pi54 were present in all the genomes, whereas PWL was absent in most of the genomes representing *Oryza* lineage. Thus, variation in effectors likely underlies the differential virulence of various isolates from different host species.

Chemical effectors, specifically secondary metabolites, are usually synthesized from multi-step biosynthetic pathways. Genes involved in secondary metabolic biosynthetic pathways are usually located together at a particular genomic locus and constitute a Biosynthetic Gene Cluster (BGC). We predicted such BGCs in all the genomes under study using fungiSMASH. We included more genomes from public domain to make the study statistically significant. We found that ~80% of BGCs are novel/uncharacterised , and lack any similarities to the known BGCs from any other organisms. In a few BGCs, we found differences in the gene content either due to gene gain and/or loss of tailoring genes, likely affecting the modification of the core product and ultimately giving rise to a different end product. Characterisation of such diverse genes and their corresponding products might give insight into metabolic strategies of the blast fungus in host adaptation.

We also checked the structural changes in BGCs between two isolates 70-15 and MZ5-1-6, representing *Oryza* and *Eleusine* lineages, respectively. Pairwise synteny among the BGCs was identified in order to understand the genomic rearrangements, if any. Homologs of SM genes from both the genomes were mapped using Circos plot (Fig. 2). While most of the SM genes were highly syntenic, a total of 36 SM gene pairs located onto different chromosomes in the two genomes. Our finding further suggests that about 48 SM related genes of MZ5-1-6 have orthologs outside the genomic regions of predicted BGCs in 70-15. This kind of rearrangement might affect the final SM product and have an evolutionary relationship leading to lineage specificity.

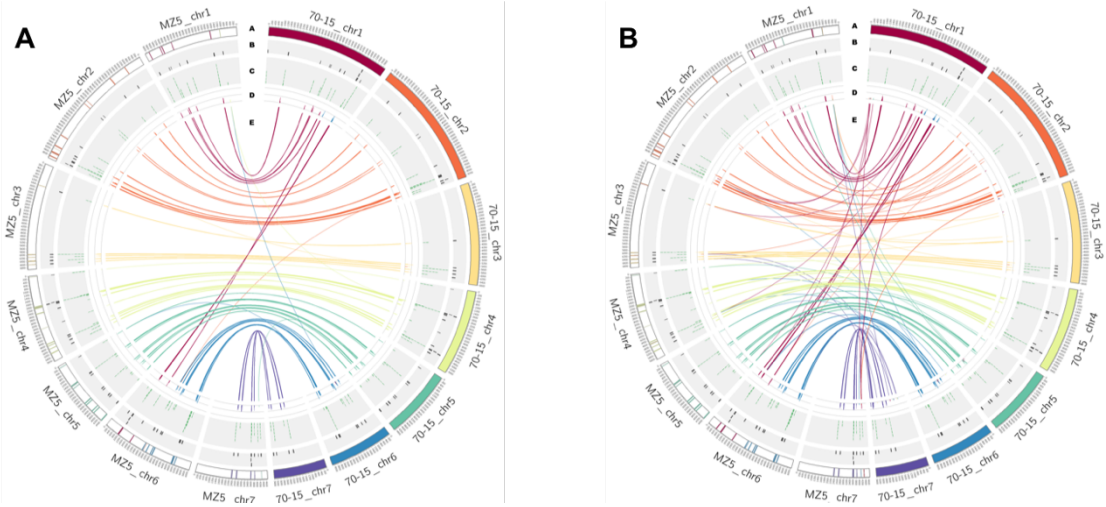

Figure 2. Circos plots showing orthologous gene pairs represented by connecting links between (A) SM-genes of both MZ5-1-6 and 70-15 and (B) SM-genes of MZ5-1-6 and total genes of 70-15. In each plot, individual tracks (A to E) denote different characteristics – Track A, chromosomes scaled according to their length; Track B, candidate clusters positioned accordingly on the chromosomes; Track C, core biosynthetic genes with respect to their positions on the chromosomes; Track D, histogram showing orthologous gene-pairs between the two genomes, with color codes according to those of the chromosomes of 70-15; and Track E, links representing orthologous gene-pairs between the two genomes, with color codes same as those for the 70-15 chromosomes.

4 Ongoing work & future direction

Currently, the biosynthetic diversity is being further explored to assess whether SMs in *M. oryzae* have an important role in host adaptation. In order to identify BGCs with possibly similar SM products, all the BGCs predicted from different isolates, belonging to host-specific lineages, are being grouped into Gene Cluster Families (GCFs), using the similarity network analysis tool BiG-SCAPE. Further, the phylogeny of core biosynthetic genes is being constructed for SM genes, associated with different classes of the chemical molecules, such as those encoding polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS) etc. Lastly, a large-scale phylogenetic analysis will be performed across the fungal kingdom to determine the evolutionary history of unique host-specific BGC.

