A DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

QTL Mapping and Genome-Wide Association Study for *Phytophthora capsici* Resistance in Pepper *(Capsicum annuum)* and Comparative Genomic Analysis of *P. capsici* Isolates

고추(*Capsicum annuum*)의 역병 저항성을 조절하는
양적조절유전자좌 및 전장유전체연관분석과
역병균주의 비교유전체분석

AUGUST, 2019

MUHAMMAD IRFAN SIDDIQUE

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

A DISSERTATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

QTL Mapping and Genome-Wide Association Study for *Phytophthora capsici* Resistance in Pepper *(Capsicum annuum)* and Comparative Genomic Analysis of *P. capsici* Isolates

고추(*Capsicum annuum*)의 역병 저항성을 조절하는
양적조절유전자좌 및 전장유전체연관분석과
역병균주의 비교유전체분석

AUGUST, 2019

MUHAMMAD IRFAN SIDDIQUE

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

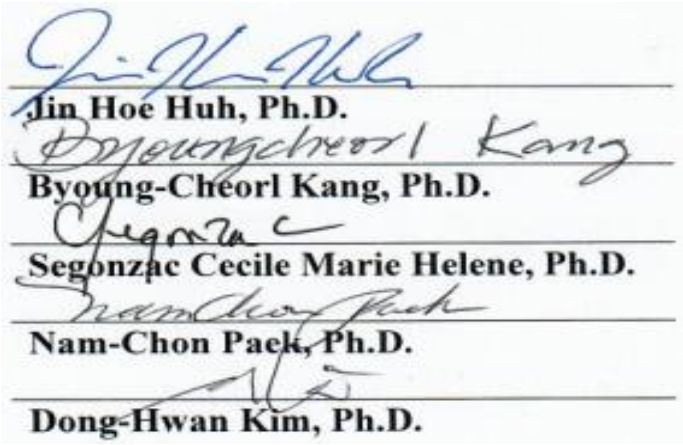
QTL Mapping and Genome-Wide Association Study for *Phytophthora capsici* Resistance in Pepper *(Capsicum annuum)* and Comparative Genomic Analysis of *P. capsici* Isolates

UNDER THE DIRECTION OF DR. BYOUNG-CHEORL KANG
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

BY
MUHAMMAD IRFAN SIDDIQUE

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
AUGUST, 2019

APPROVED AS A QUALIFIED DISSERTATION OF MUHAMMAD IRFAN SIDDIQUE
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
BY THE COMMITTEE MEMBERS



QTL Mapping and Genome-Wide Association Study for *Phytophthora capsici* Resistance in Pepper *(Capsicum annuum)* and Comparative Genomic Analysis of *P. capsici* Isolates

MUHAMMAD IRFAN SIDDIQUE

Department of Plant Science, Seoul National University

ABSTRACT

Phytophthora capsici is a globally prevalent devastating Oomycetes pathogen causing root rot in pepper. Genetic and genomic resources have recently been developed for *P. capsici* including a highquality reference genome, a dense genetic linkage map, and thousands of validated single nucleotide polymorphism (SNP) markers. *Capsicum annum* cv. CM334 is widely used as a source of resistance. A high-resolution linkage map constructed using 188 recombinant inbred lines (RILs) derived from a cross between resistant cv. CM334 and susceptible cv. ECW30R. Finally, retained 13,021 high-quality genotyping-by-sequencing derived single nucleotide polymorphism (SNP) markers were used to map resistance against low, medium and high virulent isolates of *P. capsici* in two different environments. Quantitative trait loci (QTL) analysis for *P. capsici* root rot (PcRR) mapped three major effect QTLs (*5.1*, *5.2* and *5.3*) on chromosome P5 conferring broad-spectrum resistance to PcRR isolates. In addition, QTLs with minor effects, specific to isolate and environment were detected on the

chromosomes P1, P2, P4, P7, P8 and P11. QTL analysis was complemented with a genome-wide association study (GWAS) of root rot resistance in a pepper core collection with 352 diverse accessions. A total of 507,713 SNPs derived from two genotyping-by-sequencing (GBS) libraries were used for GWAS analysis. GWAS detected 117 significant SNPs across the genome associated with resistant to PcRR, and the regions on chromosomes P5, P7, and P11 were co-located with QTLs identified in the present study. By leveraging the combined use of QTL mapping and GWAS, clusters of the nucleotide-binding site leucine-rich repeat (NBS-LRR) and receptor-like kinases (RLKs) like protein candidate genes were predicted within the QTL and GWAS regions for disease resistance mechanisms. Highly significant SNP markers and predicted candidate genes identified through QTL mapping and GWAS for PcRR resistance in present study, could accelerate the marker-assisted breeding and genomic selection for durable resistance in pepper by combining alleles of race-specific and non-race specific resistance in pepper.

In second chapter, genome sequencing and characterization was performed of three *P. capsici* isolates representing distinct virulence profiles. The *de novo* hybrid assemblies of three *P. capsici* isolates, using Illumina HiSeq and single-molecule real-time (SMRT) sequencing technologies were illustrated. In results, average number of 514 contigs with 50.96% of GC contents, 698,937 bp of N50 and 16,398 predicted genes were obtained. Genomic analysis discovered the huge number of genes encoding potential secreted effectors in the genomes, including average 60 RxLR domain containing effectors, 42 Crinklers (CRN), 536 CAZymes grouped into 7 families, and several apoplastic effectors, such as cytochrome P450, phytotoxins (PcF proteins), NPP1 families, LRR kinase as well as virulence and necrosis inducing proteins, in three *P. capsici* isolates. The comparative genomic analysis and GO term enrichment analysis for the polymorphism detection in various genes revealed, the numerous groups of

genes which showed 10 polymorphisms on one CDS sequence. In addition, the characterization of the virulence profiles of the isolates in laboratory and field experiments were assessed which showed a striking variation in virulence of three isolates. This study provides a genomic landscape of three *P. capsici* isolates for the comparative genomic analysis, which may help to identify the genes under positive and negative selections, homology analysis with other *Phytophhtora* spp. and effector assisted breeding. This characterization and genome information elucidates virulence and genome structural variations for *P. capsici* isolates, offering opportunities to better understand the virulence profiles of this oomycete pathogen, and presents possibilities for more detailed studies on genomics and virulence that may lead to future help in resistance incorporation and practical breeding applications.

Key words: Pepper, *Phytophthora capsici*, Genotyping-by-sequencing (GBS), Genome-wide association study (GWAS), Quantitative trait locus (QTL), Comparative genomic analysis, *De novo* hybrid assembly

Student number: 2015-30766

CONTENTS

LIST OF TABLES

Chapter I

Chapter II

LIST OF FIGURES

Chapter I

Chapter II

LIST OF ABBREVIATION

CTAB	Cetyltrimethylammonium Bromide
CV	Coefficient of variation
EC	ECW30R and CM334
GAPIT	Genome association and prediction integrated tool
GBS	Genotyping-by-sequencing
GWAS	Genome wide association study
H^2	Broad-sense heritability
LD	Linkage disequilibrium
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism
Pc	*Phytophthora capsici*
RILs	Recombinant inbred lines
RLK	Receptor like kinase
NBS	Nucleotide binding site
LRR	Leucine rich repeat
DPI	Days post inoculation
cM	CentiMorghan

GENERAL INTRODUCTION

The genus *Capsicum* has numerous common English names which contain chili pepper, hot pepper, bell pepper, sweet pepper, and chili belongs to the family Solanaceae. Occasionally, the plant is just called pepper. Approximately, this genus comprises of five domesticated species and twenty-two (22) wild species. The domesticated are *Capsicum annuum*, *Capsicum chinenses*, *Capsicum frutescens*, *Capsicum pubescens*., and *Capsicum baccatum* (Bosland and Votava, 2000. *Capsicum* species can be divided into diverse classes based on pod or fruit structures such as size, shape, color, pungency and flavor. In spite of their huge trait deviations, the worldwide commercially cultivated cultivars of peppers are *C. annuum* species (Bosland and Votava, 2000.

Botanically, *Capsicum* species are perennial shrubs in tropical climates, and they are commonly grown as annual and herbaceous crops in temperate climates and can also be grown as perennials under controlled environments that are climate-regulated (Olatunji and Afolayan, 2018). *Capsicum* spp. are diploids, generally having 24 chromosomes (n = x = 12) and several wild species containing of 26 chromosomes (n = x = 13). The most of the domesticated species associated to the first group (Bosland and Votava, 2000). Like most other vegetable crops, peppers required well-drained, moisture-holding loamy soil having an optimum growth and production temperatures between 18 to 30°C. However, the optimal temperatures for seed germination range from 25 to 30°C (Olatunji and Afolayan, 2018).

Peppers have progressively become essential ingredients of various cuisines in the world followed by the large acres of land dedicated to its cultivation in the countries such as Mexico, China, India, USA, Korea, and Africa (Siddique, 2012). According to nutritional fact sheet, pepper contains surplus amounts of vitamins and minerals such as Vitamin A, B2, C and

are rich in calcium, phosphorus, and potassium. A considerable share of land in pepper-producing countries is devoted to make chili powder. The increase in domestic and industrial demands of pepper has expanded its cultivation around the globe (Siddique, 2012). *C. annuum* L. includes a huge number of horticultural varieties (hot and sweet) and great economical important among cultivated species. It is popular vegetable in agriculture because of its high cash value to the growers and great taste, color, and medicinal values. *C. annuum* contains a variety of bioactive compounds and essential nutrients that exhibit various bioactivities such as antiviral, antioxidant, antimicrobial, anticancer, and anti-inflammatory activities. It has been evident that the place of origin of this species is Central and South America, and currently, it is widely cultivated on above 1.5 million hectares in several countries of the world (Russo, 2012). *C. annuum* needs tropical climates, because of it required a warm, humid climate to survive. It is a small erecting or spreading herb which grows up to 1 m in height. The leaf shapes are ovate, oblong-ovate, or ovate-lanceolate, 1.5–4 cm by 4–13 cm with entire margin, whereas the flowers are small, tinged purple, or white color. Mostly, fruits are red in color but many other types with orange, green, and yellow color also exist. The seeds have pale-yellow color, reniform, or discoid shape and are around 3–5 mm in size. *C. annuum* is day-neutral; nevertheless, photoperiodic reaction may be noticed in some forms (Hundal and Dhall, 2005).

Like other cultivated plants, peppers are also affected by many common pests and diseases that halt the yield and quality. Common pests include thrips, aphids, spider mites, nematodes, weevils, corn borer and white flies. Diseases caused by viruses, usually infected and spread by insect vectors, bacteria, fungi and oomycetes. Among these, the oomycete pathogen, *Phytophthora capsici* is the major limitation factor to pepper production throughout the pepper production regions of the world. It has been estimated that average annual losses were 42-49% between the years of 1998 - 2002 (Siddique, 2012). Pepper landraces have been

identified for genetic resistance to *P. capsici*. Utilizing these resources offer an ideal solution in managing this disease-causing pathogen as opposed to other management practices such as chemical sprays which may be harmful to the environment and create chemical-resistant pathogen strains. Several landraces of pepper have been identified for genetic resistance to *P. capsici*. Exploiting these resources in a practical breeding is most environmental-friendly and promising approach to manage this destructive pathogen as compared to controls, such as chemical and fungicide which are hazardous to environment (Barchenger et al. 2018).

On account of increasing demand, the pepper cultivation has been increased by 40 fold in the last decade. Even though the pepper cultivation is expanding, the diseases affecting the production of *Capsicum* are also increasing alarmingly. In particular, the prevalence of Phytophthora infection causes drastic yield loss in pepper. Therefore, in order to overcome the aforementioned issues, several breeding technologies integrating novel approaches employing the next-generation sequencing technologies have been regarded as the promising method. The availability of pepper genome sequence information had improved the molecular marker assisted breeding or genomics assisted breeding of pepper especially, for disease resistance. In recent days, understanding for the marker trait association has been increase and several genome-wide molecular markers have been identified in pepper. Different types of molecular marker systems have been described, among them, single nucleotide polymorphisms (SNPs) markers expedited the molecular breeding in pepper. Because of their abundant occurrence in the plant genomes, the application of SNPs has increased in the molecular marker assisted breeding. Furthermore, the SNP markers facilitate the understanding of genome-wide fine mapping of genes and linkage between a trait and marker and permits high resolution quantitative trait loci (QTL) mapping (Manivannan et al. 2018). Several SNPs marker have been developed in pepper linked to Phytophthora root rot but there is still room to develop

more closely linked and precise markers.

The virulence and pathogenicity of *P. capsici* has been described for Solanaceous, Cucurbitaceous, and Leguminous species of economic interest such as chili pepper, tomato, squash, cucumber, pumpkin snap bean, and several weeds associated with crop fields. Inoculation can be performed with active zoospores and mycelia. The ability of infection can be affected by the quantity of inoculum, the more the concentration of the inoculum the higher the rate of the infection. Humidity, temperature and the age of the fruit can affect the infection severity of *P. capsici.* Several *P. capsici* isolates has been classified into physiological races according to geographical areas and host (Castro-Rocha et al. 2014). Different virulence of *P. capsici* isolates from the same host has been reported in different species including peppers which may cause hurdle during resistant evaluation. The whole-genome sequences of several oomycetre pathogens have revealed mutual features of *Phytophthora* species genomes as well as exclusive aspects of pathogen evolution. Numerous patterns related to pathogenesis and virulence, such as genome reduction associated with biotrophy and repeat-driven virulence change, have developed from genome analysis and comparisons (Jiang and Tyler, 2012).

The purpose of this research was to uncover the genetic components underlying resistances in resistant sources as well as the virulence discrepancy among pathogen isolates using next generation sequencing technologies such as genotyping-by-sequencing, genome wide association analysis and whole genome sequencing that can facilitate breeding strategies in passing the resistant to modern cultivated varieties. The current study is unique in that it evaluated pepper resistance to multiple isolates of *P. capsici* exhibiting diverse virulence in multi-locations and uses the latest genetic-molecular tools and resources to reveal the molecular mechanism conferring resistance and virulence.

LITERATURE REVIEW

Phytophthora capsici an overview

Phytophthora derived from a Greek work which means "plant destroyer" and capsici added the pathogen as the "destroyer of *Capsicum*. The genus *Phytophthora* belongs to the group Oomycetes. In early reports, the genus was categorized under the Kingdom Fungi due to high morphological similarities of filamentous mycelium but later they were characterized as water molds (Agrios, 2005). The genus *Phytophthora* include many notorious species such as *Phytophthora infestans* which cause late blight of potatoes and was a major contributing factors of Irish famine in 19th century. *Phytophtora ramorum* and *Phytophthora sojae* which cause Sudden Oak death and root rot of Soybean, respectively. Counting them all, the genus *Phytophthora* is has undeniably been one of the most devastating groups of plant pathogens to the crop plants (Rehrigh et al. 2014).

P. capsici was first isolated from pepper in New Mexico Agricultural Research Station by Leonian in 1922 (Leonian, 1922). *P. capsici* belongs to oomycete, which physiologically and morphologically similar to fungi, but is phylogenetically different from them and more closely related to algae. *Phytophthora* species contain filamentous vegetative structure called mycelium, without hyphal septa (Erwin and Ribeiro, 1996). They are saprophytic and heterotrophic, and grow promptly in culture. *P. capsici* has both sexual and asexual production system. Three types of asexual spores are formed, chlamydospores, sporangia, and zoospores (Brett, 2002). Sporangia normally are ovoid shaped and have a protruding papilla at their apex. Sporangia can directly germinate to produce hyphae, or indirectly by releasing biflagellate zoospores, which are the major infection causing agents to the plant roots, especially when the soil is moist and flooded. Zoospores can transport in a film of water around soil particles to

reach and infect roots of plants. Mostly, zoospores encyst on the root surface of host plant from where the hyphae start penetrating the root directly from the adhesive cyst (Brett, 2002). Sporangia and zoospores can infect all part of the host plants after spreading by splashing of rainfall, running water or by irrigation. Chlamydospores are typically spherical shaped and pigmented, and have a thick cell wall to support in their role as survival structures for long duration. *P. capsici* is a heterothallic which requires isolate of different compatibility type (A1 and A2) in close proximity for sexual reproduction, which leads to the sexual crossing and producing of oospores (Brett, 2002; Erwin and Ribeiro, 1996). Oospores are dormant and resistant organelles that permit the long time survival of pathogen in soil plant debris, causing genetic variability of isolates through random genetic mating. Both mating types has been found in same region, but, until now, not detected in the same pepper production field (Brett, 2002). *P. capsici* causes damping off, root-rot as well as stem, leaf and fruit-blight. In pepper, the root rot syndrome caused by *P. capsici* is related with root darkening and lesions that can rapidly enlarge to girdle and kill the root. At seedling stage, damping off can kill plants two to 5 days after infection. In mature plants, root infections can cause in stunting, wilting, and eventually plant death in about two weeks. Root rot is the most damaging and economically important disease syndrome of pepper. Foliar blight symptoms show dark, water soaked dots on the leaves (Walker and Bosland, 1999). The infection starts with a tiny circular or irregular-shaped lesion on the leaves producing a “blistered” appearance. Afterwards, the lesions size increase getting dry, and bleach to a light tan. The disease develops towards the stem as a brownish, dark-green and water-soaked lesions. At final stage, the plant is defoliated with dried stems. Infected foliar will turn brownish or tan and defoliate later as infection spreads to the stem. Crown-rot and stem-blight symptoms of pepper are usually similar. These symptoms show typical black or purple lesions close to the soil line. The lesions promptly merge and girdle on the main branches of stem resulting in entire plant death (Erwin and Ribeiro, 1996).

