Chapter 1:
Review of Literature

1.1 Plant diseases – a threat to global food security

Plant diseases are a worldwide concern, wreaking havoc on food crop production as well as nation-state social and political stability. Over 80% of the food consumed by people and the majority of the nourishment for livestock comes from plants. Important staple crop yield losses on a global scale can reach up to 30%, costing hundreds of billions of dollars in lost food production (Rizzo et al., 2021). Plant diseases significantly reduce agricultural productivity and constitute a persistent danger to the safety of the world's food supply (Fisher et al., 2012; Pennisi, 2010). Both endemic and emerging plant diseases are expanding and worsening due to climate change, migration through international food exchange networks, pathogen spill over, and the emergence of new pathogen strains (Ristaino et al., 2021). Plant pathogens are difficult to control because their populations vary over time, space, and genotype. Most insidiously, they continue to evolve and often overcome the resistance that may have been the hard-earned achievement of the plant breeders (Strange & Scott, 2005).

Plants are prone to infection by plethora of microbial pathogens, including viruses, bacteria, fungi, oomycetes, nematodes and insect herbivores. 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. 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). Food security is often threatened in recent times by emerging plant diseases including late blight of potato, coffee rust, Panama disease of banana and blast disease caused by *Phytophthora infestans*, *Hemileia vastatrix*, *Fusarium oxysporum f. sp. cubense* (TR4) and *Magnaporthe oryzae,* respectively (Austin Bourke, 1964; Avelino et al., 2015; Stokstad, 2019). The top 10 fungal pathogens list, based on the scientific and agro-economic importance include *M. 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 mode of

reproduction is critical in order to understand the mechanisms underlying disease occurrence and spread (Taylor & Fisher, 2003).

1.2 The plant immune system in response to microbial pathogens

Plants inhabit intricate ecosystems where they engage closely with a diverse array of microbial pathogens employing various lifestyles and infection tactics. Plant pathogens 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. Plants have evolved sophisticated strategies to prevent or attenuate the invasion by microbial attackers.

First, the innate immune response identifies shared characteristics of the pathogens, such as, chitin, lipopolysaccharides, glycoproteins and flagellin (Göhre & Robatzek, 2008; Nürnberger & Kemmerling, 2018). These molecules and their derivatives are referred to as pathogen-associated molecular patterns (PAMPs; Göhre & Robatzek, 2008; Jones & Dangl, 2006). PAMPs engage pattern-recognition receptors (PRRs), which upon recognition induce diverse downstream defense signalling processes, culminating in the initiation of a basal resistance, referred to as PAMP-triggered immunity (PTI; **Fig. 1.1A**; Chisholm et al., 2006). Due to ongoing arms race between pathogens and their host plants, both co-evolve. Pathogens acquire effector molecules to enhance their own virulence that are translocated via secretion pathways into the host cell and suppress PTI, thus, resulting in effector-triggered susceptibility (ETS; **Fig. 1.1B**). Conversely, plants also develop resistance (R) proteins that identify the corresponding effectors in the pathogen, leading to a secondary and more robust immune response - effector-triggered immunity (ETI; **Fig. 1.1C**; Chisholm et al., 2006; Jones & Dangl, 2006).

Ultimately, the outcome of this arms race depends on the trade-off between the pathogen's ability to subdue the plants defense mechanism and the ability of the plant to distinguish the pathogen and to trigger effective defense processes.

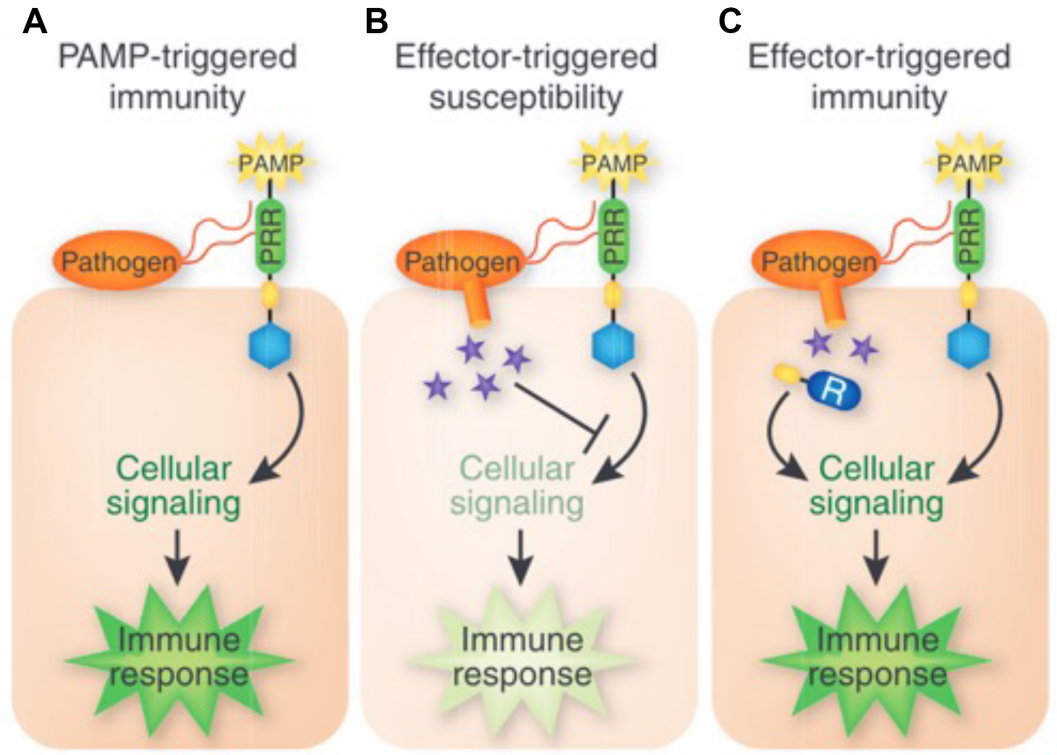

Figure 1.1: A simplified diagram illustrating the plant immune system. (A) When a pathogen attacks, pathogen-associated molecular patterns (PAMPs) activate host pattern-recognition receptors (PRRs), initiating a downstream signalling pathway leading to PAMP-triggered immunity (PTI). (B) Pathogens possess effectors (depicted as purple stars) that suppress PTI, and become virulent, in turn causing effector-triggered susceptibility (ETS). (C) Conversely, plants have developed resistance (R) proteins capable of recognizing these pathogen-specific effectors, triggering the subsequent immune response known as effector-triggered immunity (ETI) (adapted from Pieterse et al., 2009).

1.3 Blast disease caused by *Magnaporthe oryzae* - aka Cereal Killer

Ascomycete filamentous fungus *Magnaporthe oryzae* (synonym *Pyricularia oryzae*) is the causal agent of the most devastating blast disease. Typical annual yield losses range from 10-30%, which can go up to 100% when the environmental conditions are conducive. It can infect virtually all the aerial part of the cereal host plant, leading to leaf blast, neck and panicle blast, collar blast and node blast, finger blast (on millets). *M. oryzae* has the capability to invade rice host roots, propagate systemically in the plant system, and develop unique infection structures, which are different from the specialized appressoria, similar to the ones produced by related root-infecting fungi like, *Gaeumannomyces graminis* (Sesma

& Osbourn, 2004). Neck and finger blast can cause yield loss upto 70% when it occurs just before the grain formation (Ekwamu, 1991). Blast disease, initially referred to as rice fever disease in China as far back as 1637, was first documented in the United States in 1876. Rice blast disease has now been detected in 85 countries or regions around the world where rice is cultivated (**Fig. 1.2**; Wang et al., 2014). This disease has notably expanded across Sub-Saharan Africa, coinciding with the increasing demand for rice in recent years (Kihoro et al., 2013). Owing to the absence of resistance mechanisms and effective fungicides, a distinct strain of *M. oryzae* capable of infecting wheat was identified in Brazil in 1985, posing a significant threat to wheat production (Igarashi et al., 1986). The threat of wheat blast disease has extended to Bangladesh in Asia, presenting a grave peril to the continent's wheat cultivation (Islam et al., 2016; Malaker et al., 2016).

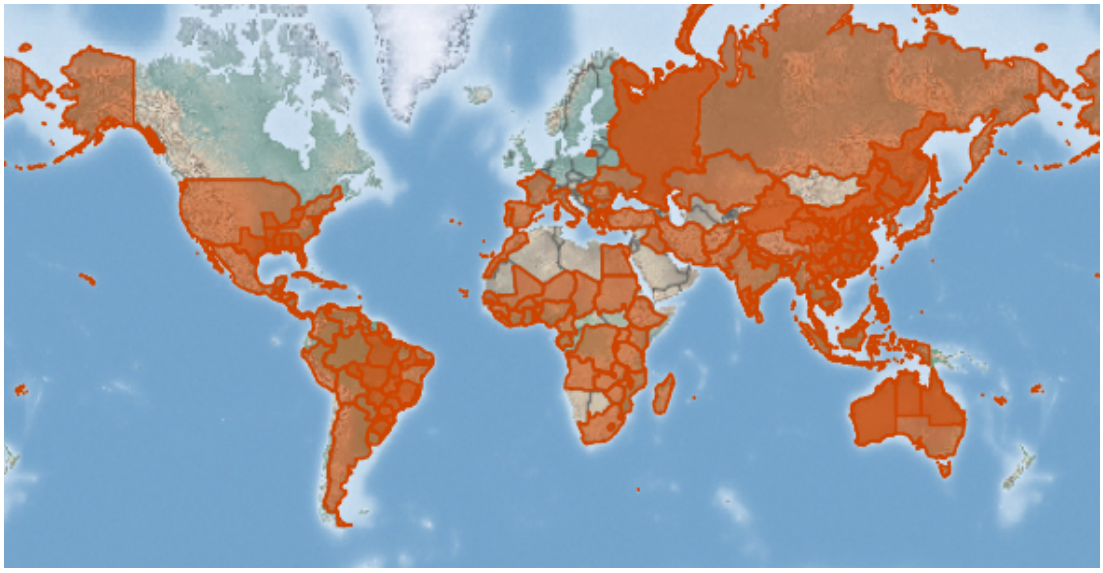

Figure 1.2: *Magnaporthe oryzae* distribution map (*Source: CABI, 2023*. https://plantwiseplusknowledgebank.org/doi/10.1079/PWKB.Species.46103). The intensity of orange colour in the map represents the number of *M. oryzae* strains identified in a certain geographic region.