5 Summary

In the current study, we collected 80 *M. oryzae* field isolates from rice, finger millet and foxtail millet cultivated in different geographic locations in India. We characterized them using phenotypic and molecular markers, and the representative 15 isolates were sequenced for whole genomes using Illumina NGS technique. *De novo* assembly and gene predictions were performed on repeat masked genomes. In order to understand genetic diversity among these 15 sequenced genomes, principal component analysis (PCA) was performed on SNPs datasets, where three major clusters were identified correlating with the aforementioned three host plants. Phylogenomic study confirmed the divergence of isolates into three host-specific lineages. Gene families identification based on orthologous gene sets revealed the unique gene sets specific to isolates belonging to host-specific lineages. Further, we found that ~80% of the BGCs are uncharacterised and do not have any similarities to the known clusters from any other organisms. In a few isolates, we found differences – through gene gain and/or loss - in the gene content of the BGCs, which correlated with isolates belonging to different host-species. Our finding thus far suggests that the differential genomic rearrangement, especially with respect to the genes involved in secondary metabolism, might have some evolutionary role in host adaptation. Overall, this study would help understand the strategy, especially pertaining to chemical effectors, used by the blast fungus to adapt to a new host.

References:

Asuke S, Tanaka M, Hyon G, Inoue Y, Vy TTP, Niwamoto D, Nakayashiki H & Tosa Y (2020) Evolution of an Eleusine -Specific Subgroup of Pyricularia oryzae Through a Gain of an Avirulence Gene. *Mol Plant-Microbe Interact* 33: 153–165

Bao J, Chen M, Zhong Z, Tang W, Lin L, Zhang X, Jiang H, Zhang D, Miao C, Tang H, *et al* (2017) PacBio Sequencing Reveals Transposable Elements as a Key Contributor to Genomic Plasticity and Virulence Variation in Magnaporthe oryzae. *Mol Plant* 10: 1465–1468 doi:10.1016/j.molp.2017.08.008 [PREPRINT]

Böhnert HU, Fudal I, Dioh W, Tharreau D, Notteghem JL & Lebrun MH (2004) A putative

polyketide synthase/peptide synthetase from Magnaporthe grisea signals pathogen attack to resistant rice. *Plant Cell* 16: 2499–2513

Bolger AM, Lohse M & Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120

Campbell MS, Holt C, Moore B & Yandell M (2014) Genome Annotation and Curation Using MAKER and MAKER-P. *Curr Protoc Bioinforma* 48: 4.11.1-39

Chiapello H, Mallet L, Guérin C, Aguileta G, Amselem J, Kroj T, Ortega-Abboud E, Lebrun M-H, Henrissat B, Gendrault A, *et al* (2015) Deciphering Genome Content and Evolutionary Relationships of Isolates from the Fungus Magnaporthe oryzae Attacking Different Host Plants. *Genome Biol Evol* 7: 2896–2912

Collemare J, O'Connell R & Lebrun MH (2019) Nonproteinaceous effectors: the terra incognita of plant–fungal interactions. *New Phytol* 223: 590–596

Collemare J, Pianfetti M, Houlle AE, Morin D, Camborde L, Gagey MJ, Barbisan C, Fudal I, Lebrun MH & Böhnert HU (2008) Magnaporthe grisea avirulence gene ACE1 belongs to an infection-specific gene cluster involved in secondary metabolism. *New Phytol* 179: 196–208

Dean R a, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu J-R, Pan H, *et al* (2005) The genome sequence of the rice blast fungus Magnaporthe grisea. *Nature* 434: 980–986

Dean R, Van Kan JAL, Pretorius ZA, Hammond-Kosack KE, Di Pietro A, Spanu PD, Rudd JJ, Dickman M, Kahmann R, Ellis J, *et al* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* 13: 414–430 doi:10.1111/j.1364-3703.2011.00783.x [PREPRINT]

Dellaporta SL, Wood J & Hicks JB (1983) A plant DNA minipreparation: Version II. *Plant Mol Biol Report* 1: 19–21

Ebbole DJ (2007) Magnaporthe as a Model for Understanding Host-Pathogen Interactions. *Annu Rev Phytopathol* 45: 437–456

Farman ML (2002) Pyricularia grisea isolates causing gray leaf spot on perennial ryegrass (Lolium perenne) in the United States: Relationship to P. grisea isolates from other host plants. *Phytopathology* 92: 245–254

Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL & Gurr SJ (2012) Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484: 186–194

doi:10.1038/nature10947 [PREPRINT]