The initial symptoms of fruit-blight show small, water-soaked, dull-colored acnes that can readily elongate under favorable conditions. Fruit-blight symptoms can continue to enlarge until most of the pepper pod is rotten, in result de-shaped and unmarketable fruit. Lesions normally appear at either the stem tips or the blossom tips of the fruit, and spread rapidly to the center of the fruit (Erwin and Ribeiro, 1996). The diseased tissue converts to dry and sunken, and often turn a tan or straw color.

Current management practices for *P. capsici* are avoidance, cultural, chemical, biological and planting resistant cultivars. These approaches include crop rotation, irrigation management, soil solarization and fungicide applications (Hausbeck and Lamour, 2004). But the most promising is planting the resistant cultivars to regional isolates. Normally, the management approaches target to reduce the losses associated with the pathogen because it very challenging to eliminate the *P. capsici*, once it has established (Lamour et al. 2011). Moreover, *P. capsici* can promptly move from one field to the neighboring field and readily epidemics in a certain region, especially when the irrigation system is surface water causing the dissemination of pathogen (Gevens et al. 2007). Humid and rainy weather events can initiate and increase new and extensive infestations. So the complete restriction of the *P. capsici* movement among sites is impossible, the finest approach to avoid *P. capsici* infection in horticultural crops is the breeding of resistant cultivars because it is most promising, less expensive, environmental friendly and a sustainable alternative to chemical applications and other management strategies (Hausbeck and Lamour, 2004).

Genetic basis and mapping *P. capsici* resistance in pepper

Genetic basis of Phytophthora resistance in *Solanaceae* is complex and intricate. Different genetic models have been reported for the late blight in potato and tomato against

Phytophthora infestans including single R gene mediated resistance and quantitative resistance (Vleeshouwers et al. 2011). Traditional breeding methods for incorporating resistance to *P. capsici* into adapted pepper germplasm has been an aim of many crop improvement and breeding programs. One major confront for pepper breeders is that diverse inheritance models have been described among the sources of resistance to *P. capsici*. Numerous sources of resistance to *P. capsici* has been described in capsicum (Table 1); among them a Mexican land race 'Criollo de Morelos-334' (CM334) has been extensively evaluated. CM334 has proven absolutely resistant to multiple races of *P. capsici* (Ortega et al. 1992; Reifschneider et al. 1992; Glosier et al. 2008; Bosland and Lindsey 1991).

Table 1. Reported *P. capsici* resistant accessions

Resistant Sources	**Reference**
'493-4-1-2'	(Yamakawa, et al., 1979)
Fyuco	(Barksdale, et al., 1984)
Grif 9109	Candole, et al., 2010
Line '29' (aka L29)	Ortega, et al., 1992
Mirasol	(Bartual, et al., 1991
NY07-8001	(Foster and Hausbeck, 2010)
P51	Bartual, et al., 1993)
Perennial (PI 631147)	Thabuis, et al., 2003)
Phyo 636	Palloix, et al., 1988)
PI 201232, PI201234	(Kimble and Grogan, 1960, Smith, et al., 1967)
Waxy Globe	(Saini and Sharma, 1978)
KC00807-1, KC01744, KC00937	(Mo, et al., 2014)
Criollo de Morelos 334 (CM 334)	Lefebvre and Palloix, 1996

However, the precise location of genomic locus is a labyrinth for researcher and still lacking any clue because of complicated mode of inheritance. The resistance intensity against *P. capsici* attributed to many diverse factors, for example plant age, inoculums density, soil moisture, temperature, pathogen isolate, inoculation and disease scoring methods (Lee et al. 2010; -Ortega et al. 1991; -Kim and Hwang 1989). Because of these factors incorporation of resistance to elite breeding material is complex and tricky (Thabuis et al. 2003). Different research reports attribute the resistance in two categories (Truong et al. 2012). Some researchers claimed that single or two dominant genes are the controlling factors for *P. capsici* resistance on the basis of inheritance study (Ortega et al. 1992; Monroy-Barbosa and Bosland, 2008; Smith et al. 1967; Sy et al. 2005). Contrarily, other research groups argued that resistance is controlled by multi-genes and involvement of epistasis effects has also been reported (Barksdale et al. 1984; Bartual et al. 1993; Bonnet et al. 2007; Palloix et al. 1988; Reifschneider et al. 1992; Thabuis et al. 2004a). This discrepancy among the inheritance study because of different environmental conditions, pathogenicity and virulence of isolates, tested breeding materials and resistance evaluation criteria (Oelke et al. 2003; Rehrig et al. 2014; Sy et al. 2008). Additionally, further intricacy is because of separate genes are involved in resistance to infection at different plants organ like root, foliar and fruits (Candole et al. 2010; Naegele et al. 2014). At the same time there are different physiological races have been reported according geographical regions (Table 2).

Table 2. *P. capsici* physiological races studies

Type of differential hosts	Number of differential hosts	Number of isolate tested	Isolates origian	Physiological races characterized	Reference
Commercial pepper	6	49	Georgia	10 root rot races	Yin et al., 2012
NM-RIL	26	20	Brazil	8 root rot races	da Costa anf bosland, 2012
NM-RIL	26	12	New Mexico, Turkey, Holland, and Argentina	12 foliar blight races	Monroy-Barbosa and Bosland, 2011
Commercial pepper lines and cultivars	31	4	Michigan, USA	4 root rot races	Foster et al., 2010
NM-RIL	26	17	New Mexico, California, and Holland	13 root rot races	Sy et al., 2008
Commercial pepper lines and cultivars	11	34	North Carolina, and Turkey California, New Mexico,	14 root rot races	Glossier et al., 2008
Commercial pepper lines and cultivars	8	5	Guangdong, China	2 root rot races	Li et al., 2007
Commercial pepper lines and cultivars	18	10	New Mexico, New Jersey, Italy, Korea, and Turkey	9 root rot races 5 foliar blight races	Oelke et al, 2003

QTL mapping and cloning by means of DNA markers are advantageous tools to determine the genetic architecture of complicated phenotypic traits. Furthermore, implementation of molecular markers delineated by QTL analysis expedites the effectiveness of breeding by providing opportunity for marker-assisted selection to locate particular trait of interest (Liu et al. 2014). Several QTLs have been reported conferring resistance to *P. capsici* in pepper, mainly with major effect on chromosome P5 (Bonnet et al. 2007; Kim et al. 2008; Lefebvre and Palloix, 1996; Liu et al. 2014; Mallard et al. 2013; Minamiyama et al. 2007; Ogundiwin et al. 2005; Pflieger et al. 2001; Quirin et al. 2005; Rehrig et al. 2014; Sugita et al. 2006; Thabuis et al. 2004b; Truong et al. 2012). Although there are many researchers have been reported about main affect QTL on P5, also two and three QTLs in close vicinity (Mallard et al. 2013; Rehrig et al. 2014) demonstrating that there is cluster for resistant loci, but not yet any reliable markers have been developed and validated for marker-assisted selection (MAS) of Phytophthora resistance in pepper predicting high accuracy. Utilization of MAS in crop breeding programs requires markers tightly linked to the desired loci to enhance the effectiveness of early selection. A sequence-amplified characterized region (SCAR) marker developed from random amplified polymorphic DNA (RAPD) marker OpD04.717 was mapped within 6cM of the peak of a major QTL present at chromosome P5 in the pepper genetic map AC99 (Quirin et al. 2005). Bacterial artificial chromosome (BAC)-derived markers converted from restriction fragment length polymorphism (RFLP) markers CD125 and CT211A linked to QTLs on chromosome P5 and P9 also have been reported (Kim et al. 2008). The SNP marker Phyto5NBS1 localizing in QTL5-1 has shown 90% accuracy for predicting resistant phenotypes against *P. capsici* with low virulent isolates, when validated by testing 100 F_1 resistant commercial cultivars (Liu et al. 2014).

However, until now, these publically accessible molecular markers are not applicable to predict resistance precisely, and certain level of phenotype and genotype disparity has been observed when they are applied in diverse germplasm. This phenotype–genotype mismatch reduces the selection effectiveness for marker assisted selection. Generally, it is challenging to incorporate *P. capsici* resistance into local high-yielding susceptible cultivars. Introducing resistance using, classic backcross methods, showed lower resistance the donor parent with threshold effects, which is probably because of the loss of secondary resistance genes (Palloix et al. 1990). Recurrent selection has been practiced to transfer polygenic resistance into well-adapted cultivars (Thabuis et al. 2004). Nevertheless, linkage drag concomitant with lower yield, undesirable shaped and fruit with less vigorous plants is a major constraint to wide adoption of resistant cultivars. In this scenario, finding new resistant sources by evaluating diverse pepper accessions with different *P. capsici* isolates, using revolutionary next generation sequencing technologies to develop high-throught put abundant molecular marker and complementation of traditional and advance breeding approaches can overcome the existing hurdles for the pepper crop improvement against *P. capsici*.

Table 3. Resistance QTLs studies against *P. capsici* in pepper

Resistance Trait	Population Type	Cross	No. of Individuals	Types of Markers	No. of Markers	Map Length	Chromosomes with Resistance QTL	Reference
Stem and root rot	Dihaploid	H3 x Vania	101	AFLP, RFLP, RAPD	135	1513 cM	P3a, P3b, P5, P7, P10a ,P11	(Thabuis, et al. 2003)
Stem and root rot	F2	Yolo Wonder x CM334	151	AFLP, RFLP, RAPD	154	1668 cM	LG1, P4, P5, P6, P11a, P12	(Thabuis, et al. 2003)
Stem and root rot	Dihaploid	Perennial x Yolo Wonder	114	AFLP, RFLP, RAPD	64	685 cM	P2, LG2, P5, P10	(Thabuis, et al. 2003)
Stem and root rot	Dihaploid	Perennial x Yolo Wonder	94	RFLP, RAPD	119	954.2 cM	LG8, LG1, BRUN, LG5, LG2, NOIR, POURPRE	(Lefebvre and Palloix, 1996)
Stem and root rot	F2	Yolo Wonder x CM334	151	RFLP	255	520 cM	P1b	(Pflieger et al. 2001)
Root rot and foliar blight	RIL (F7)	Psp-11 x PI201234	94	AFLP, RAPD, SSR, SCARs	144	1466.1 cM	LG1 (P7), LG2 (P1), LG3, LG4, LG5, LG6, LG7 (P12), LG8 (P6), LG9 (P9), LG11 (P11), LG15 (P5)	(Ogundiwin et al. 2005)
Root rot	F2	Joe E Parker x CM334	94	AFLP, RAPD, SSR, SCARs	113	1089.2 cM	LG1 (P9), LG2, LG4 (P11), LG5 (P5)	(Ogundiwin et al. 2005)
Stem and root rot	RIL (F5)	Yolo Wonder x CM334	297	AFLP, SSR, RFLP, SSAP	323	1857 cM	P1, P3, P4a, P5, P6, P11a	(Bonnet et al. 2007)
Root rot and damping-off	F2	CM334 x Chilsungcho	100	RFLP, WRKY, SSR	209	1482.3 cM	P5, P6, P8, P9-3	(Kim et al. 2008)

Whole Plant	RIL (F8)	YCM334 (F6 RIL derived from CM334 x Yolo Wonder) x CM334	126	AFLP, SSR, CAPS	249	1486.6 cM	P5, P10, P11, LG Pb, LG Pc	(Truong et al. 2012)
Root rot and damping-off	Dihaploid	Manganji x CM334	96	AFLP, SSR	369	878 cM	LG3, LG15	(Minamiyama et al. 2007)
Root rot	RIL (F8)	Early Jalapeno x CM334	66	SNP	3887		P1, P5	(Rehrig et al. 2014)
Root rot	RIL	YCM334 (F6 RIL derived from CM334 x Yolo Wonder) x CM334	128	SNP	44	131 cM	P5	(Liu et al. 2014)
Root rot	DH	K9-11 x AC2258	176	RAPD, AFLP	518	1100.5 cM	LG1, LG6, LG7	(Sugita et al. 2006)
Root rot	F2	Shanghaiyuan x PI201234	794	SSR, CAPS	719	51 cM	P5	(Wang et al. 2016)
Root rot	BC1, F2	NMCA10339 x CM334	222, 372, 259	SNP, SSR			p10	(Xu et al. 2016)

Next generation sequencing technologies in pepper breeding

By employing next-generation sequencing (NGS) technologies, many species containing economically important crops, have been subjected to whole-genome sequencing by de novo assembly and resequencing. Recently, sequencing technologies have advanced from genome sequencing projects using enormous parallel sequencing technologies. NGS technologies provide us with better opportunities for learning crop genomics and other post-genomics (transcriptomics, proteomics, metabolomics).

Pepper reference genome

The reference genome of *Capsicum annuum* cv. CM334 was sequenced with 186.6X genome coverage (Kim et al. 2014) and another sequence of cultivated pepper Zunla-1 also has been sequenced (Qin et al. 2014) using whole genome shot-gun approach in Illumina platform. The whole genome size of pepper is 3.48 Gb which is larger than the genome size of its closely relative tomato. The genome of *Capsicum annum* cv. CM334 comprise of 37,989 numbers of and 34,903 contigs. Furthermore, 34,903 genes with the average/total coding sequence span of 1,009.9/35.2 Mb were detected in the genome with 76.4% transposable elements.

Genotyping by sequencing in genetic mapping

Recent advancements in NGS technologies boosted the high throughput discovery of SNPs. Among various NGS technologies, genotyping by sequencing (GBS) is highly convenient and simple method. The GBS method integrates a multiplex sequencing approach for making reduced representative libraries for the Illumina NGS platform that uses an inexpensive barcoding system for high efficiency at a lower cost compared to other genotyping methods (Elshire et al. 2011). Using GBS, costs can be more lessened joined with imputation of missing internal SNPs in haplotype blocks. GBS significantly diminishes complexity of

genomes by using enzymes to cleave the DNA attached with DNA-barcoded adapters with small amounts of starting DNA (100-200 ng). Fractionated genomic DNA through restriction digestion decreases depiction through size selection or the explicit combinations of restriction enzymes for further targeting specific genomic regions of interest. The sequenced part of the genome is greatly consistent within a population because restriction sites are generally conserved across species. This makes the GBS procedure extremely suitable for experiments that need surveying large numbers of markers within a population for genetic mapping and population genomics. Additionally, by choosing proper restriction enzymes (e.g. methylation sensitive restriction enzymes), GBS can offer high SNP coverage in gene-rich regions of the genome in a very economical manner (Gore et al. 2009). Consequently, GBS has potential to add bulk of markers to existing or new mapping populations for detecting polymorphism, gene expression analysis and genotyping at relatively low cost (Poland et al. 2012; Spindel et al. 2013). These aspects make GBS an influential means for a number of plant genetic studies especially in *Capsicum*.

Genome wide association study an alternative of traditional mapping

The genetic sources of phenotypic variation have been a main emphasis for crop improvement intended at recognizing the causes of disease, improving agricultural productivity and understanding adaptive processes. Particularly in plants, quantitative trait loci (QTL) were mostly mapped in bi-parental crosses, but they were limited in allelic multiplicity and having inadequate genomic resolution. The detection of genetic recombinants unveiling phenotypic variation in a population is limited to the capacity to relate a particular genomic region to a certain phenotype because of the small number of recombinations that occur throughout each generation, deterring fine mapping with tinier intervals (Zheng et al. 2008). This constraint led to the development of association mapping. Genome wide association study (GWAS) provides

higher resolution than a linkage map by narrowing down the genetic and physical regions that contain the potential candidate genes due to the comparatively large number of recombinations that occur over an evolutionary process. GWAS, utilize the natural diversity bred by multigenerational recombination events in a panel (Yu and Buckler, 2006). This approach can provide enlarged resolution compared to linkage mapping of populations when enough molecular markers are provided to generate adequate data and greater coverage. GWAS entails up to several million SNPs depending on the decline of linkage disequilibrium (LD). SNP detection through GBS can provide an ample number of makers for GWAS (Poland and Rife, 2012). GBS through the NGS method has been used to examine numerous agronomic traits in specific breeding programs and genetic studies (Gore et al. 2009). Nevertheless, successful GWAS need preceding information of LD decay and an awareness of the population under examination. If evidence on LD decay is not accessible, the chance is there that target loci are situated within gaps, which do not encompass appropriate marker density, could not be omitted even though GBS appears to deliver a much higher density of markers than required for GWAS. Thus, it may be convenient to test additional enzyme combinations for GBS for use in a large association panel. GWAS has been used successfully for discovering QTLs for various traits in diverse genetic materials (Wang et al. 2014). Recently, GWAS has been employed to locate genes of different traits in various crops; for example, in Arabidopsis quantitative trait locus for climate sensitivity while in rape seed, for yield, flowering time, seed quality and resistance to *Sclerotinia sclerotiorum*. However, there is no study about GWAS being used to report resistance against *P. capsici* in pepper.