In India, while the rice blast disease was first recorded in 1913, the first devastating epidemic was reported in 1919 in Tanjore delta (Madras state) (Padmanabhan, 1965). Between 1980-1987, several blast epidemics occurred in Himachal Pradesh, Andhra Pradesh, Tamil Nadu, and Haryana, resulting in huge financial losses. Although rice blast disease occurs

throughout India, certain regions, such as sub-Himalayan regions (Jammu and Kashmir, Himachal Pradesh, hill districts of Uttaranchal, and West Bengal), upland rice growing areas in eastern part (Arunachal Pradesh, Manipur, Mizoram, Meghalaya, Assam, Chotanagpur belt, and Jaypore tract of Orissa), western part (Konkan region of Maharashtra and Gujarat) and southern regions (Andhra Pradesh, Telangana, Tamilnadu, and Coorg region of Karnataka), are identified as hotspots (Annegowda et al., 2022). Similarly, grain yield losses up to 80-90% due to blast disease in finger millet have been reported in Karnataka state in the period from 1976 to1985 (Rao, 1990). Since then, apart from rice, *M. oryzae* field strains have been isolated and studied from various cereal crops such as pearl millet, finger millet, foxtail millet, barnyard millet, and grasses across various blast hotspots in the country (Palanna et al., 2023; Saha et al., 2016; Shirke et al., 2016).

1.3.1 The life cycle of the blast fungus

M. oryzae follows a hemibiotrophic lifestyle, characterized by an initial biotrophic phase where the fungus suppresses the plant immune system, followed by a transition to a necrotrophic phase that triggers plant cell death. The infection cycle begins when a three-celled conidium adhere to the leaf surface and germinates in response to the surface cues (**Fig. 1.3**; Hamer & Talbot, 1998; Talbot et al., 1993). The germ tube tip swells and forms specialized dome-shaped infection structure called appressoria around 4-6 hours post inoculation (hpi; Ebbole, 2007; Talbot, 2003). Appressorium maturation completes by 24 hpi, marked by the collapse of the spore, reutilizing of its contents into the developing appressorium, and the build-up of turgor pressure up to 8 MPa due to glycerol accumulation (De Jong et al., 1997). By 28 hpi, the penetration peg punctures the leaf cuticle to invade the underlying host cell.

Invasive growth by *M. oryzae* involves a prolonged biotrophic phase during which the fungus resides within host plant cells, surrounded by the invaginated plant plasma membrane. Within this primary cell, the penetration peg differentiates into a primary bulbous hypha referred to as invasive hyphae (IH; Kankanala et al., 2007). Host plant cytoplasm and IH are separated by the host-derived extra-invasive hyphal membrane (EIHM), and the EIHM separates from fungal cell wall by formation of EIHM matrix **(Fig. 1.4)**. *M. oryzae* deploys secreted effector molecules to suppress host basal resistance and

host reactive oxygen species (ROS) during this early biotrophic phase (Egan et al., 2007; Mosquera et al., 2009). Effectors could be secreted through two mechanisms: (i) apoplastic effectors are released into the EIHM matrix by the conventional endoplasmic reticulum (ER)–Golgi pathway, and (ii) cytoplasmic effectors are translocated into the host through the biotrophic interfacial complex (BIC) by an unconventional secretion route (Khang et al., 2010). IH constricts and moves through plasmodesmata to the neighboring living host cells, around 44–48 hpi. Each IH tips colonized to new cells undergo BIC formation, while the EIHM integrity is lost in the initially infected host cell, leading to its death. This process continues until extensive host cell death occurs, marking the transition to necrotrophy, which becomes visually apparent as necrotic lesions on the leaf surface by 120–144 hpi.

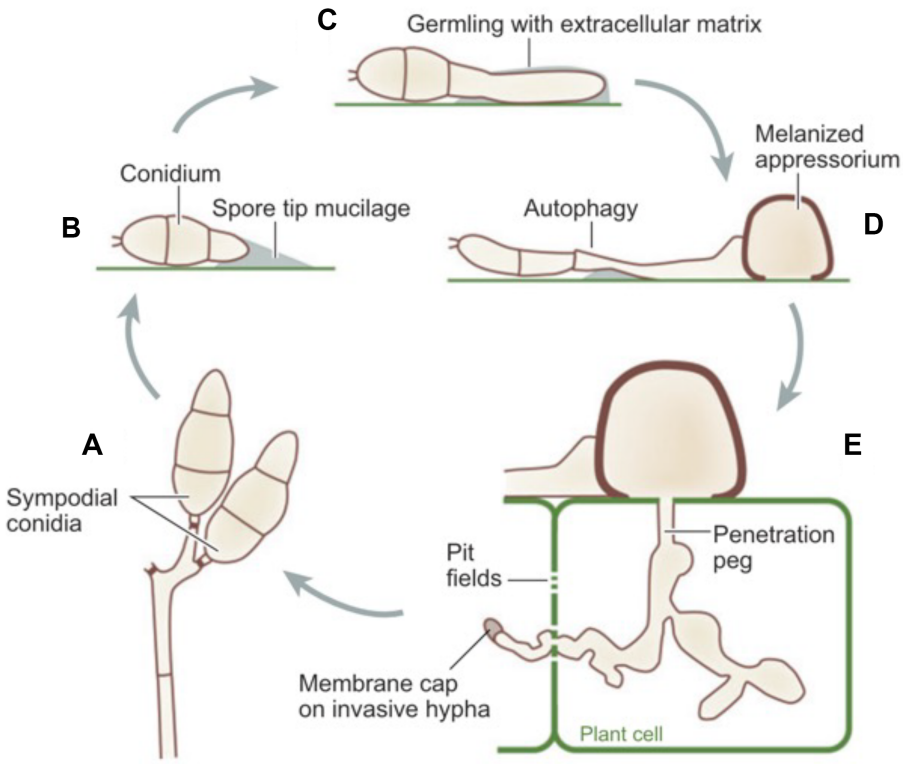

Figure 1.3: Disease cycle of *M. oryzae*. (A) Conidia are organized in a sympodial manner on the aerial conidiophore. (B) Conidium attaches to the host surface utilizing mucilage stored in a compartment located at the tip of the spore. (C) Conidia undergo germination and give rise to a hyphal filament enclosed by an extracellular matrix that adheres to the host surface. (D) The melanized appressorium is matured, comprising the autophagocytized contents of the conidium and germ tube. (E) A penetration peg employs internal turgor pressure developed in appressorium to penetrate the plant cuticle and cell wall. The

penetration peg gives rise to primary bulbous invasive hyphae (IH) and propagates within the host cell. Entry into the next cell occurs specifically at junction of plasmodesmata, depicted as pit fields. As IH traverse the plant cell wall, they constrict to a narrower diameter (adapted from Ebbole, 2007).

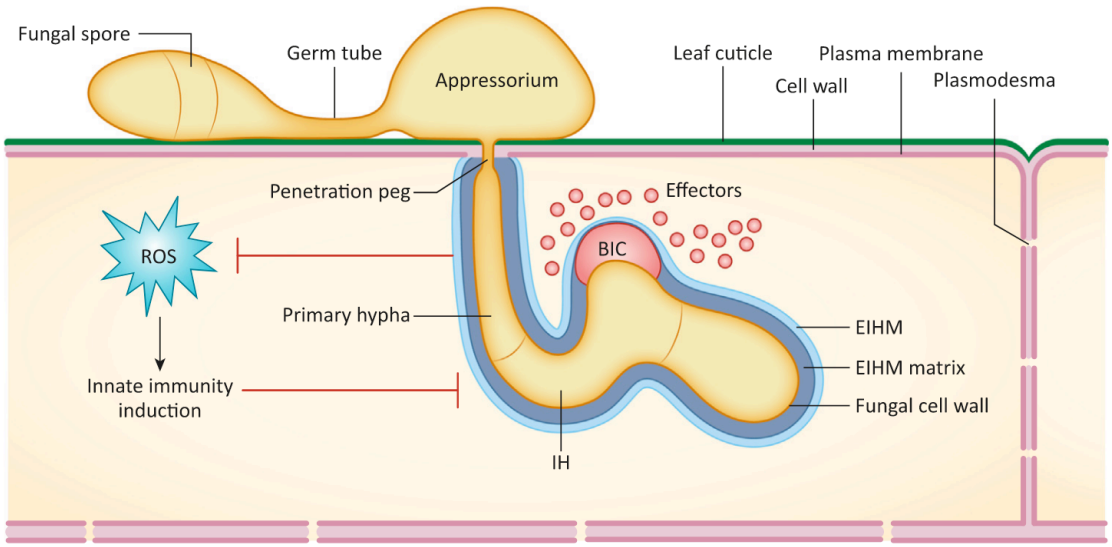

Figure 1.4: Schematic showing the early invasion into host cell by blast fungus. This diagram depicts the the early biotrophic colonization in living host rice cells by *M. oryzae* at 32–36 hours post inoculation (hpi). T-bars represent the inhibition of fungal biotrophic growth or host ROS accumulation (adapted from Wilson, 2021).