Gladieux P, Condon B, Ravel S, Soanes D, Maciel JLN, Nhani A, Chen L, Terauchi R, Lebrun M-H, Tharreau D, *et al* (2018) Gene Flow between Divergent Cereal- and Grass-Specific Lineages of the Rice Blast Fungus Magnaporthe oryzae. *MBio* 9: e01219-17

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

Inoue Y, Vy TTP, Yoshida K, Asano H, Mitsuoka C, Asuke S, Anh VL, Cumagun CJR, Chuma I, Terauchi R, *et al* (2017) Evolution of the wheat blast fungus through functional losses in a host specificity determinant. *Science (80-)* 357: 80–83

Kachroo P, Leong SA & Chattoo BB (1994) Pot2, an inverted repeat transposon from the rice blast fungus Magnaporthe grisea. *MGG Mol Gen Genet* 245: 339–348

Kachroo P, Leong SA & Chattoo BB (1995) Mg-SINE: A short interspersed nuclear element from the rice blast fungus, Magnaporthe grisea. *Proc Natl Acad Sci U S A* 92: 11125–11129

Kumar S, Stecher G, Li M, Knyaz C & Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35: 1547–1549

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

Melnyk RA, Hossain SS & Haney CH (2019) Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated Pseudomonas. *ISME J* 13: 1575–1588

Patkar RN, Benke PI, Qu Z, Chen YYC, Yang F, Swarup S & Naqvi NI (2015) A fungal monooxygenase-derived jasmonate attenuates host innate immunity. *Nat Chem Biol* 11: 733–740

Patkar RN & Naqvi NI (2017) Fungal manipulation of hormone-regulated plant defense. *PLoS Pathog* 13: 10–14

Peng Z, Garcia EO, Lin G, Hu Y, Dalby M, Migeon P, Tang H, Farman M, Cook D, White F, *et al* (2018) Effector Gene Reshuffling Involves Dispensable Mini-chromosomes in the Wheat Blast Fungus. *bioRxiv*: 359455

Plissonneau C, Benevenuto J, Mohd-Assaad N, Fouché S, Hartmann FE & Croll D (2017) Using Population and Comparative Genomics to Understand the Genetic Basis of Effector-Driven Fungal Pathogen Evolution. *Front Plant Sci* 8

Raffaele S & Kamoun S (2012) Genome evolution in filamentous plant pathogens: Why bigger can be better. *Nat Rev Microbiol* 10: 417–430 doi:10.1038/nrmicro2790 [PREPRINT]

Risks to Plant Health - Launch of EASAC Report 24 | EASAC - Science Advice for the Benefit of Europe

Saleh D, Milazzo J, Adreit H, Fournier E & Tharreau D (2014) South-East Asia is the center of origin, diversity and dispersion of the rice blast fungus, Magnaporthe oryzae. *New Phytol* 201: 1440–1456

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A & Croll D (2018) The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annu Rev Phytopathol* 56: 21–40

Shen Q, Liu Y & Naqvi NI (2018) Fungal effectors at the crossroads of phytohormone signaling. *Curr Opin Microbiol* 46: 1–6

Skamnioti P & Gurr SJ (2009) Against the grain: safeguarding rice from rice blast disease. *Trends Biotechnol* 27: 141–150

Smit A, Hubley R & Green P RepeatMasker Open-4.0. 2013-2015.

Taylor JW & Fisher MC (2003) Fungal multilocus sequence typing--it's not just for bacteria. *Curr Opin Microbiol* 6: 351–6

Xue M, Yang J, Li Z, Hu S, Yao N, Dean R a, Zhao W, Shen M, Zhang H, Li C, *et al* (2012) Comparative analysis of the genomes of two field isolates of the rice blast fungus Magnaporthe oryzae. *PLoS Genet* 8: e1002869

Yoshida K, Saitoh H, Fujisawa S, Kanzaki H, Matsumura H, Yoshida K, Tosa Y, Chuma I, Takano Y, Win J, *et al* (2009) Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen Magnaporthe oryzae. *Plant Cell* 21: 1573–1591

Yoshida K, Saunders DGO, Mitsuoka C, Natsume S, Kosugi S, Saitoh H, Inoue Y, Chuma I, Tosa Y, Cano LM, *et al* (2016) Host specialization of the blast fungus Magnaporthe oryzae is associated with dynamic gain and loss of genes linked to transposable elements. *BMC Genomics* 17: 370

Zhong Z, Norvienyeku J, Chen M, Bao J, Lin L, Chen L, Lin Y, Wu X, Cai Z, Zhang Q, *et al* (2016) Directional Selection from Host Plants Is a Major Force Driving Host Specificity in Magnaporthe Species. *Sci Rep* 6: 25591