Genomic analysis of *Phytophthora spp*.

Most of the current understanding of the evolution and mechanisms of virulence in oomycetes specially *Phytophthora* spp. has built on expressed sequence tag (EST) and genome sequencing completed over the past decade. Genome sequences include those of the late blight pathogen *P. infestans* (240 Mb) (Haas et al. 2009) *P. sojae* (95 Mb) (Tyler et al. 2006), sudden oak death pathogen *P. ramorum* (65 Mb) (Tyler et al. 2006).

***P. capsici* reference genome**

The reference genome of *P. capsici* is distinctive for the *Phytophthora* genus in size and content and offers a valuable resource for discovering the foundations of a broad host range oomycete (Lamour et al. 2012). A partially inbred stain of *P. capsici* was sequenced using 454 FLX and Titanium technology (30X paired and single reads) complemented by Sanger reads (5X) (Lamour et al. 2012). The *P. capsici* genome assembly covered the genome size of 64 Mb with 56 Mb (87.5%) of non-gap sequence. Approximately, all the genes were assembled as expected by the capture of 97% of *P. capsici* full-length cDNAs and 93% of highly conserved central eukaryotic orthologs. The assembly encompasses of 917 scaffolds with half of the genome (N50) covered in 29 scaffolds of size 706 kb or larger. A high density SNP-based genetic linkage map further arranged and oriented 84% of the genome assembly and about 90% of the predicted genes into 18 linkage groups. In total, 17,123 genes and 2,682 transposable element-like genes were discovered with 89% having cDNA-based or protein evidence. Around 86% of the *P. capsici* genome assembly was non-repetitive (Lamour et al. 2012).

The size of the *P. capsici* genome is intermediary among the sequenced oomycetes and comparable to *P. ramorum* (Table 4). Repeated regions composition was quite similar to other *Phytophthora* species with 84% of the repeat sequences being long terminal repeat retrotransposons, among them, 57% were Gypsy elements. Gene contents were also

proportionate to other *Phytophthora* sequenced genomes (Table 4). The gene density in non-repetitive noticed around 268 genes per Mb, was greater in *P. capsici* compared to the other *Phytophthora* species (Table 4). Genetic arrangement was also similar with two-thirds of the genes were detected in conserved gene-rich blocks. Over 75% of the genes could be clustered into potential gene families (Lamour et al. 2012).

P. capsici and other accessible sequenced genomes of *Phytophthora* species showed high level of synteny. Scaffold-level synteny was approximately complete for *P. capsici* compared to *P. ramorum* and *P. sojae* with negligible duplications, re-arrangements, and non-syntenic stretches. *P. capsici* and *P. infestans* also displayed high levels of synteny though 20% of the *P. infestans* genome did not contained syntenic genes in *P. capsici* (Lamour et al. 2012). Novel genes were detected in *P. capsici* genome. At a peptide level, total 365 genes did not contain homolog in the genome assemblies of *P. sojae*, *P. infestans* and *P. ramorum* (Lamour et al. 2012).

Virulence mechanism in *Phytophthora* species

The primary theme that come forth from the genomic studies is that virulence in oomycetes depends comprehensively, on large, promptly diversifying protein families (Tyler et al. 2006; Haas et al. 2009). These families comprise hydrolytic enzymes and inhibitors, extracellular toxins, and effector proteins that can penetrate into the cytoplasm of plant cells (Tyler et al. 2009). This differs with typical fungal pathogens in which secondary metabolite toxins play a major role, and virulence proteins families are less involved and diversified (Tyler et al. 2009).

Plants defend themselves from infection by pathogens like oomycetes and others through various sets of defenses that include chemical and physical barriers, antimicrobial

chemicals [phytoalexins, phytoanticipins, and reactive oxygen species (ROS)], antimicrobial enzymes and peptides, and against bio-trophic and hemi-bio-trophic pathogens by programmed cell death (PCD) (Jones and Dang, 2006).

For the successful survival of pathogens, including oomycetes, they must avoid, overcome, or tolerate these defenses, and obtain nutrition from the host (Torto-Alalibo et al. 2010). Usually, plant defenses consist of two intersecting modules by the microbial molecules that stimulate them. Generally existing microbial molecules that induce defenses in a wide variety of plants are referred to as microbe- (or pathogen-) associated molecular patterns (MAMPs or PAMPs) and the responses triggered by them as MAMP- (or PAMP-) triggered immunity (MTI or PTI). However, plants also compose intracellular receptors that can indirectly or directly perceive the presence of intracellular pathogen effectors, subsequent in effector-triggered immunity (ETI). Most plant associated pathogens including *Phytophthora* have evolved protein or chemical effectors that can suppress or reprogram the PTI and ETI (Jones and Dang, 2006). The effectors proteins from the *Phytophthora* species could infect outside the plant cell, targeting the apoplast, or at plasma membrane of the plant, also they may penetrate inside the cytoplasm of the plant cell to alter its functioning (Torto-Alalibo et al. 2019). Effectors secretions into apoplast can be done directly from the hyphae of pathogen, or with specialized intercellular structure hausroria.

Extra and intra cellular effectors in *Phytophthora spp.* genomes

Various groups of extra cellular proteins that activate cell death in host plants have been detected in *Phytophthora* species genomes, including the NLPs (necrosis and ethylene inducing peptide-like proteins) (Qutob et al. 2006) (60), the Scr family (Liu et al. 2005) and the PcF family (Orsomando et al. 2011).

Plant tissues, particularly the intercellular apoplastic regions, are abundant in complex carbohydrates. Therefore, oomycete genomes, like other plant pathogens, are found to be with the genes encoding a wide range of carbohydrate degrading enzymes, including pectate lyases, pectin esterases, cellulases and glucanases (Haas et al. 2009; Tyler et al. 2006). Additionally, a rich amount of lipases and proteases is also secreted. Nearly 30 to 60 protein families in each class were predicted to be secreted by each *Phytophthora* species (Haas et al. 2009; Tyler et al. 2006).

Since oomycetes emit a large set of virulence proteins, for the defense plants produce the proteases to degrade the virulence proteins. Consequently, oomycetes are detected with the secretion of proteinase inhibitors against the plant proteins. The genome of *Pythium* and *Phytophthora* contain 38 to 18 genes functioning for the secretion of inhibitors of serine and cysteine proteases (Haas et al. 2009; Levesque et al. 2010; Tyler et al. 2006).

The largely studied classes of virulence proteins to emerge from the genome sequencing of *Phytophthora* species are RxLR effectors, called for a conserved N-terminal amino acid sequence motif (arginine, any amino acid, leucine, arginine) (Kale and Tyler, 2011; Tyler, 2009). The *Phytophthora* genomes sequencing revealed about 350 to 550 genes coding for these instantly evolving secreted proteins (Haas et al. 2009; Jiang et al. 2008; Tyler et al. 2006). Pathogens belongs to *Phytophthora* species produce a second main class of secreted effector proteins named crinklers, called for their function to initiate crinkling and necrosis upon overexpressing in transient expression assays. Crinkler effectors are considered to the most vastly expressed genes both prior to and during infection (Jiang and Tyler, 2012).

The whole-genome sequences of *Phytophthora* pathogens have discovered common features of among the genomes as well as distinctive features of pathogen evolution. Numerous patterns associated with pathogenesis, such as genome reduction related to bio-trophy and

repeat-driven virulence variation, have emerged from genome comparisons and analysis. The most of the sequenced *Phytophthora* species display different lifestyles. A huge fraction of their genes are conserve to each pathogen family. Nevertheless, a mutual proteome can be detected by ortholog analysis throughout the phylogenetically different species (Tyler et al. 2006). The *Phytophthora* species *P. sojae*, *P. infestans*, and *P. ramorum* prorate a core set of 8,492 ortholog clusters (Haas et al. 2009). Some *Phytophthora* species exhibit a compact genome (<50Mb) (Levesque et al. 2010), while others, like *P. infestans* has the largest and highly complex genomes (>200 Mb) (Haas et al. 2009). All sequenced *Phytophthora* species genomes encode close numbers of genes, ranging from 14,000 to 19,000; but there is a large alteration in the repeat content that cause for the genome size differences. Nearly 75% of *P. infestans* genome is made up of repeats (Haas et al. 2009; Kemen et al. 2011). Large part of repetitive regions of these big genomes are very dynamic and likely to go under evolutionary changes. In summary, the pathogenesis mechanisms in *Phytophthora* species depend mainly on large protein families. These promptly diversifying families consist of extracellular toxins, hydrolytic enzymes, inhibitors, and effector proteins that can penetrate to the cytoplasm of plant cells. Thus, understanding the mechanisms of *Phytophthora* isolates virulence and adaptability by exploiting the genome alteration is important for planning management strategies to control *Phytophthora* diseases.

Table 4. Summary of the *Phytophthora* species genomes

Organism	Genome Size	Genes	Reference
P. ramorum	65 Mb	14,451	Tyler et al. 2006
P. sojae	95 Mb	16,988	Tyler et al. 2006
P. infestans	240 Mb	17,797	Haas et al. 2009
P. cinnamomi	58 Mb	26,132	Studholme et al. 2016
P. capsici	64 Mb	17,123	Lamour et al. 2012

REFERENCES

Agrios, G., N. Plant pathology / George N. Agrios Elsevier Academic Press. (2005).

Barchenger, D. W., Lamour, K. H. & Bosland, P. W. Challenges and strategies for breeding resistance in *Capsicum annuum* to the multifarious pathogen, *Phytophthora capsici*. *Front. Plant Sci*. **9,** 628 (2018).

Barksdale, T. H., Papavizas, G. C., & Johnson, S., A. Resistance to foliar blight and crown rot of pepper caused by *Phytophthora capsici*. *Plant Dis*. **68,** 506-509 (1984).

Bartual, R., Lacasa, A., Marsal, J. I & Tello, J. C. Epistasis in the resistance of pepper to phytophthora stem blight (*Phytophthora capsici* L.) and its significance in the prediction of double cross performances. *Euphytica* **72,**149-152 (1993).

Bonnet, J., Danan, S., Boudet, C., Barchi, L., Sage-Palloix, A. M., & Caromel, B. *et al.* Are the polygenetic architectures of resistance to *Phytophthora capsici* and *P. parasitica* independent in pepper? *Theor. Appl. Genet*. **115,** 253-264 (2007).

Bosland, P. W., & Lindsey, D., L. A seedling screen for Phytophthora root rot of pepper, *Capsicum annuum*. *Plant Dis*. **75,** 1048-1050. (1991).

Bosland, P. W., & Votava, E., J. Peppers: Vegetable and spice capsicums. Wallingford, CT: CABI Publishing p. 204 (2000).

Candole, B. L., Conner, P. J., and Ji, P. Screening *Capsicum annuum* accessions for resistance to six isolates of *Phytophthora capsici*. *HortScience* **45,** 254-259 (2010).

Castro-Rocha A. *et al.* Traditional and Molecular Studies of the Plant Pathogen *Phytophthora capsici*: A Review. *J Plant Pathol Microb* **5,** 245 (2014).

Erwin, D. C., & Ribeiro, O. K. Phytophthora Disease Worldwide. St Paul, MN: The American Phytopathological Society. (1996).

Gevens, A. J., Donahoo, R. S., Lamour, K. H., & Hausbeck, M. K. Characterization of *Phytophthora capsici* from Michigan surface irrigation water. *Phytopathology* **97,** 421-428 (2007).

Glosier, B. R., Ogundiwin, E. A., Sidhu, G. S., Sischo, D. R., & Prince, J. P. A differential series of pepper (*Capsicum annuum*) lines delineates fourteen physiological races of *Phytophthora capsici*. *Euphytica* **162,** 23-30. (2008).

Haas B. J. *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461,** 393-98 (2009).

Hausbeck, M. K. & K. H. Lamour. *Phytophthora capsici* on vegetable crops: Research progress and management challenges. *Plant Dis*. **88,** 1292-1303 (2004).

Hundal, J. S, & Dhall R. K. Breeding for Hybrid Hot Pepper. *J. New Seeds.* **6,** 31-50 (2005).

Jiang R. H. Y, Tripathy. S., Govers, F. & Tyler, B. M. RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving super-family with more than 700 members. *Proc. Natl. Acad. Sci*. **105,** 4874-79 (2008).

Jiang R. H. Y., & Tyler B. M. Mechanisms and evolution of virulence in oomycetes. *Annu. Rev. Phytopathol*. 50: 295–318 (2012).

Jones J. D. & Dangl J. L. The plant immune system. *Nature* **444,** 323-29 (2006).

Kale, S. D. & Tyler B. M. Entry of oomycete and fungal effectors into plant and animal cells. *Cell. Microbiol*. **13,** 1839-48 (2011).

Kemen E. *et al.* Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of Arabidopsis thaliana. *PLoS Biol*. 9:e1001094 (2011).

Kim, Y. J. & Hwang B. K. Expression of age-related resistance in pepper plants infected with *Phytophthora capsici*. *Plant Dis* **73,** 745-747 (1989).

Kim, H. J.et *al.* BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). *Theor. Appl. Genet*. **118,** 15-27. (2008).

Lamour, K. *et al.* Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Mol. Plant Microbe Interact*. **25,** 1350-1360 (2012).

Lamour, K. H., Stam, R., Jupe, J., & Huitema, E. The oomycete broadhost-range pathogen *Phytophthora capsici*. *Mol. Plant Pathol*. **13,** 329-337 (2011).

Lee, S. J., Park, Y. J., Kim, H. T., & Kim, B. S. The race differentiation of *Phytophthora capsici* in Korea. Res. *Plant Dis*. **16,** 153-157 (2010).

Lefebvre, A., & A. Palloix. Both epistatic and additive effects of QTLs are involved in polygenic induced resistance to disease: A case study, the interaction pepper *Phytophthora capsici* Leonian. *Theor. Appl. Genet*. **93,** 503-511 (1996).

Leonian, L. H. Stem and fruit blight of peppers caused by *Phytophthora capsici* sp. nov. *Phytopathology* **12,** 401-408 (1922).

Levesque, C. A. *et al.* Genome sequence of the necrotrophic plant pathogen Pythium ultimum reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol*. **11,** R73 (2010).

Liu, Z. *et al.* Patterns of diversifying selection in the phytotoxin-like scr74 gene family of

Phytophthora infestans. *Mol. Biol. Evol*. **22,** 659-72 (2005).

Liu, W. Y. *et al.* Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. *Theor. Appl. Genet*. **127,** 2503-2513. (2014).

Mallard, S., Cantet, M., Massire, A., Bachellez, A., Ewert, S. & V. Lefebvre. A key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to *Phytophthora capsici*: A valuable locus for pepper breeding. *Mol. Breed*. **32,** 349-364. (2013).

Minamiyama, Y., Tsuro, M., Kubo, T. & M. Hirai. QTL analysis for resistance to *Phytophthora capsici* in pepper using a high density SSR-based map. *Breed. Sci*. **57,** 129-134. (2007).

Monroy-Barbosa, A., & Bosland, P. W. Genetic analysis of Phytophthora root rot race-specific resistance in chile pepper. *J. Am. Hortic. Sci*. **133,** 825-829 (2008).

Naegele, R. P. *et al.* QTL mapping of fruit rot resistance to the plant pathogen *Phytophthora capsici* L. In a recombinant inbred line *Capsicum annuum* L. population. *Phytopathology* **104,** 479-483 (2014).

Oelke, L. M., Bosland, P. W., & Steiner, R. Differentiation of race specific resistance to Phytophthora root rot and foliar blight in *Capsicum annuum*. *J. Am. Soc. Hortic. Sci*. **128,** 213-218 (2003).

Ogundiwin, E. A. *et al.* Construction of 2 intraspecific linkage maps and identification of resistance QTL for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum* L.) *Genome* **48,** 698-711 (2005).

Olatunji, T. L. & Afolayan A. J. The suitability of chili pepper (*Capsicum annuum* L.) for alleviating human micronutrient dietary deficiencies: A review. *Food Sci Nutr.* **6,** 2239-2251 (2018).

Orsomando, G., Brunetti, L., Pucci, K., Ruggeri, B., & Ruggieri, S. Comparative structural and functional characterization of putative protein effectors belonging to the PcF toxin family from *Phytophthora* spp. *Protein Sci*. **20,** 2047-599 (2011).

Ortega, R. G., Espanol, C. P., & Zueco, J. C. Genetic relationships among four pepper genotypes resistant to *Phytophthora capsici*. *Plant Breed*. **108,** 118-125 (1992).

Ortega, R. G., Espanol, C. P., & Zueco, J. C. Genetics of resistance to *Phytophthora capsici* in the pepper line 'SCM-334'. *Plant Breed*. **107,** 50-55. (1991).

