Physiological and Molecular Plant Pathology 95 (2016) 4–7

Contents lists available at ScienceDirect

Physiological and Molecular Plant Pathology

journal homepage: www.elsevier.com/locate/pmpp

Whole genome sequencing approaches to understand *Magnaporthe*-rice interactions☆

Ryohei Terauchi*, Hiroyuki Kanzaki, Koki Fujisaki, Hiroki Takagi, Akira Abe, Kentaro Yoshida, Yudai Okuyama, Muluneh Tamiru, Hiromasa Saitoh

Iwate Biotechnology Research Center, 22-174-4 Narita, Kitakami, Iwate 024-0003, Japan

ARTICLE INFO

Article history:
Received 16 February 2016
Received in revised form
16 March 2016
Accepted 17 March 2016
Available online 18 March 2016

Keywords:
Oryza sativa
Magnaporthe oryzae
Rice blast
Coevolution
Genomics

ABSTRACT

Due to the recent advances in DNA sequencing technologies, whole genome sequencing (WGS)-based approaches are now accelerating the pace of our research toward a better understanding of host-pathogen interactions. Using WGS-based methods, we have isolated three *avirulence* (*AVR*) genes: *AVR-Pia*, *AVR-Pii* and *AVR-Pik* from *Magnaporthe oryzae* and two cognate rice resistance (*R*-) genes: *Pia* and *Pii*. We briefly review our current understanding of the interactions between *AVR-Pia* and *Pia*, *AVR-Pik* and *Pik*, and *AVR-Pii* and *Pii*.

1. Introduction

Thanks to the recent development of next generation sequencing (NGS) technologies, whole genome sequencing (WGS) of organisms has become easier, allowing us to carry out genetic analysis with unprecedented speed and precision. This is enabling us to identify genetic changes responsible for particular phenotypes, including those related to plant defense and pathogen virulence.

Natural selection leaves its signature on the DNA sequences of the genes involved. WGS allows the detection of such selection signatures on a genome-wide scale. The burgeoning fields of evolutionary and population genomics have already started providing us with a wealth of new insights [1]. It is widely recognized that infectious diseases impose strong natural selection on the host organisms [2]. Hosts' defense also exerts strong natural selection on the pathogens. Such reciprocal selection between hosts and pathogens has led to co-evolution of the organisms involved, leaving signatures of selection in their genomes. Such selection signatures can readily be studied by WGS of the genomes of both hosts and pathogens (reviewed in Ref. [3]).

In this short note, we provide an overview of our application of WGS to understanding *Magnaprothe*-rice interactions and briefly discuss three examples of molecular interactions between *Magnaporthe oryzae avriulence* (*AVR*) genes and cognate rice *resistance* (*R*-) genes.

2. *Magnaporthe*-rice interactions

Rice blast caused by an Ascomycete, *Magnaporthe oryzae*, is the most devastating disease of rice worldwide [4]. Deployment of blast-resistant cultivars has been the most effective means of controlling the disease. For this purpose, rice breeders and geneticists have made considerable efforts to identify blast resistance genes. So far, more than 40 *R*-genes have been named, of which 25 have been cloned [5,6]. In addition, at least nine *AVR* genes have been cloned from *M. oryzae* [6].

Additionally, the cloning of nine pairs of *M. oryzae AVR* and the cognate rice *R*-genes has been reported (Table 1), providing opportunities for elucidating their molecular interactions. These genes have been isolated mainly by conventional map-based approaches. However, recently WGS-based approaches have been increasingly used for the same purpose.

☆ This article is part of a special issue entitled "The U.S.-Japan Scientific Seminar: Molecular Contact Points in Host-Pathogen Co-evolution".

* Corresponding author.
E-mail addresses: terauchi@ibrc.or.jp (R. Terauchi), hkanzaki@ibrc.or.jp (H. Kanzaki), k-fujisaki@ibrc.or.jp (K. Fujisaki), h-takagi@ibrc.or.jp (H. Takagi), a-abe@ibrc.or.jp (A. Abe), kentaro.yoshida@port.kobe-u.ac.jp (K. Yoshida), yokuyama@kahaku.go.jp (Y. Okuyama), olimt@ibrc.or.jp (M. Tamiru), saitoh@ibrc.or.jp (H. Saitoh).

http://dx.doi.org/10.1016/j.pmpp.2016.03.007
0885-5765/

Table 1
Summary of cloned *Magnaprothe oryzae AVR* genes and cognate rice *R*-genes.

Magnaporhte oryzae AVR gene			*Oryza* spp. *R-gene*		
Gene	Note	Reference	Gene	Note	Reference
AVR-Pita	metalloprotease-like	[7]	*Pita*		[13]
AVR1-CO39	Direct binding to NLR	[8]	*Pi-CO39* (=*Pia*)	Paired NLRs	[14] [15]
AVR-Pizt		[9]	*Pizt*		[16]
AVR-Pi9		[6]	*Pi9*		[17]
AVR-Pia	Direct binding to NLR	[10]	*Pia*	Paired NLRs	[15]
AVR-Pii		[10]	*Pii*	Paired NLRs	[18]
AVR-Pik	Direct binding to NLR	[10]	*Pik*	Paired NLRs	[19]
AVR-Pib		[11]	*Pib*		[20]
AVR-Pi54		[12]	*Pi54*		[12]