1.3.2 Sexual and asexual reproduction in *M. oryzae*

M. oryzae is classified as a heterothallic species, with two distinct mating types, MAT1-1 and MAT1-2, determined by specific genetic regions on chromosome 7 (Kanamori et al., 2007). For sexual reproduction to take place between two strains, it is necessary that at least one of the involved strains, regardless of its mating type, is female fertile and thus able to produce perithecia (Valent et al., 1991). Perithecia are flask shaped bodies containing ascospores arranged in unordered octads.

In the natural field conditions, *M. oryzae* primarily undergoes an asexual pathogenic cycle, and the sexual stage has been observed rarely. Although certain strains, from the centre of origin – southeast Asia - of this fungus, have demonstrated the ability to reproduce sexually under laboratory conditions, such female-fertile strains are rare in nature. Indeed, only a

few of them have been isolated from Himalayan foothills and Yunnan province in China (Zeigler, 1998). Population genetics studies have also indicated that sexual reproduction likely occurred in nature, in limited areas within this region (Saleh, Xu, et al., 2012). Consequently, *M. oryzae* comprises populations from its centre of origin where individuals exhibit a combination of sexual and asexual reproduction, and populations outside the centre of origin are exclusively clonal and exhibit asexual reproduction. The most plausible hypothesis could be that *M. oryzae* may have gradually lost its ability to undergo sexual reproduction, as it went on spreading from its centre of origin to the other parts of the world.

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

Due to ever-increasing fungal genomes and advancement in computational biology, we now have the capability to conduct large-scale investigations into population structures. The revolutionary progress in genomics-related tools has effectively bridged the gaps between molecular biology, evolutionary genetics, and epidemiology in the context of various diseases (Plissonneau et al., 2017).

Noteworthy achievements in comparative and population genomics have enabled a deeper understanding of effector biology and its evolutionary dynamics. Key insights from studies on plant pathogen genomes encompass the highly flexible genome structure, the presence of extensive repertoires of rapidly evolving candidate effector genes, and the non-random distribution of these effector genes within the genome. Interestingly, effector genes tend to be preferentially situated in genome regions that are rich in repeats and lacking in other genes (Raffaele & Kamoun, 2012). This compartmentalization of the genome, often referred to as the "two-speed" genome, and the favoured placement of effector genes in rapidly evolving compartments have independently emerged in various fungal lineages (Torres et al., 2020). This phenomenon likely plays a pivotal role in enabling plant pathogens to swiftly adapt to selective pressures (Dong et al., 2015). Genomics approaches are also emerging as the gold standard for deducing the genetic basis of adaptive divergence and for identifying the genomic determinants governing host adaptation and the expansion of host range. **Figure 1.5** provides an overview of the tools associated with comparative and population genomics, offering the potential to address diverse questions and significantly enhance our comprehension of pathogen biology. These tools encompass the assembly of multiple

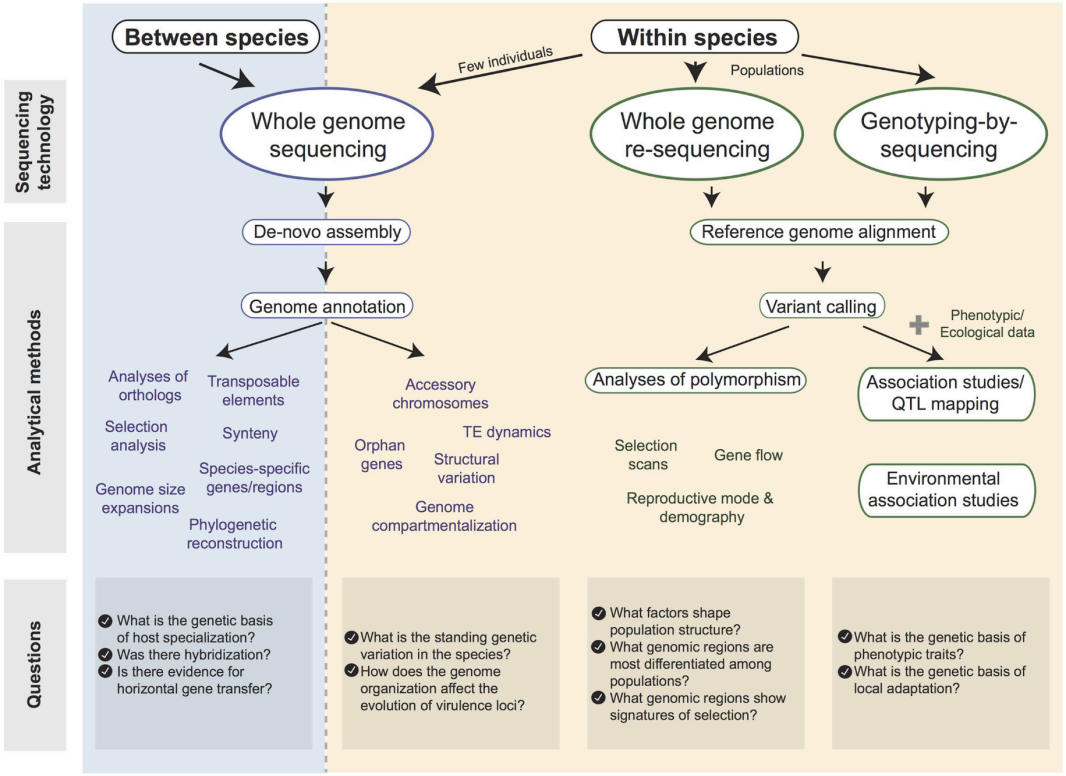

Figure 1.5: Diagram depicting the array of tools for examining fungal pathogen genomes. The diagram categorizes datasets as "between-species" (in blue) and "within-species" (in yellow). It presents various sequencing technologies and potential analytical approaches. Additionally, it outlines the primary research inquiries linked to each dataset type at the bottom (adapted from Plissonneau et al., 2017).

reference genomes and association mapping, which are pivotal for advancing our knowledge.

For various factors, *M. oryzae* stands out as a primary model organism for delving into the molecular intricacies of host-pathogen interactions (Ebbole, 2007). The first draft of reference genome of *M. oryzae* from 70-15 strain was made accessible via the Broad Institute, using sanger sequencing (Dean et al., 2005). This fungus possesses a relatively compact genome, measuring approximately 40 megabases (Mb), distributed across seven chromosomes (Dean et al., 2005). Subsequently, with the widespread availability of the next generation sequencing (NGS) technologies, whole-genome sequencing has become cost-effective for a broader spectrum of research teams. Currently, there are approximately 300

M. oryzae genomes accessible in public domain. 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 (Chiapello et al., 2015; Dean et al., 2005; Xue et al., 2012). Furthermore, additional genome analyses have revealed that major alterations in host specificity are primarily attributed to gene gains/losses and chromosomal structural changes, often instigated by the activity of transposable elements (Bao et al., 2017; Latorre et al., 2020; Peng et al., 2018; Yoshida et al., 2016).

1.5 Population dynamics of a multi-host blast fungal pathogen

Comparative genomics analyses has proven to be quite useful in understanding the population structure and host-specialization of *M. oryzae* (Chiapello et al., 2015; Gladieux, Condon, et al., 2018; Yoshida et al., 2016). Blast fungal pathogen comprises a complex mix of genetically distinguished lineages, each tend to be corelated with certain host genera. Notably, *M. oryzae* was established as a distinct species, separate from *Magnaporthe grisea*, through comprehensive studies involving multi-locus genealogy and mating experiments (Couch & Kohn, 2002). Phylogenetic analysis classified *Magnaporthe* isolates into two clades, one primarily infecting *Digitaria* (crabgrass) referred to as *M. grisea*, and another associated with infections in cereal crop plants (rice, millets, wheat, oat, maize, rye etc.) and other grasses, known as *M. oryzae* (Couch and Kohn, 2002).

The lineage responsible for infection on rice (*Oryza sativa*) diverged from the one infecting foxtail millet *(Setaria italica)* roughly 2500–7500 years ago, during or right after rice domestication in East Asia (Couch, 2005). In Asia, the population structure of *M. oryzae* comprised of numerous diverse genetic lineages, predominantly in countries like India, China, and Thailand, in contrast to simpler clonal populations observed in the USA, Europe, Japan, and the Philippines (Skamnioti & Gurr, 2009). Interestingly, South-East Asia has been recognized as the centre of origin of the blast fungus, with most *M. oryzae* populations presumably migrated/radiated from Asia to the other continents (Saleh et al., 2014). Two centres of genetic diversity were revealed in the Himalayan foothills: South China–Laos–North Thailand, and western Nepal. Sexual reproduction persisted only in the South China–Laos–North Thailand region, which was identified as the putative centre of origin of all *M. oryzae* populations on rice (Saleh et al., 2014).