Palloix, A., Daubeze, A. M., & Pochard, E. Phytophthora root rot of pepper influence of host genotype and pathogen strain on the inoculum density disease severity relationships. *J. Phytopathol*. **123,** 2533. (1988).

Palloix, A., Daubeze, A. M., Phaly, T., & Pochard, E. Breeding transgressive lines of pepper for resistance to *Phytophthora capsici* in a recurrent selection system. *Euphytica* **51,** 141-15 (1990).

Pflieger, S., Palloix, A., Caranta, C., Blattes, A., & Lefebvre, V. Defense response genes co-localize with the quantitative disease resistance loci in pepper. *Theor. Appl. Genet*. **103,** 920-929 (2001).

Poland, J. A., & Rife, T. W. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* **5,** 92-102 (2012).

Quirin, E. A. *et al.* Development of sequence characterized amplified region (SCAR) primers for the detection of *Phyto.5.2*, a major QTL for resistance to *Phytophthora capsici* Leon. in pepper. *Theor Appl Genet* **110,** 605-612 (2005).

Qutob, D., *et al.* Phytotoxicity and innate immune responses induced by Nep1-like proteins. *Plant Cell* 18:3721–44 (2006).

Rehrig, W. Z., Ashrafi, H., Hill, T., Prince, J., & Van Deynze, A. CaDMR1 cosegregates with QTL Pc5.1 for resistance to *Phytophthora capsici* in pepper (*Capsicum annuum*). *Plant Genome* **7,** 1-12 (2014).

Reifschneider, F. J. B., Boiteux, L. S., Dellavecchia, P. T., Poulos, J. M., & Kuroda, N. Inheritance of adult plant resistance to *Phytophthora capsici* in pepper. *Euphytica* **62,** 45-49 (1992).

Russo, V. M. Peppers, botany production and uses, CABI Publishing p. 14 (2012).

Siddique, M. I. Evaluation and characterization of Phytophthora resistant breeds of Yeongyang local cultivars of pepper, *Capsicum annuum* L. Master Dissertation. (2012).

Smith, P. G., Kimble, K. A., Grogan, R. G., & Millett, A. H. Inheritance of resistance in pepper to Phytophthora root rot. *Phytopathology* **57,** 377-379 (1967).

Spindel, J. *et al.* Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet*. **126,** 2699-2716 (2013).

Sugita, T. *et al.* QTL analysis for resistance to Phytophthora blight (*Phytophthora capsici* Leon.) using an intraspecific doubled-haploid population of *Capsicum annuum*. *Breed Sci* **56,** 137-145 (2006).

Sy, O., Steiner, R. & Bosland, P. W. Inheritance of Phytophthora stem blight resistance as compared to Phytophthora root rot and Phytophthora foliar blight resistance in *Capsicum annuum* L. *J. Am. Hortic. Soc*. **130,** 75–78 (2005).

Thabuis, A. *et al.* Marker-assisted introgression of 4 *Phytophthora capsici* resistance QTL alleles into a bell pepper line: validation of additive and epistatic effects. *Mol. Breed*. **14,** 9-20 (2004b).

Thabuis, A. *et al.* Phenotypic and molecular evaluation of a recurrent selection program for a polygenic resistance to *Phytophthora capsici* in pepper. *Theor. Appl. Genet*. **109,** 342-351 (2004).

Thabuis, A., Palloix, A., Pflieger, S., Daubeze, A. M., Caranta, C., and Legebvre, V. Comparative mapping of Phytophthora resistance loci in pepper germplasm: evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. *Theor. Appl. Genet*. **106,** 1473-1485 (2003).

Torto-Alalibo, T. *et al.* Unifying themes in microbial associations with animal and plant hosts described using the gene ontology. *Microbiol. Mol. Biol. Rev*. **74,** 479-503 (2010).

Torto-Alalibo, T., Collmer, C. W., Lindeberg, M., Bird, D., Collmer, A., & Tyler B. M. Common and contrasting themes in host cell–targeted effectors from bacterial, fungal, oomycete and nematode plant symbionts described using the gene ontology. *BMC Microbiol*. **9,** (2009).

Truong, H. T. H. et al. Identification of isolate-specific resistance QTLs to phytophthora root rot using an intraspecific recombinant inbred line population of pepper (*Capsicum annuum*). *Plant Pathol* **61,** 48-56 (2012).

Tyler, B. M, *et al.* Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313,** 1261-66 (2006).

Tyler, B. M. Entering and breaking: virulence effector proteins of oomycete plant pathogens. *Cellular Microbiol*. **11,** 13-20 (2009).

Tyler, B. M. Effectors in oomycete genetics and genomics: Diversity, Interactions and Research Tools, ed. K Lamour, S Kamoun, pp. 361-85. Hoboken, New Jersey: Wiley-Blackwell (2009).

Tyler, B. M. Molecular basis of recognition between Phytophthora pathogens and their hosts. *Annu. Rev. Phytopathol*. **40,** 137-167 (2002).

Vleeshouwers, V. G. A. A. Understanding and exploiting late blight resistance in the age of effectors. *Annual Review of Phytopathology*. **49,** 507-31 (2011).

Walker, S. J. and Bosland, P. W. Inheritance of Phytophthora root rot and foliar blight resistance in pepper. *J. Amer. Soc. Hort. Sci.* **124,** 14-18 (1999).

Wang, C. *et al.* Genome-wide association study of blast resistance in indica rice. *BMC Plant Biology*, **14,** 311 (2014).

CHAPTER I

Identifying Candidate Genes for *Phytophthora capsici* Resistance in Pepper (*Capsicum annuum*) Via Genotyping-By-Sequencing-Based QTL Mapping and Genome-Wide Association Study

The research described here has been published in *Scientific reports journal*

ABSTRACT

Phytophthora capsici (Leon.) is a globally prevalent, devastating oomycete pathogen that causes root rot in pepper (*Capsicum annuum*). Several studies have identified quantitative trait loci (QTL) underlying resistance to *P. capsici* root rot (PcRR). However, breeding for pepper cultivars resistant to PcRR remains challenging due to the complexity of PcRR resistance. Here, I combined traditional QTL mapping with GWAS to broaden our understanding of PcRR resistance in pepper. Three major-effect loci (*5.1*, *5.2*, and *5.3*) conferring broad-spectrum resistance to three isolates of *P. capsici* were mapped to pepper chromosome P5. In addition, QTLs with epistatic interactions and minor effects specific to isolate and environment were detected on other chromosomes. GWAS detected 117 significant SNPs across the genome associated with PcRR resistance, including SNPs on chromosomes P5, P7, and P11 that colocalized with the QTLs identified here and in previous studies. Clusters of candidate nucleotide-binding site-leucine-rich repeat (NBS-LRR) and receptor-like kinase (RLK) genes were predicted within the QTL and GWAS regions; such genes often function in disease resistance. These candidate genes lay the foundation for the molecular dissection of PcRR resistance. SNP markers associated with QTLs for PcRR resistance will be useful for marker-assisted breeding and genomic selection in pepper breeding.

INTRODUCTION

Pepper is an economically important vegetable crop grown in tropical and temperate regions for fresh and processed use. *Phytophthora capsici* is an invasive oomycete, poses a serious threat to pepper production across the globe (Bosland and Lindsey 1991). The pathogen infects pepper plant at almost all growth and developmental stages causing damping off, root rot, stem rot, collar rot, fruit rot and foliar blight. *P. capsici* root rot (PcRR) is a most devastating disease causing 100% yield losses under favorable environmental conditions (Foster and Hausbeck 2010). The pathogen may enter into the plant system through roots or stem collar causing water-soaked lesions development and girdling of the stem, eventually leading to the plant wilting and death. The broad ranges of host species including members of *Solanaceae*, *Fabaceae*, and *Cucurbitaceae* as well as soil born nature and random mating potential of *P. capsici* isolates render it more difficult to control. Severe epidemic of *P. capsici* are favored by humid and warm conditions (25-28˚C) leading to the high yield losses (41-49%) acceding 100 billion USD in both field and greenhouse cultivations (Kim and Hwang 2012; Bosland and Lindsey 1991). Cultural practices and chemical control measures, including the application of fungicides, such as phenylamide and metalaxyl has proven to be ineffective (Lamour and Hausbeck 2000). Use of resistant cultivars is the most sought-after due to their eco-friendliness and cost effectiveness for managing PcRR.

Resistance to PcRR is influenced by several factors including environmental cues, virulence of different *P. capsici* isolates, existence of various physiological races and source of resistance (Foster and Hausbeck 2010; Rehrig et al. 2014). *P. capsici* with short life cycles and rapid evolution of complex races offer them a selective advantage over their relatively slower evolving plant hosts in the evolutionary arms race (Hausbeck and Lamour 2004; Kamoun et al.

2015). Moreover, separate genes are involved for resistance to infection from plants organ like root, leaf and fruits. More than 45 physiological races have been described corresponding to the *Phytophthora* root rot, crown rot, foliar-blight and fruit rot disease syndromes, involving different loci governing the resistance phenotype, against each disease syndromes for various physiological race of *P. capsici* using commercial cultivars and recombinant inbred lines as a differential host (Barchenger et al. 2018). The isolate virulence of *P. capsici* is another factor, which may affect the genetic study of resistance. Isolates with low virulence would lead to the presence of false-resistance phenotype, while highly virulent isolates could overcome the genetic resistance of resistant line (Wang et al. 2016). Several resistance sources including *C. annuum* cultivars CM334, PI201232, PI201234, AC2258, perennial and YCM334 have been reported globally. Varied levels of resistance response for different *P. capsici* isolates have been documented in these resistance sources (Thabuis et al. 2003). Among these sources, cv. CM334 is the most stable having the highest resistance levels, and has been extensively utilized in breeding programs (Candole and Conner 2010). Nevertheless, breeding for PcRR resistance is intricate due to the genetic flexibility of *P. capsici* with the presence of numerous physiological races, and the existence of complex polygenic inheritance in genetic resistance sources. Furthermore, the resistance levels in commercial pepper cultivars are not comparable to that of original resistant sources (Foster and Hausbeck 2010; Thabuis et al. 2003). Therefore, genetic improvement of pepper for resistance to PcRR still is an important objective in most pepper breeding programs.

In *Solanaceae,* resistance to the *Phytophthora* has been reported to be *race-specific and quantitative in nature. Several race-specific resistance (R) genes against P. infestans including R1*, *R2*, *R3a*, *R4*, and *Rpi-blb2* on the chromosome 4, 5, 8 and 11 in potato and *Ph-1*, *Ph-2*, *Ph-3*, *Ph-5-1* and *Ph-5-2* on chromosome 1, 7, 9 and 10 in tomato have been reported (Vleeshouwers et al. 2011; Brouwer et al. 2004; Chen et al. 2014). In previous studies, resistant

QTLs to *P. infestans* on potato chromosomes V, VIII, and IX and tomato chromosomes 1, 2, 3, 5, 9 and 11 have been reported (Vleeshouwers et al. 2011; Brouwer et al. 2004; Chen et al. 2014). In pepper, two dominant race specific resistance genes *CaPhyto* from *C. annuum* cv. PI201234 and *PhR10* from cv. CM334 were mapped on chromosome P5 and P10, respectively (Wang et al. 2016; Xu et al. 2016). Race non-specific quantitative resistance is controlled by multiple genes at different loci, which is considered durable (Vleeshouwers et al. 2011; Mallard et al. 2013). In pepper, major and minor resistant QTLs against PcRR has been mapped on chromosomes P1, P3, P4, P6, P8, P9, P10, P11 and P12 utilizing various resistance sources available in *Capsicum* species (Thabuis et al. 2003; Bonnet et al. 2007; Truong et al. 2012). Two or three major effect QTLs conferring resistance to PcRR was consistently identified in close vicinity on chromosome P5 irrespective of the resistant sources and *P. capsici* isolates in different studies, however the exact location of these QTLs were reported to be inconsistent (Bonnet et al. 2007; Kim et al. 2008; Liu et al. 2014; Mallard et al. 2013; Rehrig et al. 2014). Based on the meta-analysis of QTLs, the major effect QTL, MetaPc5.1 was found to be located between 22.4–24.6 Mb-region, and MetaPc5.2 and MetaPc5.3 were located in 53.0–162.6 Mb and 9.7–13.3 Mb on chromosome P5, respectively (Mallard et al. 2013; Wang et al. 2016). Depending on different genetic backgrounds and *P. capsici* isolates several minor and isolate-specific QTLs associated with PcRR have also been detected on chromosomes P1, P3, P6, P11, and P12 (Rehrig et al. 2014; Thabuis et al. 2004; Truong et al. 2012). Several molecular markers linked to PcRR resistance, such as Phyto5NBS1 and ZL6726 with high accuracy have been developed and implemented for MAS in pepper but their use is limited against high virulent and race non-specific *P. capsici* isolates (Liu et al. 2014; Wang et al. 2016).

Ultrahigh-density linkage maps play a crucial role in the genetics and genomics studies. Utilization of ultrahigh-density linkage maps has expedited the discovery of functional genes

controlling number of traits (Han et al. 2018; Sonah et al. 2015; Yu et al. 2017), genome assembly (Guo et al. 2013; Consortium, 2012), and comparative genome analysis of important crops (Yang et al. 2017). The emergence of next-generation sequencing (NGS) technologies and high throughput genotyping tools have made possible the rapidly discovery of bulk of markers. The NGS-based genotyping approaches, such as RAD (restriction site associated DNA) genotyping (Chen et al. 2014), SLAF-seq (specific-locus amplified fragment sequencing) (Xu et al. 2016), and GBS (genotyping-by-sequencing) (Elshire et al. 2011) enable discovery and genotyping of thousands of genetic markers across the genome at relatively lower prices. Among the numerous NGS technologies, GBS is a simple, efficient and rapid approach having the potential to detect polymorphism at relatively low cost. GBS approaches have been widely used in several crops for bi-parental QTL mapping and genome-wide association studies (GWAS) to uncover loci controlling various traits (Han et al. 2018; Sonah et al. 2015; Yu et al. 2017).

Previous genetic studies in pepper for PcRR have mainly focused on bi-parental QTL mapping which mainly depends on the genetic diversity of two parental lines. Therefore, the detected QTLs were encompassed large genomic regions with low map resolution. Thus, hampering the development of tightly linked markers as well as identification of candidate genes. These limitations can be overcome by GWAS approach that allow narrowing down the QTL regions to the candidate gene level using natural populations. However, GWAS also has limitation of high false positive QTLs detection requiring additional validation of the results (Korte and Farlow, 2013; Zhu et al. 2008). Shortcomings of bi-parental QTL mapping and GWAS can be overcome by a combination of these two approaches (Han et al. 2018). Combination of these two approaches were effectively applied for genetic dissection of several important traits, including frost resistance, flowering time, panicle architecture, leaf

architecture, pungency and seed-related traits (Brachi et al. 2010; Crowell et al. 2016; Han et al. 2018; Sallam et al. 2016; Sonah et al. 2015; Tian et al. 2011) .

In our present study, I developed a large population of recombinant inbred lines (RILs) $F_{7:8}$ derived from a cross between resistant source *C. annuum* cv. CM334 and susceptible cv. ECWR30 for genetic dissection of PcRR resistance against multiple *P. capsici* isolates. Furthermore, the previously reported core collections (Lee et al. 2016) were used to evaluate PcRR resistance for GWAS. The high-density SNP markers developed through GBS were used for bi-parental QTL mapping and GWAS to identify and validate novel QTLs associated with the PcRR resistance. Based on combined approach GWAS-QTL analyses candidate resistant genes including NBS-LRR, RLK, and MLO-like protein genes in the QTL regions were identified.

MATERIALS AND METHODS

Plant materials and DNA extraction

The ECRIL mapping population comprised of one hundred and eighty-eight $F_{7:8}$ ECRILs from a cross between resistant source *C. annuum* cv. CM334, and susceptible line cv. ECW30R were developed. Single seed descent method used for RILs development and the generation advancement was carried out with shuttle breeding concept in South Korea in spring and Thailand during autumn to reduce time. The plants were grown in green house conditions using plastic pots. GWAS was carried out on a core collection population consisting of 352 *Capsicum* accessions (Lee et al. 2016). Both parental lines of ECRIL population were included in the association mapping-panel. Genomic DNA from RILs and the association panel was extracted from young leaf tissues at seedling stage using hexadecyl trimethyl ammonium bromide (CTAB) method as described by Yang et al. (2012).