3. Isolation of *M. oryzae avirulence* (*AVR*) genes by using whole genome sequencing and association genetics

The first whole genome sequence of *M. oryzae* was published in 2005 for the laboratory strain 70-15 [21]. From the predicted 11,109 genes encoded in the genome, we selected 1032 genes encoding putative secreted proteins as effector candidates. Focusing on these 1032 genes, we tried to identify DNA polymorphisms that show association with AVR phenotypes among 23 Japanese *M. oryzae* isolates. These isolates were previously inoculated onto 11 rice cultivars or varieties, each harboring specific *R*-genes, allowing the presence and absence of cognate *AVR*-genes to each *R*-gene to be determined for each isolate. However, this approach failed to identify any association between DNA polymorphisms and the AVR phenotype. Accordingly, we reasoned that the *AVR* genes were not represented in the genome of the 70−15 reference isolate. We therefore decided to obtain the genome sequence of another isolate, Ina168, which is known to have at least 11 AVRs by inoculation studies. Using the 454 Next Generation Sequencing (NGS) platform, we performed re-sequencing of Ina168, and aligned the resulting sequence reads to the 70−15 reference genome. This analysis allowed us to identify a total of 1.68 Mb of Ina168-specific genomic regions that could not be aligned to the genome of 70−15. Gene prediction in the Ina168-specific regions identified 316 putative secreted protein genes. Using PCR primers generated for the 316 genes, we amplified the gene fragments from the 23 Japanese *M. oryzae* isolates and studied presence/absence-type DNA polymorphism of each gene. Presence/absence of the PCR products in three genes exhibited association with *AVR* phenotypes of *AVR-Pia*, *AVR-Pii* and *AVR-Pik* [10]. Genetic complementation studies validated the *AVR* genes. All the three AVRs were short proteins (AVR-Pia: 85 amino acids; AVR-Pii: 70 amino acids; AVR-Pik: 113 amino acids) with no known functional domains.

4. Isolation of rice resistance (*R*-) genes using WGS information

Pik, the cognate *R*-gene for *AVR-Pik*, was isolated by Ashikawa et al. [19] using map-based cloning. *Pik* comprises two tightly linked genes *Pik-1* and *Pik-2*, both encoding nucleotide binding leucine-rich repeat (NLR) proteins. *Pia* was isolated by a combination of association study and mutant screening [15]. The approximate position of *Pia* on rice chromosome 11 was previously known [22]. From this region, we identified 12 NLR-like *R*-gene analogs (*RGA*s). DNA polymorphism study of these *RGA* genes among seven rice cultivars with known *Pia* phenotype showed that DNA polymorphisms in three genes, *RGA3*, *4* and *5* exhibited perfect association with the phenotype. We then screened ethyl methane sulfonate (EMS) mutants generated by mutagenesis of the rice cultivar Sasanishiki harboring *Pia*, and identified two mutants that lost the *Pia* phenotype. DNA sequencing of *RGA3*, *4* and *5* from these mutants identified two independent mutations in the *RGA4* gene. Genetic complementation with the wild type allele of *RGA4* restored the *Pia* phenotype, suggesting that *RGA4* corresponds to *Pia*. However, when *RGA4* was transferred to the cultivar Hitomebore lacking *Pia*, it could not confer the *Pia* phenotype. The *Pia* phenotype was only obtained when another NLR gene, *RGA5*, located at close proximity to *RGA4*, was co-transformed with *RGA4*. This result suggested that *Pia* is comprised of two NLR genes, *RGA4* and *RGA5* [15].

The rice cultivar Hitomebore has the *Pii R*-gene. To isolate the *Pii* gene, we utilized an EMS-mutagenized population of Hitomebore and MutMap [23], a WGS-based method for simultaneous mapping and identification candidate mutations. EMS mutagenesis causes mainly transition nucleotide substitutions (G->A; C->T) resulting in single nucleotide polymorphisms (SNPs). Our mutagenesis procedure [24] was shown to cause on average 1500 homozygous SNPs per rice genome (370 Mb) in the M4 generation. MutMap is a mutation identification method in which WGS is applied to bulked-segregant analysis (BSA) [25,26]. In MutMap, we cross the mutant of interest to the parental line that was used for mutagenesis. If the mutation causing the phenotype is a recessive single SNP, the F2 progeny segregate to wild and mutant types in a 3:1 ratio. DNA from multiple (>20) individuals of the mutant type F2 progeny are pooled to make a bulked DNA. This DNA is subjected to WGS using the Illumina NGS platform and the resulting short sequence reads are aligned to the reference genome sequence of the parental line. Following crossing of the mutant to the parental line, all induced SNPs of the mutant will be inherited by the F1 in heterozygous state. By self-fertilization of the F1 individual, we obtain F2 progeny in which the expected frequency of the mutant-type allele is 0.5. However, if we focus only on the F2 progeny showing the mutant phenotype, they should all inherit the causative SNP in mutant-type/mutant-type homozogyous state, resulting in the frequency of the mutant allele being 1.0. MutMap identifies SNPs different from the parental line with a frequency of 1.0 in the DNA of the population of mutant-type progeny [23]. A screen of 3000 Hitomebore EMS mutant lines identified two independent mutants that lost the *Pii* phenotype. Using the two mutants, we applied MutMap-Gap, a modified version of MutMap, and identified the *Pii R*-gene encoding an NLR [18]. The *Pii* gene seems to be allelic to the previously described *Pi-5 R*-gene [27] and comprises a pair of NLR genes, *Pii-1* and *Pii-2*.

5. *AVR-Pia/Pia* interactions

Cesari et al. [14] showed by yeast two-hybrid (Y2H) assay that AVR-Pia directly binds the C-terminal domain of RGA5, a