The evolutionary history of the global population of rice infecting strains, revealed their divergence into the three clonal lineages. Each of these lineages exhibits a precise patterns of presence/absence variations in the effector repertoire, potentially influencing their specialization to the rice host. (Gladieux, Ravel, et al., 2018; Latorre et al., 2020; Zhong et al., 2018).

Based on a population genomics based analysis on a globally collected dataset encompassing 76 genomes of *M. oryzae* from 12 grass host genera, it was found that *M. oryzae* has diversified into multiple host-specific lineages, while inter-lineage gene flow maintains them as a single species (**Fig. 1.6**; Gladieux, Condon, et al., 2018). These host-specific lineages within *M. oryzae* show a little gene flow but transposon activities in repeat-rich genomic regions show frequent gene gain and loss. This process of speciation often results from significant changes in the pathogen's genome, enabling adaptation to new host environments (Raffaele & Kamoun, 2012). The blast fungus's ability to overcome host resistance leads to the emergence of new strains with heightened virulence and multiple host-adapted lineages within the same species (Gladieux, Condon, et al., 2018). This speciation process primarily arises from the interaction between the pathogen and its host-plants, and poses a substantial threat to crop production, as exemplified by the recent emergence of wheat blast due to a host-shift in the fungal pathogen *M. oryzae* (Inoue et al., 2017; Islam et al., 2016).

Gladieux and colleagues' genome-wide analyses support a model where *M. oryzae* is a single species consisting of well-differentiated lineages that still exhibit some level of genetic exchanges and recombination. This reflects a pattern of incipient speciation driven by reproductive isolation, followed by ecological specialization on specific hosts, counterbalanced by the ability of distinct lineages to infect common host plants, allowing occasional genetic exchanges (Gladieux, Condon, et al., 2018). Co-infection on shared host plants can facilitate genetic exchange between different *M. oryzae* lineages **(Fig. 1.7)**. While individual lineages often demonstrate host specificity, under specific ecological conditions, distinct lineages may coexist on a common host, enabling genetic exchange through processes like recombination and gene flow. Subsequent selection pressures from the original primary hosts can either eliminate disadvantageous segments or facilitate the emergence of new sub-lineages.

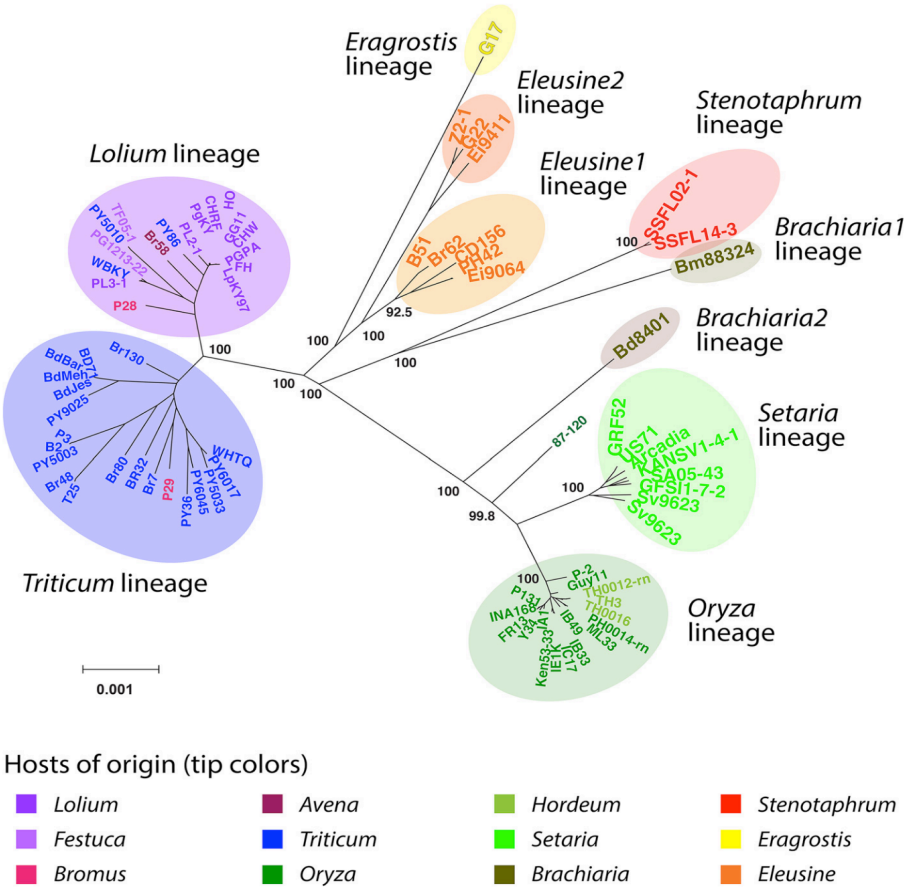

Figure 1.6: Species tree of the blast fungus showing its divergence into multiple host-specific lineages. The strains of *M. oryzae* were collected across worldwide and their hosts of origin are colour coded (adapted from Gladieux, Condon, et al., 2018).

Furthermore, horizontal transfer of genetic material and potential occurrence of sexual production in certain geographical regions may contribute to the genetic exchanges among different lineages of *M. oryzae*. The concept of "mobile effectors" was introduced in response to the horizontal transfer of the AVR-Pita effector genes among the different lineages of *M. oryzae*, a process that permits asexual lineages to regain lost genes (Chuma et al., 2011). While the blast fungus has long considered to be exclusively asexual, there are evidences suggested that blast fungus did undergone sexual reproduction in certain regions within its geographic range (Saleh, Xu, et al., 2012). Thus, Parasexual recombination and sexual mating are additional mechanisms that can lead to the exchange of genetic material, contributing to the variability in pathogenicity observed in the blast fungus.

1.6 Role of Effectors in pathobiology

During interactions between the plants and fungal pathogens, fungi secrete hundreds of proteins known as effectors, which manipulate host physiology or suppress host defences to promote infection (Franceschetti et al., 2017). Historically, the first effectors to be recognized were small, secreted proteins that caused avirulence as a result of their recognition by plant immune receptors. This discovery influenced the systematic investigations to pinpoint their specific targets and receptors responsible for recognition within host plants (Domazakis et al., 2017). As a result, the field of effector biology has largely concentrated exclusively on these secreted proteins (Varden et al., 2017), while largely overlooking the potential roles played by fungal secondary metabolites (SMs), also referred to as chemical effectors, and small noncoding RNAs (sRNA effectors). These non-proteinaceous effectors (NPEs) may have similar functions but have received less attention thus far (Collemare et al., 2019).

1.6.1 Protein effectors

In accordance with the gene-for-gene concept, host plant cultivars possessing major blast resistance (R) genes can effectively prevent the infection caused by strains carrying the cognate avirulence (AVR) genes. The recognition between R genes and AVR genes underpins this defense mechanism (Notteghem & Silue, 1992). Pathogen deploy certain small secreted proteins (SSP) as AVR factors, as their recognition by host R proteins leads to incompatible interactions between a host and its pathogen. Nonetheless, AVR genes exhibit a high degree of variability, and mutated AVR proteins can evade recognition by their cognate R proteins, thereby compromising the resistance conferred by the R gene (Deng et al., 2017). In agricultural settings, the relentless cultivation of a single resistant cultivar over extended periods exerts strong selection pressure on the blast pathogen, enabling it to overcome R gene-mediated resistance (Skamnioti & Gurr, 2009; Valent & Khang, 2010). Consequently, assessing the variations in AVR genes serves as an effective means to gauge the effectiveness of corresponding resistance genes. Therefore, monitoring AVR gene variations within the field's blast fungal population is a proactive approach to predicting R gene efficacy and averting blast disease outbreaks.

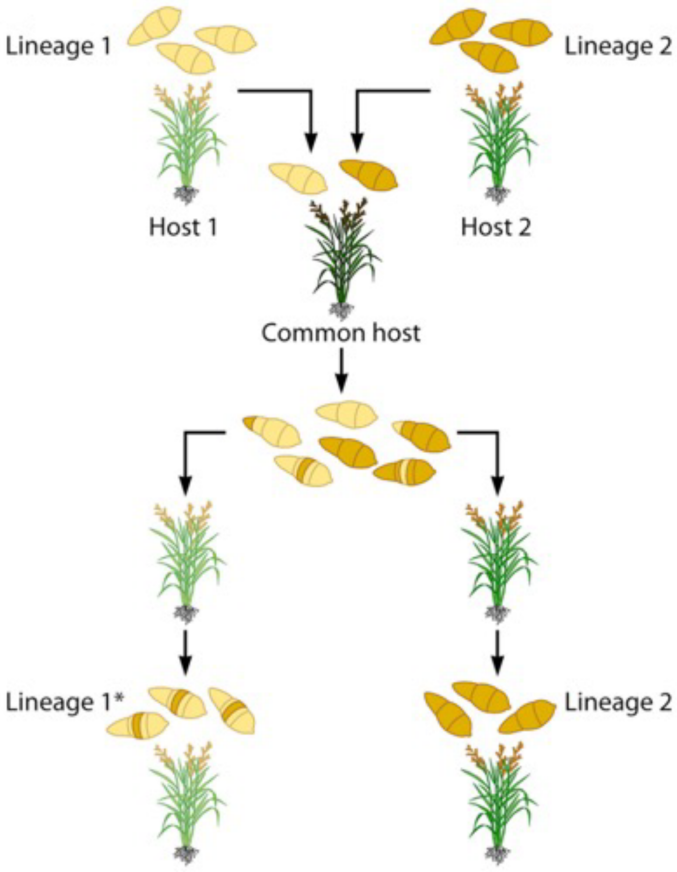

Figure 1.7: Coinfection on common host plants as one of the potential mechanisms for genetic exchange between distinct *M. oryzae* lineages. Lineage 1 (yellow) infects host 1 (light green) and lineage 2 (brown) infects host 2 (green). However, under certain ecologic conditions, two separate lineages may occasionally invade or shift on a common host (dark green). Coinfection on a mutual host enables inter-lineage genetic exchanges. Subsequently, host selection pressure can either eliminate the introgressed segments (lineage 2) or evolve in to a new sub-lineage (lineage 1*; adapted from Langner et al., 2018)