Disease assay and resistance scoring

The evaluation of resistance against PcRR in RILs were tested under two different environments; Seoul National University (SNU) farm at Suwon, Republic of Korea (2017) (environment one) and Korea Research Institute of Chemical Technology (KRICT), Daejeon, Republic of Korea (2018) (environment two). On the other hand, association panel consisting of 352 *Capsicum* accessions were also evaluated in two different environments. Resistance screening was carried out in 2017 against KPC-7 at Rural development authority (RDA), Jeonju, Republic of Korea (environment a) while against JHAI1-7 and MY-1 isolates in 2018 at SNU. *P. capsici* isolates MY-1, JHAI1-7, and KPC-7 (Table I-1) used in the current study were kindly provided by Dr. Choi (KRICT). MY-1, JHAI1-7 and KPC-7 isolates were reported

to have low, medium, and high virulence, respectively (Liu et al. 2014; Jo et al. 2014). The virulence of *P. capsici* isolates was reconfirmed by conducting preliminary experiment and isolates were re-separated from the diseased plant tissues before they were used for inoculation. For inoculum preparation, *P. capsici* isolates were cultured on V8-agar medium and incubated at 27°C for mycelia growth for five days. To get abundant sporangia, V8 agar plugs in 8 mm square size was interred into summer squash *Cucurbita pepo* fruits, and incubated at room temperature for 5 days ahead of inoculation (Kim et al. 2012; Siddique et al. 2017). Sporangial mass developed on the surface of the squash fruit was scraped into water with scalpel and straining with two layers of gauze to eliminate the debris. To inoculate core collection accessions, the isolates were cultured on the PDA (Potato dextrose agar) medium for mycelium growth at 28°C under fluorescent light illumination until the whole plates covered with mycelia. The PDA plates with mycelium were flooded with distilled water to harvest mycelium. The spore concentration was then quantified using Haemocytometer and adjusted spore suspension density to approximately 5×10^4 sporangia per mL prior to inoculation. Before inoculation, the sporangial suspension was cooled at 4°C for 1 to 2 h and then moved to 25°C for 1 h to encourage zoospore release. One hundred and eighty-eight $F_{7:8}$ ECRIL lines and GWAS population consisting of 352 core collection accessions with 10 replications for each line were scored for disease resistance. The test plants were grown into 50 cell plastic trays and inoculation was carried out at the 4 - 6 true leaf stage with 5 mL spore suspension drenching at the base of the stem by a dispenser. The inoculated plants were kept under protected conditions at 24-30°C and frequently watered to facilitate the disease establishment. *C. annuum* cv. CM334, PI201234, and Mohanjilju (commercial F_1 hybrid, Syngenta Korea) were used as resistant controls, while *C. annuum* cv. ECW30R, Tean, and Geumsugangsan (commercial F1 hybrid, Takii Korea) were used as susceptible controls to compare the resistance levels. The disease severity was scored finally at three weeks of post inoculation based on the disease scale

1-4 where 1 = no visible symptoms observed; 2 = dark lesion visible on the base of stem, but surviving without wilting; 3 = wilting with dark lesion at the base of stem; 4 = wilting and death of whole plant as described by Kim et al. (2012) (Fig. I-1).

Table I- 1. Details about the three *P. capsici* isolates used in this study to screen the ECRILs and GWAS core collection.

Isolate	Collection Site	Host	Virulence	Mating Type	Source	References
KPC-7	Kangwon/Korea	*C. annuum*	High	A1	KRICT	Jo et al. (2014)
JHAI1-7	Chungbuk/Korea	*C. annuum*	Moderate	A1	KRICT	Liu et al. (2014)
MY-1	Kangwon/Korea	*C. annuum*	Low	A2	KRICT	Jo et al. (2014)

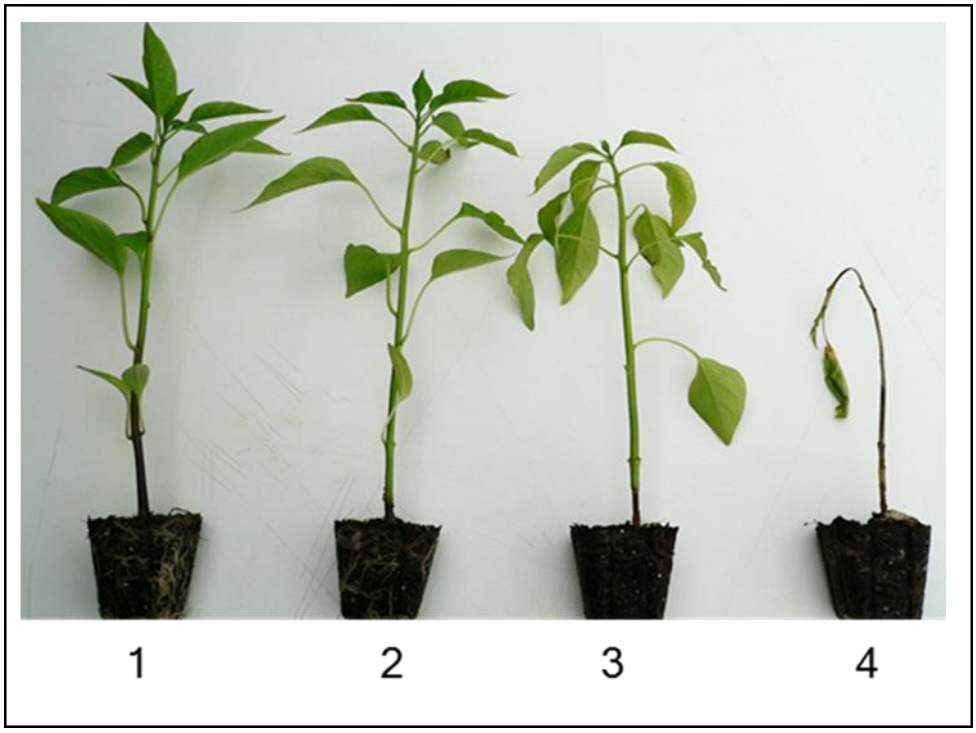


Figure I- 1. Disease scale (1 to 4) used to distinguish the resistant and susceptible genotypes, where 1 = no visible symptoms, 2 = dark lesion visible on the base of the stem but surviving without wilting, 3 = wilting, with a dark lesion at the base of the stem, and 4 = wilting and death of the whole plant**.**

Genotyping-by-sequencing

Genotyping-by-sequencing was performed as described in Truong et al. (2012). Briefly, genomic DNA diluted and adjusted at 50ng/µl from ECRILs population was digested with *EcoR*I and *Mse*I resection enzymes using a Bioanalyzer DNA 1000 Chip (Agilent Technologies, Santa Clara, CA, USA. For GWAS population, GBS libraries were constructed manually using *Pst*I/*Mse*I and *EcoR*I/*Mse*I as described by Jo et al. (2017) and Han et al. (2018). Followed by the ligation of selective adaptors, libraries were then amplified with selective primers, which contain 'GA' for RIL and 'TA' for GWAS population. Amplified libraries consisting of 188 ECRILs and two replicates of susceptible parents were pooled into a single tube. On the other hand, libraries of core collection populations were pooled into five tubes. Pooled libraries were sequenced using an Illumina HiSeq2000 at Macrogen (Macrogen, Inc. Seoul, Korea). Trimming and quality control of GBS raw data were performed using the CLC Genomics Workbench v6.5 (QIAGEN, Aarhus, Denmark) with a minimum length of reads 80 bp and minimum quality Q20. Filtered reads were aligned to *C. annuum* cv. CM334 reference genome (chromosome v1.6, http://peppergenome.snu.ac.kr) by Burrows-Wheeler Aligner (BWA) (Li and Durbin 2010). Filtring and SNP calling was performed using the Genome Analysis Toolkit (GATK) UnifiedGenotyper ver. 3.3-0. In ECRILs SNPs were filtered with QUAL value greater than 20 and minimum read depth of 3. In association panel, SNPs were filtered with minor allele frequency (MAF) > 0.03, calling rate > 0.6, and inbreeding coefficient (F) > 0.8.

Construction of bin map and QTL analysis

The SNPs with unequal segregation and missing more than 10% data were excluded for the linkage map construction. For the construction of a linkage map, bins were treated as genetic markers. Sliding window approach was used to impute the missing data and to find recombination break points as described previously (Han et al. 2016). To assign genetic

positions to the bins, arranged bins were mapped with a LOD (logarithm of the odds) threshold of 3.0 and a distance threshold of 30 cM using CarthaGene software. The Kosambi mapping function was used to convert genetic distances between markers. QTL analysis was performed using composite interval mapping using Windows QTL Cartographer 2.5 (Wang et al. 2011) The 1000-permutation tests ($P < 0.05$) was performed to determine the LOD threshold for significance of each QTL. Explanations of the phenotypic variance (PV) and additive effects for each QTL were also obtained using this software.

One-way analysis of variance ANOVA was performed to test the epistatic interaction of major and minor QTLs. The physical locations of the QTLs were compared by physical and genetic location of bins linked to the QTLs. *C. annuum* reference genome (chromosome v1.6, http://peppergenome.snu.ac.kr) was used to compare the positions of significant SNPs under QTLs and GWAS of present study and with the previously reported linked markers. The markers sequence of previous studies were retrieved using https://solgenomics.net/search/markers, and BLAST into the reference genomes. If the sequence information of linked markers were not publically accessible, the closest marker was used.

Genome-wide association analyses for PcRR

Filtered 168,714 SNPs discovered from 352 GWAS accessions were used in association mapping. Population structure (PCA and Kinship matrixes), and genome-wide association (GWAS) was estimated with the compressed mixed linear model by R package of Genomic Association and Prediction Integrated Tool (GAPIT) (Lipka et al. 2012) with default parameters. All of the probabilities generated in the association runs were transformed by -$\log_{10}$ P (0.05). Scores for individual chromosome were then inspected in Manhattan plots to determine whether the SNPs reached the significance threshold. The -$\log_{10}$ P values of SNPs

from GWAS were adjusted by Bonferroni (1936) multiple test correction and the adjusted cut-off for accepting thresholds were set to -$\log_{10}$ (P) value ≥ 7.0.

Identification of candidate genes

To identify putative candidate genes associated with PcRR resistance, position of highly significant and linked bin markers within QTLs on the genetic map was compared with their physical position on the pepper genome (v1.6) and 1 Mb upstream and downstream sequence were mined for genes associated with disease resistance. Similarly, the 1 Mb sequence information of the highly significant SNPs detected from the GWAS peaks was used for BLAST (blastn) analysis at *C. annuum cv.* CM334 reference genome database (chromosome v1.6, http://peppergenome.snu.ac.kr) to find associations with other candidate genes.

Statistical analysis for estimating genetic variability against PcRR

Mean disease score of ten plants per ECRIL and GWAS accession were used for calculating frequency distribution, coefficient of variation (CV%), range and genetic variability parameters in three different isolates. Broad-sense heritability (h^2_{bs}) was calculated with the formula suggested by Allard (1960) $h^2 = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2_p} \times 100$ where, $\hat{\sigma}^2_g$ is the genetic variance and $\hat{\sigma}^2_p$ is the phenotypic variance. Further, expected genetic advance (GA) was measured with proposed formula by Lush (1949) $GA = \frac{V_g}{V_p} \times \sqrt{V_p} \times k$, $= \frac{V_g}{\sqrt{V_p}} \times k$, Where, GA is the expected genetic advance, V_g is the genotypic variance, V_p is the phenotypic variance and *K* is the selection differential (constant) i.e. 2.06 at 5% selection intensity (Allard, 1960).

RESULTS

Evaluation of PcRR resistance in ECRILs and the GWAS core collection.

I evaluated 188 $F_{7:8}$ ECRILs, parental lines, and additional resistant and susceptible controls in two environments (E1 and E2) for resistance to three *P. capsici* isolates (Fig. I-2). The three isolates showed a significant difference in terms of symptom development and virulence. For the highly virulent isolate KPC-7, disease symptoms, including wilting and the appearance of water-soaked lesions at stem collars, were observed within 72 h of inoculation in the susceptible controls (ECW30R, Tean, and Geumsugangsan), and the plants completely wilted and died at 7 to 10 days post inoculation (DPI). By contrast, in plants infected with the moderately virulent isolate JHAI1-7, I noticed disease symptoms on the susceptible controls at 5 to 7 DPI, and the plants completely wilted and died at two weeks post inoculation. For the isolate MY-1, with low virulence, symptoms were observed on the susceptible controls at 10 to 14 DPI, but complete plant death was not detected up to three weeks post inoculation. The resistant controls CM334 and Mohanjilju (commercial F_1 hybrid, Syngenta Korea) remained symptomless after inoculation with all three isolates, even at three weeks post inoculation (Fig. I-2a). However, PI201234 was partially resistant to isolates KPC-7 and JHAI1-7 and completely resistant to isolate MY-1 (Fig. I-2a).

I evaluated PcRR resistance in the ECRIL and GWAS core collection and scored disease severity on a scale of 1–4 at three weeks post inoculation in their respective environments, as described by[33] (Fig. I-1). Ten plants from each ECRIL were grown and the resistance level of each plant scored individually. I calculated the frequency distribution based on the mean resistance values of each RIL (Fig. I-2b and 2c). Overall, the frequency distribution of disease severity showed a normal distribution for KPC-7 and JHAI-7 in both environments, although

more ECRILs showed susceptibility to KPC-7 in E1 than E2 (Fig. I-2b and 2c). By contrast, the distribution of disease severity of the RILs skewed towards resistance to the low-virulence isolate MY-1, and no ECRILs reached the disease rating of 4 during the experiment in either environment, except for two ECRILs in environment E2 (Fig. I-2b and 2c and Fig. I-3). Correlation analysis indicated a significant, positive correlation between the two environments for PcRR resistance; however, the correlations between different isolates varied (Fig. I-3).

I evaluated the GWAS accessions for PcRR resistance against the same three isolates as the ECRILs in two different environments (EA and E1). Among GWAS accessions, CM334 showed complete resistance, whereas PI201234, YCM334, and nine other accessions showed partial resistance, with slight disease symptoms at the stem collars upon infection with isolates KPC-7 and JHAI-7 (Fig. I-2a and 2d and Table I-2). Most of these PcRR-resistant accessions include known resistance sources (Table I-2). In the case of the GWAS core collection, the isolate JHAI1-7 showed slightly higher virulence compared to KPC-7 in a few lines, but a higher mortality rate was observed for KPC-7 (Fig. I-2d and Fig. I-5). The disease response to the low-virulence isolate MY-1 was less severe, with 52 accessions showing no PcRR symptoms (Fig. I-2d). Correlation analysis revealed significant, positive correlations between the KPC-7 and JHAI1-7 isolates and both environments for PcRR resistance (Fig. I-4). Highly virulent isolate KPC-7 for ECRILs in E2 and moderately virulent isolate JHAI1-7 for the GWAS accessions showed a minimum coefficient of variation (CV) of 38.55 and 28.09%, respectively. Weak isolate MY-1 showed the highest (CV 57.0 to 62.02%) disease score in both screened populations in all environments (Table I-3). Overall, broad-sense heritability (H^2%) was high for all PcRR isolates. In E1, isolate KPC-7 showed the highest H^2 (82.48%) among ECRILs, whereas isolate JAHI-7 showed an H^2 of 87.40% and 84.16% for ECRILs in E2 and the GWAS accessions, respectively (Table I-3).

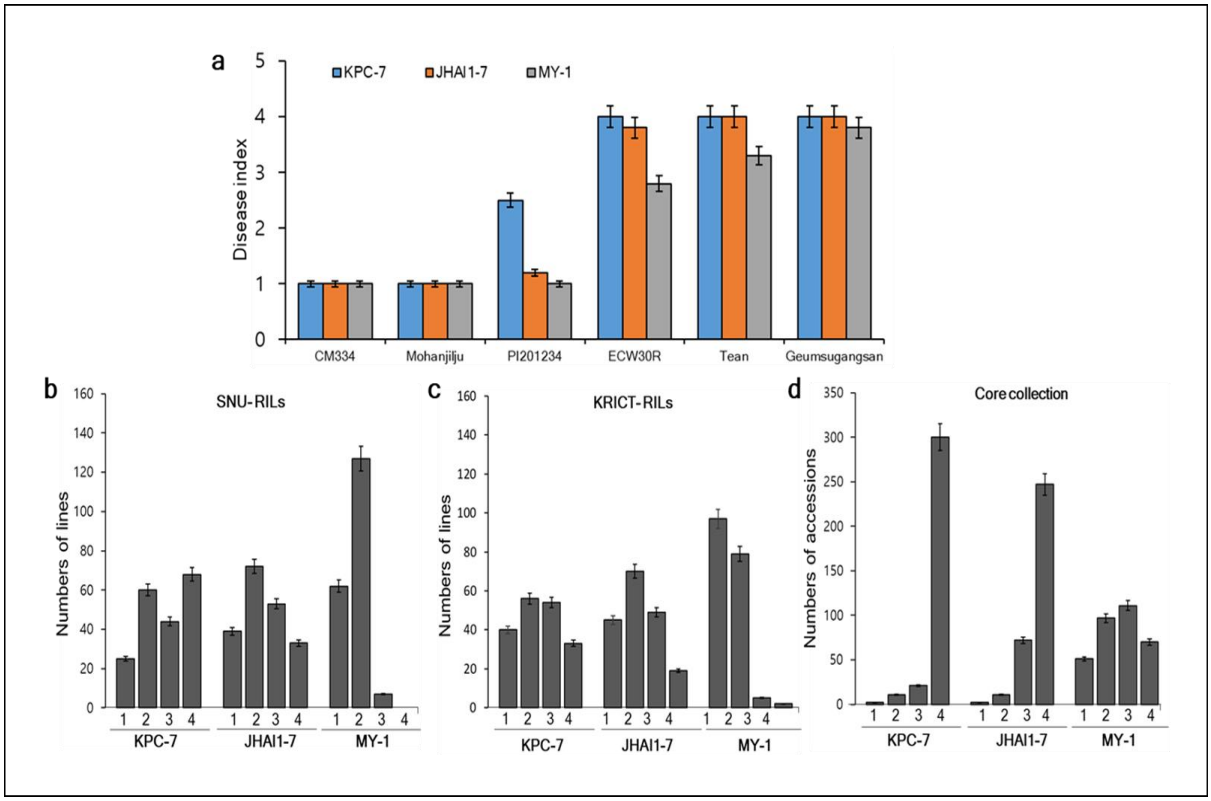


Figure I- 2. Frequency distribution of PcRR in the parental lines, controls, ECRILs (SNU-RILs/KRICT-RILs), and GWAS (core collection). (a) Disease index of resistant controls (cv. CM334, Mohanjilju, and PI201234) and susceptible parents (cv. ECW30R, Tean, and Geumsugangsan) at three weeks post inoculation. (b) Frequency distribution of ECRILs against isolates KPC-7, JHAI1-7, and MY-1 in environment E1 (SNU-RILs). (c) Frequency distribution of ECRIL against isolates KPC-7, JHAI1-7, and MY-1 in environment E2 (KRICT-RILs). (d) Frequency distribution of GWAS (core collection) against isolates KPC-7, JHAI1-7, and MY-1. Bars denote the standard deviation.

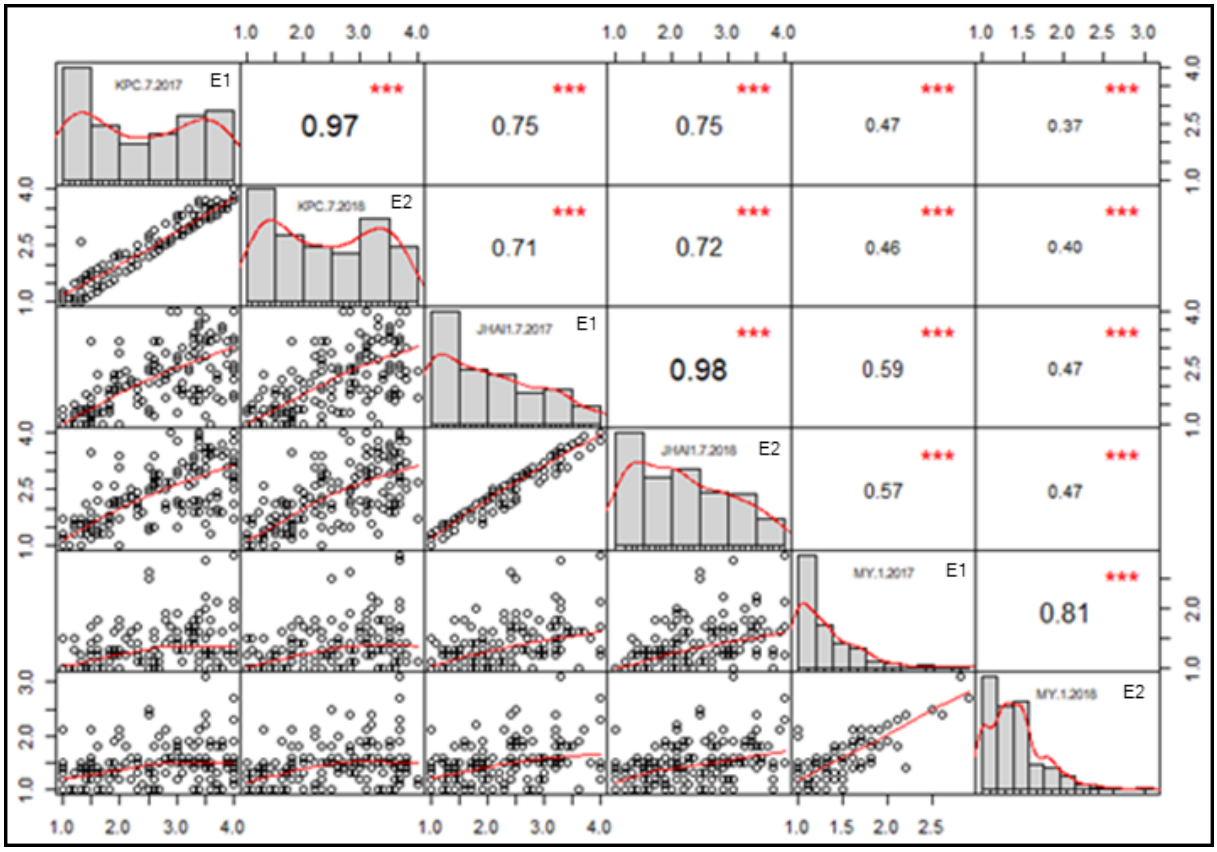


Figure I- 3. Pearson correlation matrix of PcRR disease index for ECRILs in two environments (E1 and E2).

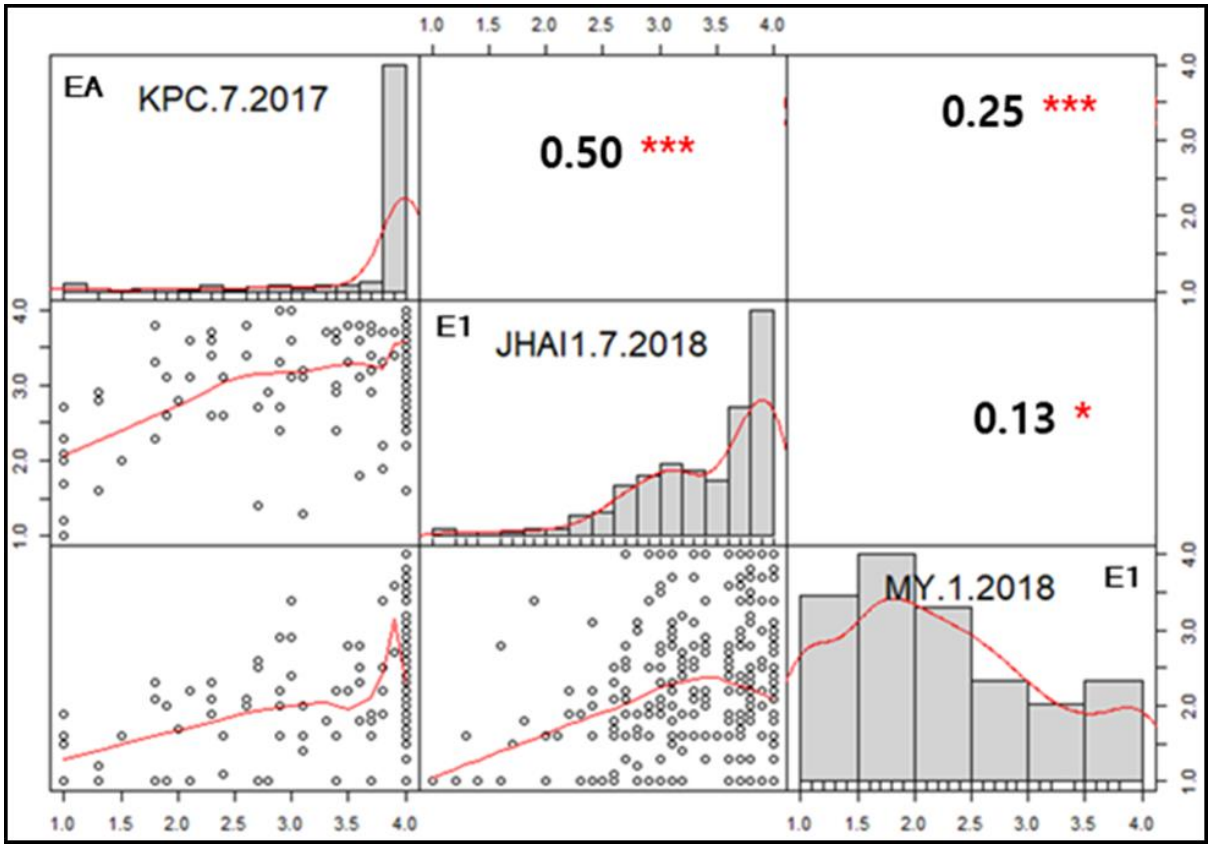


Figure I- 4. Pearson correlation matrix of PcRR disease index for the GWAS (core collection) in two environments (EA and E1).

Table I- 2. Details about the resistant accessions from the GWAS core collection.

Accession name	IT number[X]	Species	Origin	Resistance score[Y]		
				MY-1	JHAI1-7	KPC-7
Szechwan 8	IT 158280	*C. annuum*	Taiwan	1	1.8	1
Pangalengan-1	IT 158714	*C. annuum*	Indonesia	1	1.7	1
Ikdosan	IT 236666	*C. annuum*	China	1	1.9	1
Jungangjongmyo-2000-5032	IT 236424	*C. annuum*	-	1	2	1
Ibam	IT 231151	*C. annuum*	-	1	1.2	1
CM334	IT 250209	*C. annuum*	Mexico	1	1	1
CM331	IT 236422	*C. annuum*	Mexico	1	1	1
PI 439449	IT 264121	*C. chinense*	Peru	1	2	1
AC2212	IT 294726	*C. chinense*	-	1	1.2	1
YCM334	K150013	*C. annuum*	-	1	1.5	1.6
PI 201234	IT 294726	*C. annuum*	-	1	1.7	1.3

[X]GenBank accession number.
[Y]Disease index.

Table I- 3. Genetic variation among ECRILs and GWAS accessions inoculated with three isolates of PcRR in different environments/years

Population (Environment-year)	Isolate	Mean ± SE	Range (Max-Min)	CV (%)	h^2 (Broad-sense) (%)	Expected genetic Advancement (5%)
SNU-ECRILs	KPC-7	2.12 ± 0.28	4-1	42.05	82.48	170.75
(E1-2017)	JHAI-7	1.75 ± 0.25	4-1	46.10	80.22	171.30
	MY-1	1.16 ± 0.21	2.09-1	57.00	52.45	89.32
KRICT-ECRILs	KPC-7	2.38 ± 0.91	4-1	38.55	80.21	153.17
(E2-2018)	JHAI-7	2,19 ± 0.84	4-1	38.60	87.40	124.16
	MY-1	1.43 ± 0.36	4-1	56.29	75.82	178.83
GWAS	KPC-7	2.86 ± 0.45	4-1	50.09	59.61	96.79
Population	JHAI-7	3.03 ± 0.26	4-1	28.09	84.16	122.39
(EA-2017, E1-2018)	MY-1	1.89 ± 0.37	4-1	62.06	72.85	178.73

SNP discovery through GBS and bin map construction for ECRILs.

I performed genotyping of the ECRILs using GBS after *Eco*RI and *Mse*I digestion. Sequencing of the prepared GBS libraries of 188 ECRILs and parental lines resulted in 525.62 million raw reads. After trimming the raw reads, I obtained an average of approximately 4.4 million reads per sample (Table I-4). After aligning the reads to the CM334 reference genome V1.6, 66,405 SNPs were detected (Fig. I-5a). The SNPs were distributed more densely at the ends of the chromosomes than in the middle regions (Fig. I-5a). After removing more than 90% of missing data and filtering unequally distributed SNPs, I obtained 13,021 high-quality SNPs that were equally distributed across the genome, which were used for bin linkage map construction (Table I-4 and Table I-5 and Fig. I-6). To impute missing data and genotyping errors, I employed a sliding window approach (Han et al. 2016). Recombination breakpoints were determined by sliding 25 SNPs consecutively as one window. As a result, a high-density bin map consisting of 2,663 bins covering a total genetic distance of 1,428.8 cM was constructed (Table I-5 and Fig. I-6). The average genetic distance between bins was estimated to be 0.6 cM (Table I-5). Among the 12 linkage groups, maximum and minimum genetic distances of 195.6 and 44.7 cM were obtained for chromosomes P1 and P8, respectively (Table I-5).

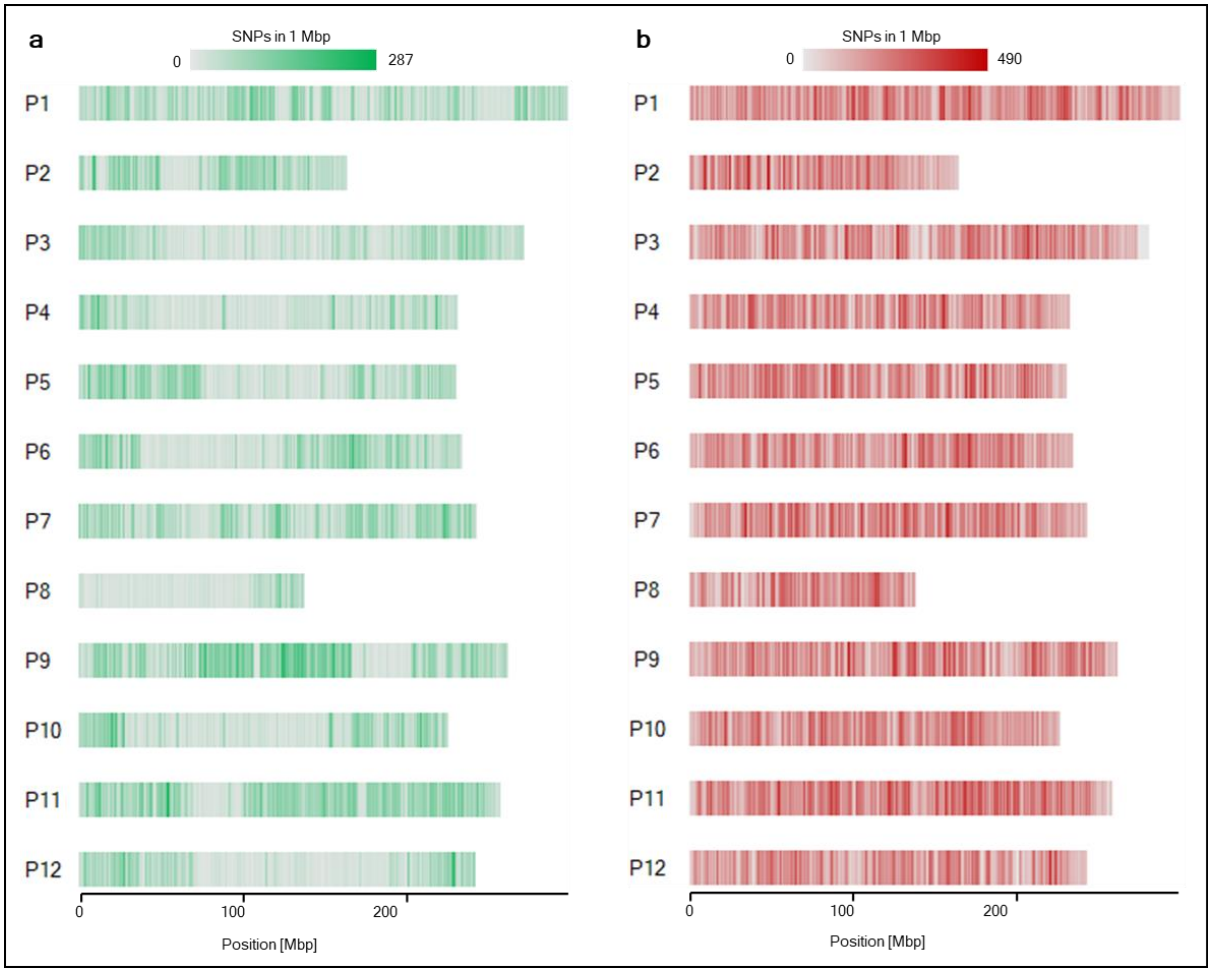


Figure I- 5. SNP density (number of SNPs per 1 Mbp). (a) SNP density for ECRILs digested with Eco*R*I/Mse*I*. (b) SNP density for the GWAS (core collection) digested with Pst*I*/Mse*I* and Eco*R*I/Mse*I*

Table I- 4. Genotyping data for GWAS and QTL mapping

	ECRIL	**GWAS Core Collection**
Number of accessions (lines)	188	352
Genotyping method	GBS (*Eco*RI/*Mse*I)	GBS (*Pst*I/*Mse*I and *Eco*RI/*Mse*I)
Average number of reads per sample	4,450,271*	3,326,422
Total number of retained SNPs	13,021	507,713
Average distance between SNPs (bp)	366,349	16,305

*Number of spots sequenced by paired-end reads

Table I- 5. Comparison of the physical length and genetic distance of bins for biparental QTL mapping

Chr.	Number of SNPs	Number of Bins	Physical Length of Bin (Mb)		Genetic Distance of Bin (cM)	
			Mean	Total	Mean	Total
P1	1551	329	1.2	272.7	0.6	195.6
P2	925	190	1.0	171.1	0.5	94.1
P3	887	275	1.1	257.9	0.6	162.9
P4	613	191	1.5	222.6	0.6	125.6
P5	1009	183	1.2	233.5	0.5	91.6
P6	1034	219	1.1	236.9	0.5	114.4
P7	1323	259	1.1	231.9	0.5	129.4
P8	212	52	2.1	145.1	0.9	44.7
P9	1864	260	1.2	252.8	0.5	119
P10	808	207	1.3	233.6	0.5	115.3
P11	2127	323	0.9	259.7	0.4	116.1
P12	668	175	1.2	235.7	0.7	120.1
Total	13,021	2,663	1.2	2,753.5	0.6	1,428.8