As of now, over 100 blast resistance (R) genes have been genetically mapped, with more than 40 AVR genes identified within the blast fungus. Notably, 12 of these AVR genes (AVR-Pita1, AVR1-CO39, AVR-Pik, AVR-Pizt, AVR-Pia, AVR-Pib, AVR-Pii, AVR-Pi9, AVR-Pi54, PWL1, PWL2 and ACE1) have been successfully cloned (Böhnert et al., 2004; Kang Seogchan et al., 1995; Kanzaki et al., 2012; W. Li et al., 2009; Miki et al., 2009; Orbach et al., 2000; Park et al., 2012; Ray et al., 2016; Ribot et al., 2013; Sweigard et al., 1995; Wu et al., 2015; Yoshida et al., 2009; S. Zhang et al., 2015).

Table 1.1: *Magnaporthe oryzae* effector proteins identified till date (adapted from Fernandez & Orth, 2018)

Effector name	ID number	Size[a]	Function/expression/ localization[b]	R[c]	References
Avirulent effectors					
AVR-Pita	AF207841	224	Zinc metalloprotease/in planta[d] / BIC; cytoplasm	Pita	(Chuma et al., 2011; Jia et al., 2000; Khang et al., 2010; Orbach et al., 2000)
AVR-Pizt	HE578813	108	Target rice U3 ubiquitin ligase; reduces Flg-22 and chitin induced ROS production/in planta/cytoplasm	Piz-t	(Park et al., 2012)
AVR-CO39	AF463528	89	Interacts with RGA4/RGA5 /in planta/ cytoplasm	CO39	(Cesari et al., 2013; Ribot et al., 2013)
AVR-Pia	AB498873	85	–/In planta/cytoplasm	Pia	(Yoshida et al., 2009)
AVR-Pii	AB498874	70	C2H2 Zinc-finger motif/ in planta/cytoplasm	Pii	(Yoshida et al., 2009)
AVR-Pik/km/kp	AB498875-AB498879	113	–/In planta/cytoplasm	Pik	(Yoshida et al., 2009)
AVR-Pib	KM887844	74	–/–/Cytoplasm predicted	Pib	(S. Zhang et al., 2015)
AVR-Pi9	MGG_12655	91	–/In planta/BIC; cytoplasm	Pi9	(Wu et al., 2015)
PWL1	AB480169	147	–/In planta/BIC; cytoplasm	-	(Kang Seogchan et al., 1995)
PWL2	MGG_04301	145	–/In planta/BIC; cytoplasm	-	(Giraldo & Valent, 2013; Kang Seogchan et al., 1995; Khang et al., 2010; Sweigard et al., 1995)
ACE1	AJ704622	4035	Polyketide synthase-peptide synthetase/appressorium penetration/appressorium; cytoplasm	Pi33	(Böhnert et al., 2004)
AVR-Pi54	MGG_01947	153	–/In planta/appressorium	Pi54	(Ray et al., 2016)
Secreted proteins					
SLP1	MGG_10097	162	LysM domains, binds to chitin oligosaccharides; suppresses chitin-induced immunity in rice/in planta/ apoplast	–	(Mentlak et al., 2012)

MC69	MGG_02848	54	–/in planta/apoplast	–	(Saitoh et al., 2012)
MSP1	MGG_05344	137	Cerato-platanin family; triggers autophagic cell death and elicits host defense responses/in plant/apoplast	–	(Y. Wang et al., 2016)
Biotrophy-associated secreted proteins					
BAS1	MGG_04795	115	–/In planta/BIC; cytoplasm	–	(Khang et al., 2010; Mosquera et al., 2009)
BAS2	MGG_09693	102	–/In planta/ BIC; cell wall crossing points	–	(Khang et al., 2010; Mosquera et al., 2009)
BAS3	MGG_11610	113	–/In planta /BIC; cell wall crossing points	–	(Khang et al., 2010; Mosquera et al., 2009)
BAS4	MGG_10914	102	–/In planta/EIHM	–	(Khang et al., 2010; Mosquera et al., 2009)
BAS 107	MGG_10020	132	–/In planta /BIC; cytoplasm	–	(Giraldo & Valent, 2013)
Group of predicted effector proteins					
IUGs			Isolate unique genes; suppression of SA and ET signaling/in planta/BIC; cytoplasm		(Dong et al., 2015)
MoHEGs			*M. oryzae* hypothetical effector genes; suppress plant defenses; virulence function in barley plants/ in planta/appressorium		(Mogga et al., 2016)
MoCDIPs			Induced plant cell death in rice protoplasts/in planta/apoplast		(Chen et al., 2013)
SDPs			Suppression of plant host defenses/in planta/ apoplast and cytoplasm		(Sharpee et al., 2017)

[a] Indicates the number of amino acid residues in the predicted protein.
[b] Inside the host plant cell upon rice blast infection.
[c] Cognate resistance protein in rice plants.
[d] In planta: expression during biotrophic growth – unknown function or host interactor.

However, our understanding of the virulence functions of only a limited subset of AVR proteins is currently available **(Table 1.1)**. These findings suggest that AVR proteins may target various host factors to manipulate host immunity. Nevertheless, the precise functions of AVR proteins remain largely uncharted territory.

The blast effectors can be categorized into two groups based on their localization in host cells. Some well-known AVR effectors including like AVR-Pita, AVR-Pizt, PWL1, and PWL2, along with the secretory protein effectors MC69 and BAS1 (biotrophic-associated secreted protein), initially accumulate within the BIC structure of the primary IH before they are translocated into the host cell cytoplasm. Remarkably, once these cytoplasmic effectors are transferred into the primary invaded cell, some of them gather around the points where the host cell wall is breached or migrate into neighbouring uninfected cells, seemingly priming these adjacent uninvaded host cells for subsequent infection. Notably, BIC structures are observed in every newly infected cell and serve as sites of cytoplasmic effector accumulation. In contrast, apoplastic effectors do not secrete through the BIC structure. These apoplastic effectors, including the SLP1 (secreted LysM protein 1), BAS4, and BAS113, accumulate within the enclosed apoplastic region located between the fungal cell wall and the extra-invasive hyphal membrane (EIHM), effectively outlining the entire invasive hyphae structure (Fernandez & Orth, 2018).

1.6.1.1 MAX effectors

One notable characteristic of effector proteins is their tendency to possess unique sequences with limited similarity to the known proteins. Surprising similarities have been discovered while investigating the three-dimensional structures of effector proteins that are not closely related in terms of their sequences. For instance, the three-dimensional structures of *M. oryzae* protein effectors, such as AVR-PikD, AVR-Pia, AVR-CO39 and AVR-Pizt, all share a common core fold, despite lacking significant sequence similarity (**Fig. 1.8**; De Guillen et al., 2015; Maqbool et al., 2015; Ose et al., 2015; Z.-M. Zhang et al., 2013).

Surprisingly, the host-selective toxin ToxB from a phylogenetically distant pathogen *Pyrenophora tritici-repentis*, also exhibits this same structure—a six-stranded β-sandwich consisting of two antiparallel β-sheets (Nyarko et al., 2014). These effectors having the conserved core made of β-sandwich were designated as MAX effectors (for *Magnaporthe*

AVRs and ToxB like) and identified more effector-like proteins of this family from various other ascomycete fungi. MAX effectors constitute 5 to 10% of the total effector repertoire in *M. oryzae* and related species *M. grisea*, the gene family is significantly expanded through diversifying evolution in blast fungus.

It is important to note that although MAX effectors show common structural fold, they need not have similar molecular functions. Further, a common fold with sequence diversity in different lineages could be associated with the adaptive evolution and development of new functions for these effectors (De Guillen et al., 2015; Yoshida et al., 2016). For example, while both AVR-PikD and AVR-Pia individually interact with proteins containing heavy metal-associated (HMA) domains, AVR-Pizt targets the E3 ubiquitin ligases (Cesari et al., 2013; Park et al., 2012). The common structural fold of MAX effectors seems to likely provide a fundamental framework, which might confer stability and enable effective translocation into the host during invasion.