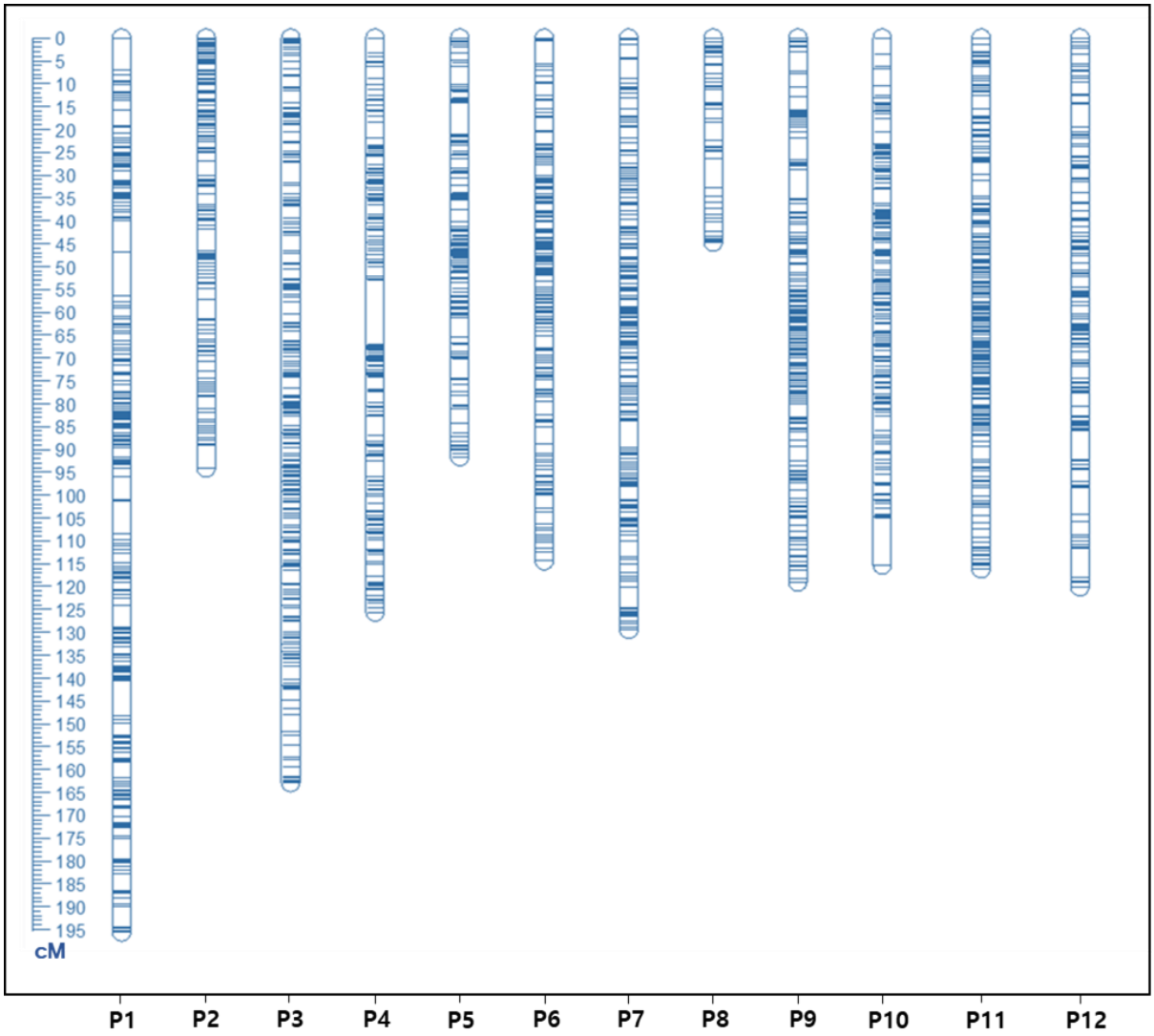


Figure I- 6. High-density genetic linkage map of ECRILs comprising 2,663 bins covering a genetic distance of 1,428 cM based on genotyping-by-sequencing.

PcRR resistance QTLs in the ECRILs.

I identified QTLs for PcRR resistance to three isolates, KPC-7, JHAI1-7, and MY-1, using 188 ECRILs and the high-density SNP linkage map in two environments, E1 and E2 (Fig. I-7 and Table I-6). In total, 14 QTLs, including those with major and minor effects, those commonly detected, those specific to isolates, and those specific to environments were detected across the genome, explaining phenotypic variation (R^2) ranging from 3.73 to 38.99% (Table I-6). The PcRR resistance QTLs were detected on chromosomes P1, P2, P4, P5, P7, P8, and P11.

Two major QTLs to highly virulent isolate KPC-7 were detected on chromosome P5, *E1Kpc-5.2* and *E2Kpc-5.3*, at 28.91 and 34.61 cM, corresponding to 27.3–29.2 Mb and 34.6–37 Mb, respectively, in the *C. annuum* 'CM334' reference genome. *E1Kpc-5.2* explained 33.77% (LOD of 21.32) and 37.12% (LOD of 23.25) and *E2Kpc-5.3* explained 25.40% (LOD of 14.87) and 25.14% (LOD of 14.17) of phenotypic variation (R^2) in E1 and E2, respectively (Fig. I-7 and Table I-6). Three minor QTLs were detected in both environments, but only one QTL (on chromosome P1 at 77.61 cM [9.5–9.6 Mb]) was commonly detected in both environments (Fig. I-7 and Table I-6).

Two major QTLs to moderately virulent isolate JHAI1-7 on chromosome P5, *E1Jha-5.1* and *E2Jha-5.2*, were detected at 22.61 and 28.91 cM, corresponding to 18.7–19.5 and 27.3–29.2 Mb, respectively, in the *C. annuum* 'CM334' reference genome. *E1Jha-5.1* explained 27.31% (LOD of 18.30) and 26.62% (LOD of 18.13) and *E2Jha-5.2* explained 38.99% (LOD of 29.73) and 38.21% (LOD of 29.66) of phenotypic variation (R^2) in E1 and E2, respectively (Fig. I-7 and Table I-6). Four minor QTLs on P8 and P11 were detected in both environments (Fig. I-7 and Table I-6).

For low-virulence isolate MY-1, two major-effect QTLs on chromosome P5, *E1My-5.1*

and *E2My-5.2*, were detected at 22.61 and 28.61 cM, corresponding to 18.7–19.5 Mb and 27.3–29.2 Mb, respectively, in the *C. annuum* 'CM334' reference genome. *E1My-5.1* explained 18.22% (LOD of 8.94) and 10.02% (LOD of 4.94) and *E2My-5.2* explained 19.68% (LOD of 10.80) and 12.61% (LOD of 6.33) of phenotypic variation (R^2) in E1 and E2, respectively (Fig. I-7 and Table I-6). One major QTL on chromosome P5, *E1My-5.3*, was detected in E2 only at 34.61 cM, corresponding to 34.6–37 Mb, which explained 8.90% (LOD 4.38) of phenotypic variation (R^2). Two minor QTLs were detected in both environments, but only one minor QTL was commonly detected in both environments (on chromosome P11 at 66.41 cM corresponding to 189–190 Mb) (Fig. I-7 and Table I-6). Notably, all QTLs associated with the MY-1 isolate showed lower R^2 and LOD values when compared to environment E1 and the QTLs detected against the two other isolates used in this study (Fig. I-7 and Table I-6).

Taken together, three major QTLs on P5 were commonly detected for all three isolates in two environments, whereas minor QTLs on different chromosomes were isolate or environment specific (Fig. I-7 and Table I-6). Among the major QTLs, *QTL5.2*, which is located at 27.3–29.2 Mb, was commonly detected for all three isolates and was named according to its genomic position on chromosome P5. The major QTL at 18.7–19.5 Mb (common to JHAI1-7 and MY-1) was named *QTL5.1*. The major QTL at 34.6–37 Mb (common to KPC-7 and MY-1) was named *QTL5.3*. Even though minor QTLs were commonly detected on chromosome 11 for all three isolates, they were located at different positions. I detected isolate-specific minor QTLs on other chromosomes, such as the minor QTLs *Kpc1* (for KPC-7) on P1, *Jha8.1*, *Jha8.2*, *Jha11.1*, and *Jha11.2* (for JHAI1-7) on P8 and P11, and *My11* (for MY-1) on P11.

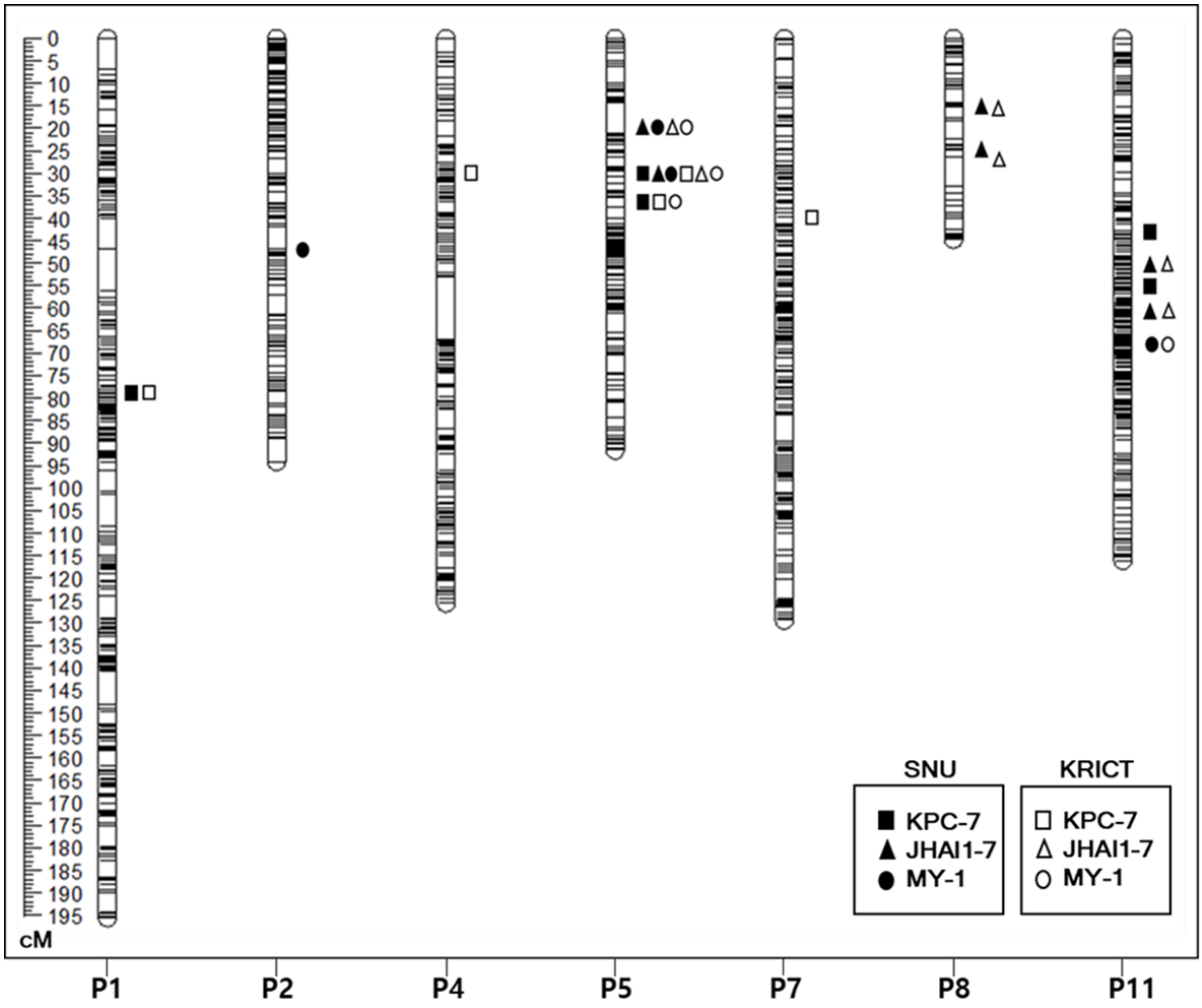


Figure I- 7. Bin linkage chromosomal map showing the locations of PcRR resistance QTLs with the genetic distance shown in Centimorgans (cM) for the ECRIL population evaluated with three *P. capsici* isolates in two respective environments.

Table I- 6. PcRR resistance QTLs detected by composite interval mapping in the ECRIL population inoculated with three isolates evaluated in their respective environments.

Isolate	Evaluating Environment	QTLs	Chr.	Position[a] (cM)	Location[b] (Mb)	LOD [c]	R^2 (%)[d]	Additive Effect
KPC-7	E1-SNU*2017	*E1Kpc-1*	P1	77.61	9.5-9.63	3.97	4.57	2.27
		E1Kpc-5.2	P5	28.91	27.3-29.2	21.32	33.77	6.26
		E1Kpc-5.3	P5	34.61	34.6-37	14.87	25.40	5.41
		E1Kpc-11.1	P11	41.40	49.5-50.5	4.78	5.59	2.53
		E1Kpc-11.2	P11	52.01	74-76	4.20	4.94	2.40
	E2-KRICT*2018	*E2Kpc-1*	P1	77.61	9.5-9.63	6.69	7.81	4.42
		E2Kpc-4	P4	28.41	87-88.5	4.64	5.41	2.96
		E2Kpc-5.2	P5	29.21	27.3-29.2	23.25	37.12	5.78
		E2Kpc-5.3	P5	34.61	34.6-37	14.17	25.14	4.70
		E2Kpc-7	P7	38.71	24.5-26.4	3.60	3.73	1.80
JHAI1-7	E1-SNU*2017	*E1Jha-5.1*	P5	22.61	18.7-19.5	18.30	27.31	4.80
		E1Jha-5.2	P5	28.91	27.3-29.2	29.73	38.99	5.72
		E1Jha-8.1	P8	13.31	126.4-126.6	7.25	6.85	2.40
		E1Jha-8.2	P8	23.71	128.2-130.3	5.82	5.27	2.06
		E1Jha-11.1	P11	48.41	61.2-63	9.01	8.47	2.62
		E1Jha-11.2	P11	59.31	111.4-112.8	7.71	7.37	2.47
	E2-KRICT*2018	*E2Jha-5.1*	P5	22.61	18.7-19.5	18.13	26.62	4.54
		E2Jha-5.2	P5	28.91	27.3-29.2	29.66	38.21	5.44
		E2Jha-8.1	P8	12.31	126.4-126.6	4.61	4.36	1.83
		E2Jha-8.2	P8	19.51	128.2-130.3	4.57	4.26	1.77
		E2Jha-11.1	P11	48.41	61.2-63	9.44	8.82	2.57
		E2Jha-11.2	P11	60.51	111.4-112.8	7.34	7.03	2.33
MY-1	E1-SNU*2017	*E1My-2*	P2	45.91	129-131	4.30	6.99	1.00
		E1My-5.1	P5	22.61	18.7-19.5	8.94	18.22	1.60
		E1My-5.2	P5	28.61	27.3-29.2	10.80	19.68	1.66
		E1My-11	P11	66.41	189-190	4.89	8.13	1.07
	E2-KRICT*2018	*E2My-5.1*	P5	22.31	18.7-19.5	4.97	10.02	1.19
		E2My-5.2	P5	28.61	27.3-29.2	6.36	12.61	1.35
		E2My-5.3	P5	34.61	34.6-37	4.38	8.90	1.12
		E2My-11	P11	64.41	189-190	3.33	6.17	1.50

[a]Positions of the markers on the linkage map in centimorgans (cM).
[b]Position of right and left bins in the pepper genome.
[c]Maximum log-likelihood (LOD) value of QTL.
[d]Phenotypic variation (R^2) explained by a QTL at the linked marker.
*E1-SNU, Environment 1, Seoul National University; *E2-KRICT, Environment 2, Korea Research Institute of Chemical Technology.

Epistatic interactions of PcRR resistance QTLs.

I detected additive-by-additive epistatic interactions against highly virulent isolate KPC-7 between major QTLs *QTL5.2* and *QTL5.3* and *QTL5.2* and *E1Kpc-11.1* in environment E1 and between QTLs *QTL5.2* and *QTL5.3* in environment E2 (Table I-7). For the moderately virulent isolate JHAI1-7, I detected additive-by-additive epistatic interactions between *QTL-5.2* and *Jha8.1* and *QTL5.2* and *Jha11.2* in environment E1 and between QTLs *QTL5.2* and *Jha11.1* in environment E2 (Table I-7). By contrast, for the low-virulence isolate, I detected additive-by-additive epistatic interactions only between *QTL5.2* and *My11* in both E1 and E2 (Table I-7). The total R^2 effects of individual QTLs and their interactions ranged from 16.6 to 42.0% (Table I-7). The detection of epistatic interactions between loci on chromosomes P5 and P11 points to multilocus epistatic control of PcRR resistance in pepper.