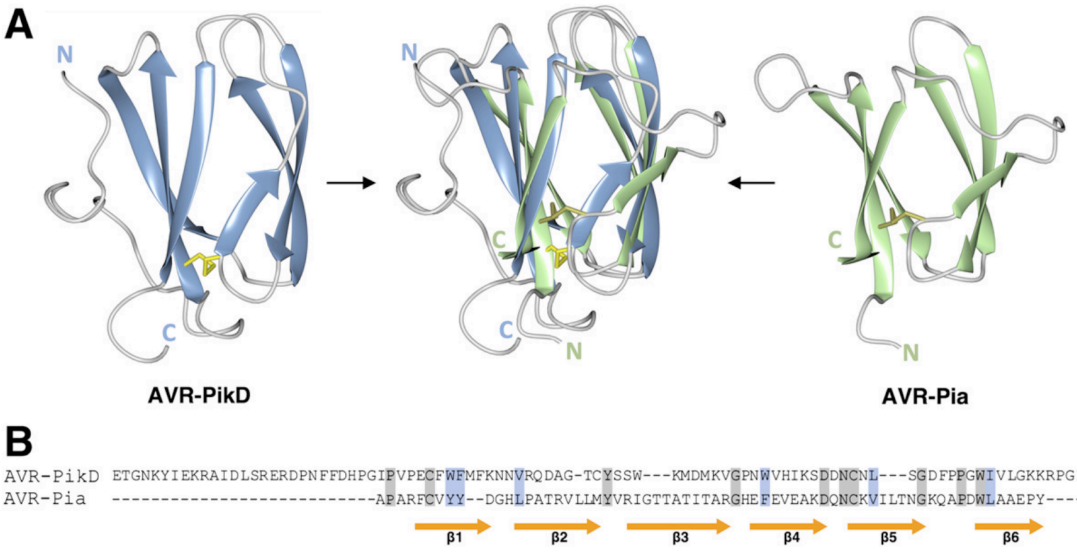

Figure 1.8: MAX effectors AVR-PikD and AVR-Pia exhibit the same structural fold. (A) AVR-PikD shown on the left, AVR-Pia on the right, while their superposition in the centre, demonstrating the conserved β-sandwich MAX fold represented as a cartoon. The cysteine residues forming disulphide bonds are depicted as sticks, and the loops are shaded in grey. The amino and carboxyl termini are 24abelled N and C, respectively. (B) Alignment of amino acid sequences of AVR-PikD and AVR-Pia, displaying the absence of sequence conservation between these two effectors. Shaded regions indicate amino acids that are either conserved or share similar properties. The structurally conserved β-strands (from β1 to β6) are depicted as yellow arrows (adapted from Białas et al., 2018).

1.6.1.2 Role of protein effectors in host specialization

The ability of the blast pathogen to infect various hosts largely depends on the effector repertoire, which can serve as major determinants in host specialization (Raffaele et al., 2010; Sánchez-Vallet et al., 2018). A successful pathogen must continuously maintain the capacity to evade host recognition and sustain its virulence on the host. Evolution toward evasion of recognition and functional optimization occurs through processes like sequence modification, gene deletion, alteration in the expression of existing effector genes, and the acquisition of new effectors (**Fig. 1.9**; Lo Presti et al., 2015). Through these mechanisms, pathogens either mutate existing effectors or acquire new ones to specialize on a new host (Poppe et al., 2015). Functional redundancy often arises from recent duplications of effector genes and is believed to play a crucial role in the risk management strategy of plant pathogens. The ability of the pathogen to cause disease can be severely harmed by the emergence of a new host genotype that recognizes a particular effector. However, pathogen populations can adjust by losing the gene encoding the identified effector if many effectors attack the same host route. The absence of an effector aids the pathogen's ability to avoid detection without impairing its ability to target the host route (Lo Presti et al., 2015; Mosquera et al., 2009; Win et al., 2012).

The avirulent characteristic of AVR genes can be nullified through a variety of mechanisms, including mutations (SNPs, deletions, insertions), genetic recombination, and sexual mating (Noguchi et al., 2006). Transposable elements (TEs) can increase the diversity in effector repertoire by deleting/inactivating genes or horizontal gene transfer (Chuma et al., 2011; Yoshida et al., 2016). The promoter or genomic regions encompassing important pathogenicity related genes have been frequently observed with transposon. For instance, Pot3 has been recognized as a major factor for the generation of several variants of AVR-Pita1, AVR-Pia, AVR-Pib, AVR-Pii and AVR-Pizt (Kang, 2001; W. Li et al., 2009; Singh et al., 2014; Yoshida et al., 2009). Point mutations of AVR-Pita1 and AVR-Pik may result in the generation of novel alleles, some of which may be resistant to detection by the cognate R gene (Kanzaki et al., 2012; Yoshida et al., 2009).

The pathogenic mutation of AVR-Pita1 and AVR-Pib is also caused by segmental loss of the coding sequence (Orbach et al., 2000; S. Zhang et al., 2015). AVR genes were reported to be completely deleted in AVR-Pita1, AVR-Pia and AVR-Pii, so that resulting strains continue to be pathogenic to cultivars retaining their cognate R genes. Because AVR-Pita1

and AVR-Pii are located next to their chromosome telomeres, chromosomal tip loss is the cause of their spontaneous loss (Khang et al., 2008; W. Li et al., 2009; Yasuda et al., 2006). Furthermore, homologous recombination between strains can result in genomic rearrangement, translocation, deletion and even horizontal gene transfer. Homologous recombination between two repeated sequences resulted in the loss of AVR-Pia. As a result of these changes in AVR genes in a population, dynamic adaptation and development of rice blast fungus could be achieved, leading to the breakdown of R genes (Hu et al., 2022).

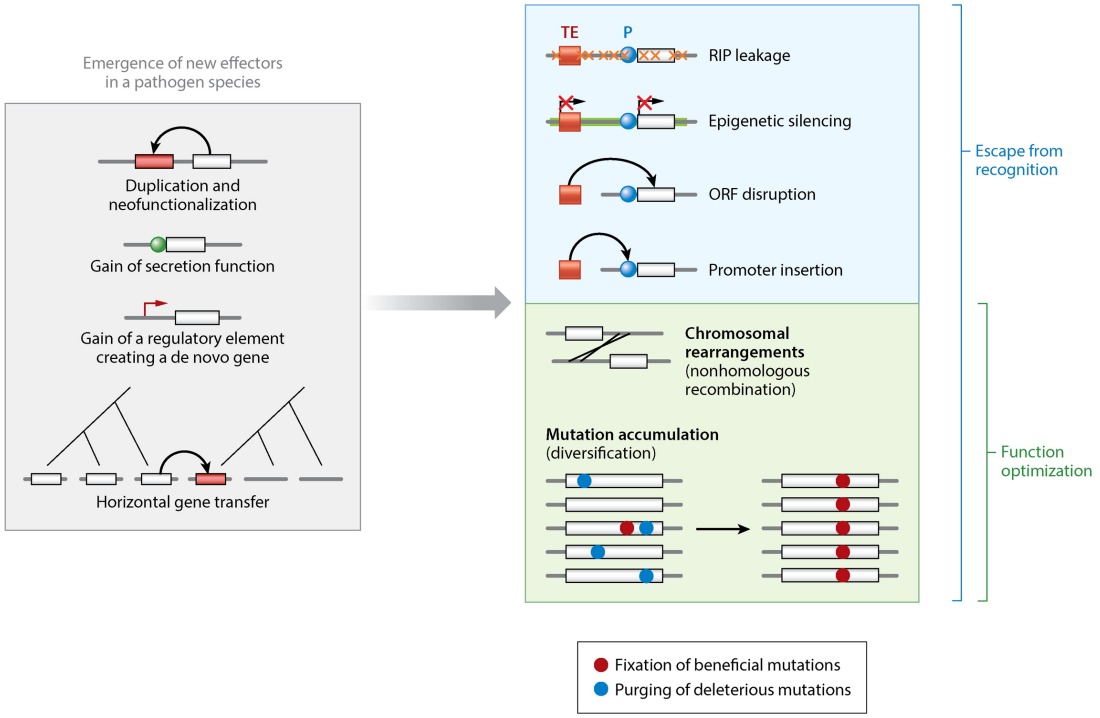

Figure 1.9: The evolutionary dynamics of effectors. New effectors can emerge through either gene duplication or the acquisition of a secretion function. Effector genes may also evolve de novo from noncoding sequences through the gain of a regulatory element or be acquired via horizontal transfer from a different pathogen species. When the host recognizes the encoded effector, effector genes can undergo rapid sequence evolution. The primary mechanism leading to the loss of an effector gene is the presence and activity of nearby transposable elements (TEs). These transposable elements can induce repeat-induced point (RIP) mutations, trigger epigenetic silencing, or disrupt the gene sequence Evading host recognition can also occur through chromosomal rearrangements or the fixation of advantageous mutations. Rearrangements and the selection for advantageous mutations are also significant pathways through which effectors fine-tune their functions. ORF: open reading frame; P: promoter regions (adapted from 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. 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 harbouring the corresponding *RWT3* resistance gene (Inoue et al., 2017).

1.6.2 Chemical effectors

Based on extensive genome-wide expression data, it is well-documented that fungal pathogens produce a wide variety of secondary metabolites (SMs) when infecting plants, particularly during the early biotrophic/asymptomatic stages of infection, such as penetration and primary infection. During these stages, the plant cells are typically still alive, suggesting that these fungal SMs are unlikely to function as cytotoxins that directly kill host cells. However, it's important to note that many of these fungal molecules have not been fully characterized chemically, and their specific targets within the plant remain unidentified.

Nonetheless, we can infer a broad range of potential plant cellular targets for these fungal SMs based on the known biological activities of SMs produced by fungi during plant colonization **(Fig. 1.10)**. For instance, various fungal SMs, like cytochalasans, have been shown to inhibit actin polymerization, a crucial process in cell structure and movement (Skellam, 2017). Additionally, several fungal SMs affect protein biosynthesis, processing, and delivery. For example, compounds such as brefeldin A and related macrolides, which are produced by numerous fungal plant pathogens and endophytes, disrupt protein secretion by targeting specific guanine nucleotide exchange factors associated with ADP-ribosylation factor GTPases (ARF-GEFs), which regulate vesicle formation (Kwon et al., 2008; Nielsen et al., 2012; Wang et al., 2002). Others, like trichothecenes and tenuazonic acid, interfere with ribosomal protein synthesis (Audenaert et al., 2013; Chen & Qiang, 2017), and radicicol inhibits the Hsp90 protein-folding machinery (Wicklow et al., 2009). Fungal SMs also disrupt plant regulatory networks by inhibiting processes such as calmodulin signaling

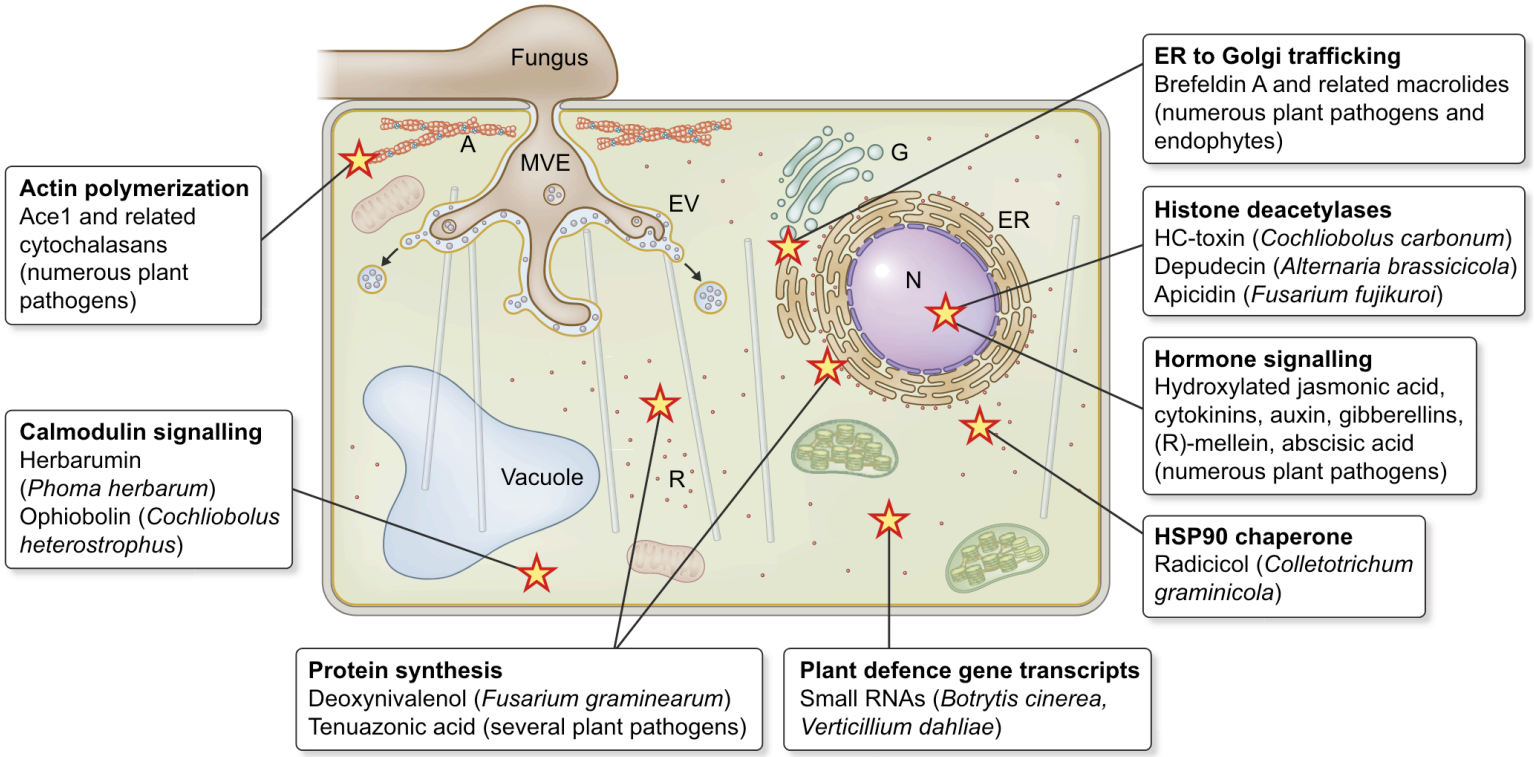

Figure 1.10: Fungal non-proteinaceous effectors (NPEs) illustrating potential effector-like functions at the interface of plant–fungal interactions. Yellow–red stars depict the putative or characterized host cellular targets of NPEs. The labels in the schematic corresponds to different cellular organelles: A, actin; ER, endoplasmic reticulum; G, Golgi apparatus; MVE, multivesicular endosome; N, nucleus; R, ribosome (adapted from Collemare et al., 2019).

(e.g., ophiobolin and herbarumin; Au et al., 2000) or histone acetylation (e.g., HC toxin and depudecin; Walton, 2006; Wight et al., 2009). Furthermore, nearly all fungi are capable of either producing or modifying plant hormones, including jasmonic acid, auxin, cytokinins, gibberellic acids, and abscisic acid, thereby disrupting corresponding plant hormone signaling pathways (Patkar et al., 2015; Shen et al., 2018). In some cases, the modification of signaling pathways controlling plant root development has been suggested, such as in the case of the ectomycorrhizal fungus *Laccaria bicolor*, which produces volatile sesquiterpenes involved in symbiosis establishment (Ditengou et al., 2015).

The diversity of plant cellular functions potentially targeted by fungal SMs is reminiscent of proteinaceous effectors, which inhibit multiple plant targets to suppress plant defenses and facilitate infection. It's important to note that known fungal toxins likely represent only a small fraction of the spectrum of non-proteinaceous effectors (NPEs), and the precise functions of the vast majority of fungal SMs produced during plant colonization are yet to be fully understood.

1.6.2.1 Organization of biosynthetic gene cluster (BGC) producing SMs

Fungal SMs are typically derived from a limited number of precursors originating from primary metabolism. These precursors follow specific metabolic pathways, which can be categorized based on the core enzyme responsible for initiating the synthesis of the first intermediate and stable core product **(Fig. 1.11A)**. Polyketide synthases (PKSs), non-ribosomal peptide synthetases (NRPSs), dimethylallyl tryptophan synthases, and terpene cyclases are responsible for the production of polyketides, non-ribosomal peptides, indole alkaloids, and terpenes, respectively (Keller, 2019). Hybrids of each class are also very frequently found in fungal pathogens. PKSs and NRPSs are large proteins with multiple functional domains, serving as core enzymes, where all their domains work in tandem to facilitate the formation of the developing intermediate. Further tailoring enzymes modify these core enzyme products by addition of reactive groups, leading to the final active SM product. In the fungal kingdom, all the genes responsible for encoding both the core and tailoring enzymes of a specific biosynthetic pathway are often grouped together in the genome and exhibit coordinated regulation. This arrangement is referred to as a biosynthetic gene cluster (BGC) (**Fig. 1.11B**; Keller & Hohn, 1997). Fungal BGCs are reminiscent to

arrangement in bacterial SM operons (Keller et al., 2005). These biosynthetic gene clusters may also encompass genes for transporters required for the efflux of SMs or self-defense mechanisms, as well as transcription factors that control the expression of genes within the biosynthetic pathway (Gardiner et al., 2005; Keller et al., 2005). This physical linkage of SM-related genes has the potential to enable tight co-regulation, facilitating highly coordinated interactions among enzymes participating within the same metabolic pathway (McGary et al., 2013; Thomma et al., 2016). BGCs are often regulated via biosynthetic-pathway-specific transcription factors, which are mostly present in/around the cluster (Brown et al., 2015; Fernandes et al., 1998; Proctor et al., 1995). Additionally, BGCs could be regulated by global TFs, which are responsive to carbon, nitrogen, light and pH, and could also be under epigenetic control (Macheleidt et al., 2016; Monroy et al., 2017; Tilburn et al., 1995; Tudzynski et al., 1999). Indeed, a likely histone methyl transferase LaeA has been implied in regulation of certain BGCs, including that for sterigmatocystin biosynthesis, in a few species of *Aspergillus* (Bok & Keller, 2004).

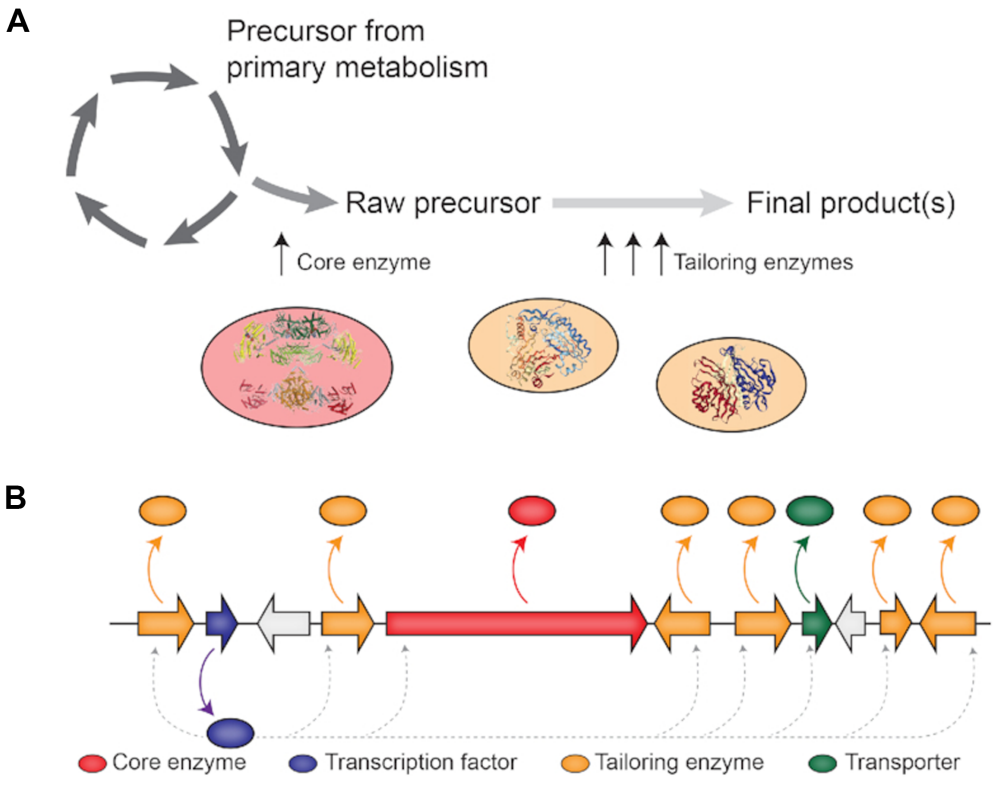

Figure 1.11: Secondary metabolites are produced via metabolic pathways, encoded by biosynthetic gene clusters. (A) Secondary metabolites (SMs) are produced from a limited number of precursors via biochemical pathways centered around distinctive core enzymes. The tailoring enzymes alter the core biosynthetic product further so to produce the final biologically active SM product. (B) Genes encoding core (red) and tailoring enzymes (orange) as well as transporters (green) involved in SM biosynthesis are physically clustered together in the genome. This arrangement is referred to as biosynthetic gene cluster (BGC). BGC may also contain transcription factors (blue) that directly oversee the expression of genes involved in the biosynthetic pathway (dashed arrows) (adapted from Collemare & Seidl, 2019).

1.6.2.2 Importance of population genomics in identifying chemical effectors

Functional studies for effector molecules are often limited to a single isolate or a few isolates, which may not fully consider the various genetic polymorphisms present among the different populations (Kuhnert & Collemare, 2022). However, it is important to note that different isolates of the same pathogen do not possess identical sets of pathogenicity factors. In the field of plant pathology, avirulence genes exhibit polymorphisms within populations due to their coevolution with host plants and their resistance genes (Frantzeskakis et al., 2020). Despite this understanding, there has been a notable lack of population genomics studies focused on secondary metabolite biosynthetic gene clusters (SM BGCs). These studies have only recently begun to emerge, opening new insights into the role of secondary metabolites in fungal pathogens and shedding light on the evolution of biosynthetic pathways in fungi **(Fig. 1.12)**. Population genomics-based research can be highly valuable in identifying novel effector genes or gene clusters that exhibit variations in presence or absence among different populations of a single species. Presence-absence variations (PAV) within secondary metabolite pathways have been frequently observed at the genus level, particularly in genera like *Aspergillus* (Kjærbølling et al., 2020; Theobald et al., 2018; Vesth et al., 2018) and *Fusarium* (Tralamazza et al., 2019; Villani et al., 2019), which has spurred comparative genomics studies between species with differing lifestyles or host specificities. Therefore, while we advocate for a comprehensive approach, we propose that utilizing population genomics to identify new effectors and host species-specific determinants, followed by the functional characterization of these potential molecules, could offer valuable insights into the molecular mechanisms underlying host-pathogen interactions.

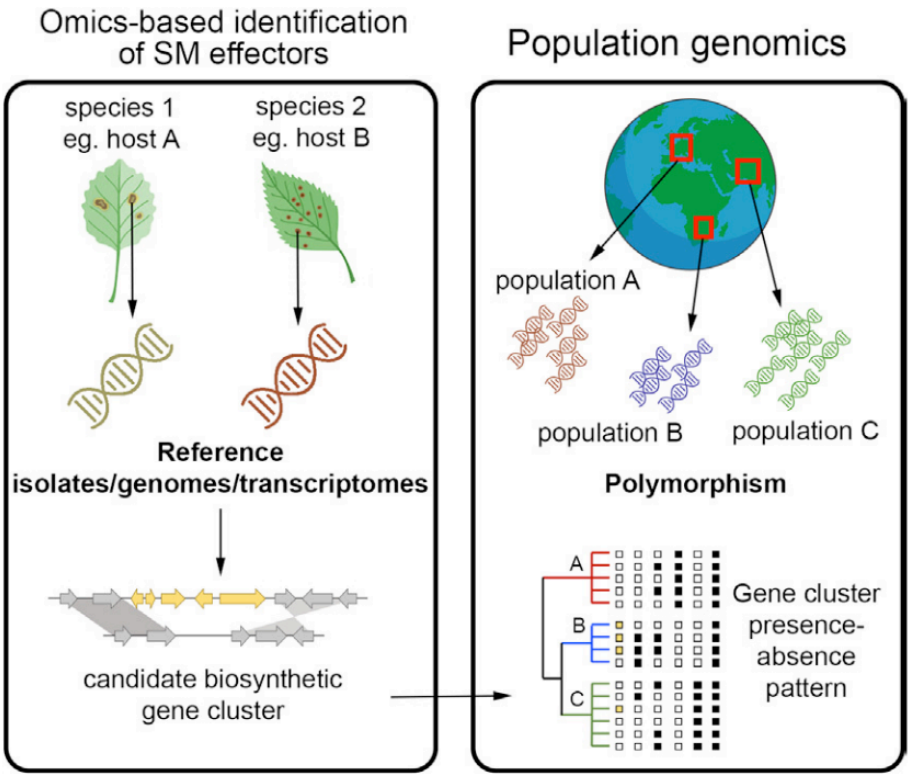

Figure 1.12: Population genomics to identify chemical effectors. The conventional approach involves the comparison of reference isolates to pinpoint biosynthetic pathways linked to specific characteristics, such as the host preference of a particular pathogenic fungus. However, this method can lead to inaccuracies when a species displays a population structure, as the set of secondary metabolites (SMs) may vary among individuals. In this instance, the potential gene cluster identified through reference isolates displays a pattern of presence and absence (highlighted in yellow) across different populations (adapted from Kuhnert & Collemare, 2022).

1.6.2.3 Known SMs of the blast fungus

M. oryzae a substantial number of genes related to secondary metabolism, with a reported count of 22 polyketide synthase (PKS) genes and eight non-ribosomal peptide synthetase (NRPS) genes (Collemare, Pianfetti, et al., 2008; Dean et al., 2005). However, it is worth noting that detailed characterization has only been accomplished for biosynthetic genes linked to few specific SM products in *M. oryzae*, namely melanin, pyriculols, pyrichalasin H, nectriapyrone and tenuazonic acid. Melanin plays a pivotal role in the infection strategy of the blast fungus. The inner melanin layer enables matured appressoria to withstand the turgor pressure and facilitates the puncturing of the plant cuticle. Mutant strains incapable

of melanin production have been observed to lose their pathogenicity on rice hosts (Chumley & Valent, 1990). *M. grisea*, a related rice blast fungus species, has been found to produce various phytotoxic secondary metabolites dependent on its in vitro cultivation conditions. For instance, *M. grisea* can produce a species-specific metabolite known as pyrichalasin. The virulence of *M. grisea* strains that infect Digitaria plants appears to correlate with the quantity of pyrichalasin H (Tsurushima et al., 2005). Pyriculol, another secondary metabolite, induces lesion formation on rice leaves and may contribute to virulence (Jacob et al., 2017). Tenuazonic acid (TA), a mycotoxin produced by several plant pathogenic fungi, can also be synthesized by the blast fungus. Interestingly, mycoviruses have been found to stimulate TA production. Application of TA on leaves resulted in disease resistance, as it led to the generation of reactive oxygen species by rice leaves, inhibiting pathogenic growth (Aver'yanov et al., 2007; Ninomiya et al., 2020).

Nectriapyrone production in *M. oryzae* can be induced by perturbing the two-component signal transduction system. Overexpression of two specific genes leads to an excess of nectriapyrone and its analogs. It's worth noting that nectriapyrone production is not essential for rice infection. The structure of nectriapyrone resembles that of germicidins produced by *Streptomyces* spp., and nectriapyrone has been found to inhibit the growth of *Streptomyces griseus*, suggesting a potential role in microbe-microbe interactions within the same environment (Motoyama et al., 2019).

The ACE1-derived secondary metabolite has an AVR effector function in *M. oryzae*, as it is necessary for full virulence on most host cultivars, except those carrying the resistance gene Pi33 (Böhnert et al., 2004; Collemare, Pianfetti, et al., 2008). Various fungal pathogens, including *M. oryzae*, are known to produce analogs of phytohormones such as jasmonic acid (JA), gibberellin, cytokinins and ethylene. Fungal analogs of JA and cytokinins are known to manipulate plant hormone-based defense signalling and subvert plant growth, respectively, during tissue colonization (Patkar & Naqvi, 2017; Shen et al., 2018). Interestingly, fungal derivative of JA produced by *M. oryzae* appears to play a significant role specifically during rice-blast fungus interactions (Patkar et al., 2015).

Overall, this chapter summarizes various aspects of *M. oryzae* population structure and possible molecular determinants responsible for the underlying genetic divergence and evolutionary mechanisms shaping its virulence and specialization on various hosts.