Table I- 7. Epistatic effects of major QTLs in ECRILs against three PcRR isolates evaluated in two environments.

Isolate	Environment*year	QTLs and their Epistatic Interactions	R^2 (%)	Total R^2 (%)
KPC-7	E1-SNU*2017	*QTL5.2*	24.4	42
		QTL5.3	15.1	
		QTL5.2 x *QTL5.3*	2.5	
		QTL5.2	24.4	41.1
		E1Kpc-11.1	5.5	
		QTL5.2 x *E1Kpc-11.1*	11.2	
	E2-KRICT*2018	*QTL5.2*	25.6	39.3
		QTL5.3	7.5	
		QTL5.2 x *QTL5.3*	6.2	
JHAI1-7	E1-SNU*2017	*QTL5.2*	28.2	35.1
		Jha8.1	4.5	
		QTL5.2 x *Jha8.1*	2.4	
		QTL5.2	28.2	44
		Jha11.2	13.5	
		QTL5.2 x *Jha11.2*	2.3	
	E2-KRICT*2018	*QTL5.2*	32.5	43
		Jha11.1	8.4	
		QTL5.2 x *Jha11.1*	2.1	
MY-1	E1-SNU*2017	*QTL5.2*	8.7	17.4
		My11	4.8	
		QTL5.2 x *My11*	3.9	
	E2-KRICT*2018	*QTL5.2*	11.6	16.6
		My11	2.7	
		QTL5.2 x *My11*	2.3	

Genome-wide association study of PcRR resistance.

I aligned the sequences derived from GBS to *C. annuum* reference genome v.1.6. GBS genotyping of 352 accessions using two sets of libraries derived from double digestion with restriction enzyme pairs *Pst*I/*Mse*I and *Eco*RI/*Mse*I generated a total of 6,126,403 of SNPs. The SNPs were evenly distributed using this combination of restriction enzymes. SNPs in the GWAS core collection were relatively uniformly distributed across the chromosomes (Fig. I-5b). I filtered out the SNPs with minor allele frequency > 0.05, SNP coverage > 0.6, and inbreeding coefficient > 0.8, resulting in 507,713 high-quality SNPs distributed evenly on chromosomes P1 to P12 (Fig. I-5b and Table I-4).

Of the SNPs association with resistance to PcRR against isolate KPC-7, 117 were significant SNPs, with -log10 (p) value > 7.0, explaining phenotypic variation (R^2) > 0.2 (Fig. I-8a and Table I-8). An association study revealed significant SNPs linked to PcRR resistance to KPC-7 on chromosomes P2 (1 SNP at 149 Mb), P3 (1 SNP at 5 Mb), P5 (64 SNPs at 27.7–29.7 and 206–212 Mb), P6 (2 SNP at 171 Mb), P7 (13 SNPs at 24 and 90 Mb), P9 (1 SNP at 263 Mb), P10 (14 SNPs at 8, 14, 70, 112, 148, and 196 Mb), P11 (8 SNPs at 75, 90, 92, and 112 Mb), and P12 (13 SNPs at34, 42, 57, 224, and 230 Mb) (Fig. I-8a and Table I-8). GWAS of isolate JHAI1-7 revealed an associated SNP peak on chromosome P5 with a similar position to that of the KPC-7 isolate (27.7 to 29.7 Mb), but only two SNPs showed a significance level above the adjusted threshold level at -log10(p) value > 7.0, explaining phenotypic variation (R^2) > 0.2 (Fig. I-9a and Table I-8). The two significant SNPs were located at 28.2 to 28.6 Mb on chromosome P5. By contrast, GWAS against the low-virulence isolate MY-1 did not detect significant SNPs (Fig. I-9b). Notably, only a few or no significant SNPs were associated with isolates JHAI1-7 and MY-1, perhaps due to environment and genotype interactions or to the low virulence levels of the isolates.

Importantly, 117 significant SNPs associated with resistance to isolate KPC-7 on three regions of chromosomes P5, P7, and P11 colocalized with QTLs detected in this study (Fig. I-8a and Table I-6). On chromosome P5, 49 SNPs colocalized with commonly detected major QTL region *QTL5.2* (27.3–29.2 Mb) (Fig. I-8b and Table I-6). One SNP detected on chromosome P7 at 25 Mb corresponded to the location of QTL *E2Kpc-7* (Fig. I-8c and Table I-6). Another region on chromosome P11 at 75.2 Mb corresponded to QTL *E1Kpc-11.2* (Fig. I-8g and Table I-6). The GWAS-SNPs at three regions on chromosomes P5, P10, and P11 colocalized with previously detected QTLs and linked markers (Fig. I-8b and 8f). The GWAS-SNPs on chromosome P5 in a 2.5-Mb region at position 27.0–29.5 Mb colocalized with a QTL detected in earlier reports (Mallard et al. 2013; Liu et al. 2014; Rehrig et al. 2014; Wang et al. 2016) (Fig. I-8b). SNP marker S10_14299301 located at 14,299 kb on chromosome P10 colocalized with the previously identified QTL *Ph051-10-2* (*RAUDPC*) for PcRR resistance (Truong et al. 2012) (Fig. I-8f). Regions detected by GWAS with significant SNPs on chromosome P10 at 196 Mb (Fig. I-8f) colocalized with the region containing *PhR10*, a race-specific PcRR resistance locus (Xu et al. 2016). I identified 13 significant GWAS-SNPs linked to PcRR resistance at the end of the long arm (206–212 Mb) of chromosome P5 that were not detected in previous studies, emphasizing the involvement of this genomic region in PcRR resistance (Fig. I-8a). I used the significant SNP S05_208290460, with a -log10(p) value of 9.72 identified from the GWAS-SNPs (Table I-8), to compare the resistance levels of the GWAS accessions. As shown in the box plots in (Fig. I-10) the homozygous genotype S05_208290460 AA is associated with an increased resistance level against all three isolates compared to the alternative homozygous genotype GG.

I conducted extensive linkage disequilibrium (LD) analysis of the GWAS core collection (N=352) based on all adjacent marker pairs within a chromosome or within a

haplotype block (Table I-9). I identified 29,269 haplotypes, with an average of 2,439 per chromosome. The average haplotype was 81.2 kb in size, with an average of 16.3 SNPs per haplotype (Table I-9). Further haplotype block analysis carried out with significant SNPs on chromosome P5 revealed 4 blocks containing 64 significant SNPs (Fig. I-8d and Table I-8). The three blocks on P5 colocalized with *QTL 5.2* detected in this study and with previously identified QTLs (Mallard et al. 2013; Liu et al. 2014; Rehrig et al. 2014; Wang et al. 2016) (Fig. I-8b and 8d). Two haplotype blocks each were obtained on chromosomes P7, P10 and P11, including regions with significant SNPs (Fig. I-8e and 8h and 8i). These results indicate that the identified haplotype regions are strongly associated with the PcRR resistance.

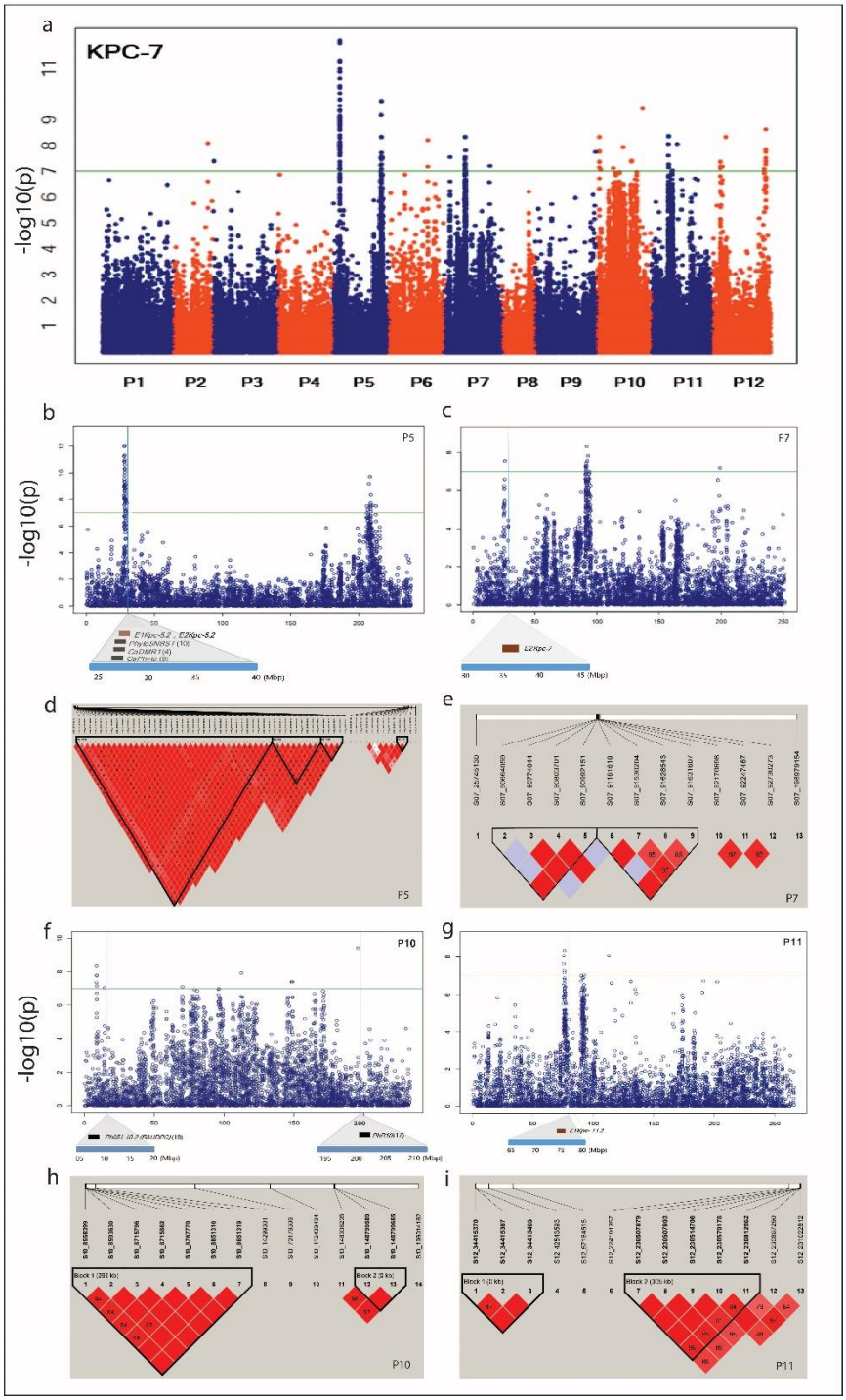


Figure I- 8. Manhattan plots based on GBS-GWAS showing the significant SNPs associated with PcRR resistance and haplotype analysis. (a) Significant SNPs associated with PcRR isolate KPC-7. (b, c, f, g) SNPs detected in common regions of GWAS and ECRIL QTL maps from this and previous studies. QTL positions are marked with blocks under the SNP positions. Brown and black blocks indicate QTLs detected in this and previous studies, respectively. (d, e, h, i) Haplotype blocks containing significant SNPs on chromosomes P5, P7, P10 and P11

Table I- 8. Significant SNPs associated above the threshold -log10 p-values for PcRR isolates KPC-7 and JHAI1-7 (bold font) in the GWAS core collection.

SNP-ID	Chr.	Position (bp)	Alleles	*p*-value	-log10(p)	FDR_*p*-value	MAF	R^2 with SNP	Effect Est.
S02_149068682	P2	149,068,682	G/A	8.13E-09	8.09	8.77E-05	0.0430	0.2332	0.5525
S03_5274098	P3	5,274,098	G/A	4.01E-08	7.40	2.43E-04	0.0415	0.2244	0.5288
S05_27703815	P5	27,703,815	G/A	5.84E-12	11.23	7.41E-07	0.0920	0.2745	0.4714
S05_27761792	P5	27,761,792	C/T	8.67E-10	9.06	1.69E-05	0.0623	0.2457	-0.4830
S05_27807397	P5	27,807,397	T/G	4.97E-09	8.30	5.87E-05	0.3635	0.2359	0.3283
S05_27816851	P5	27,816,851	T/C	1.16E-09	8.93	2.11E-05	0.0890	0.2441	0.4075
S05_27816914	P5	27,816,914	T/C	1.45E-11	10.84	9.20E-07	0.0742	0.2692	0.4948
S05_27816918	P5	27,816,918	G/A	1.45E-11	10.84	9.20E-07	0.0742	0.2692	0.4948
S05_27879870	P5	27,879,870	A/T	1.05E-12	11.98	2.66E-07	0.0682	0.2846	-0.5507
S05_27950402	P5	27,950,402	A/G	2.49E-08	7.60	1.81E-04	0.3220	0.2270	-0.2961
S05_27950407	P5	27,950,407	A/G	8.81E-09	8.06	8.77E-05	0.3650	0.2327	-0.3247
S05_27950442	P5	27,950,442	T/C	8.81E-09	8.06	8.77E-05	0.3650	0.2327	0.3247
S05_27950452	P5	27,950,452	A/T	7.43E-09	8.13	8.20E-05	0.3591	0.2337	-0.3236
S05_27950453	P5	27,950,453	G/A	2.92E-10	9.54	9.25E-06	0.2582	0.2519	0.4017
S05_28046041	P5	28,046,041	G/A	1.56E-10	9.81	5.67E-06	0.1172	0.2555	0.4176
S05_28046047	P5	28,046,047	A/G	9.78E-09	8.01	9.19E-05	0.2745	0.2322	-0.3710
S05_28046048	P5	28,046,048	G/A	4.81E-12	11.32	7.41E-07	0.0816	0.2756	0.4878
S05_28122588	P5	28,122,588	A/T	2.18E-09	8.66	3.35E-05	0.0593	0.2405	-0.4749
S05_28122629	P5	28,122,629	G/A	2.18E-09	8.66	3.35E-05	0.0593	0.2405	0.4749
S05_28123055	P5	28,123,055	G/A	1.27E-09	8.90	2.22E-05	0.2211	0.2436	0.3980
S05_28129446	P5	28,129,446	C/T	1.41E-11	10.85	9.20E-07	0.0786	0.2694	-0.4893
S05_28129512	P5	28,129,512	T/C	8.18E-10	9.09	1.66E-05	0.3739	0.2461	0.3859
S05_28134305	P5	28,134,305	G/A	4.72E-10	9.33	1.09E-05	0.0772	0.2492	0.4413
S05_28134313	P5	28,134,313	C/T	7.10E-09	8.15	8.01E-05	0.0742	0.2339	-0.4136
S05_28134344	P5	28,134,344	C/T	4.72E-10	9.33	1.09E-05	0.0772	0.2492	-0.4413
S05_28134355	P5	28,134,355	C/T	4.72E-10	9.33	1.09E-05	0.0772	0.2492	-0.4413
S05_28134380	P5	28,134,380	T/C	9.72E-09	8.01	9.19E-05	0.1157	0.2322	0.3691
S05_28136336	P5	28,136,336	T/C	4.32E-10	9.36	1.09E-05	0.0653	0.2497	0.4770
S05_28136347	P5	28,136,347	G/A	4.32E-10	9.36	1.09E-05	0.0653	0.2497	0.4770
S05_28163580	P5	28,163,580	G/A	1.06E-11	10.98	9.20E-07	0.0712	0.2710	0.5088
S05_28163627	P5	28,163,627	G/A	4.78E-11	10.32	2.21E-06	0.0757	0.2623	0.4852
S05_28225171	P5	28,225,171	A/G	4.32E-08	7.36	2.52E-04	0.2389	0.2240	-0.3286
S05_28225177	P5	28,225,177	A/G	2.53E-09	8.60	3.56E-05	0.1276	0.2397	-0.4253
S05_28225178	P5	28,225,178	C/T	5.50E-11	10.26	2.32E-06	0.0712	0.2615	-0.4874
S05_28225198	**P5**	***28,225,198**	**A/G**	**8.62E-13**	**12.06**	**2.66E-07**	**0.0653**	**0.2858**	**-0.5642**
S05_28261683	P5	28,261,683	G/C	9.74E-10	9.01	1.83E-05	0.1053	0.2451	0.4090
S05_28261689	P5	28,261,689	C/T	2.51E-11	10.60	1.28E-06	0.0727	0.2660	-0.4909
S05_28261695	P5	28,261,695	T/C	9.16E-11	10.04	3.58E-06	0.1261	0.2585	0.4061
S05_28261710	P5	28,261,710	C/T	2.51E-11	10.60	1.28E-06	0.0727	0.2660	-0.4909
S05_28274277	P5	28,274,277	C/G	1.97E-08	7.71	1.47E-04	0.0712	0.2283	-0.4071
S05_28393613	P5	28,393,613	C/T	1.69E-09	8.77	2.77E-05	0.0801	0.2420	-0.4358
S05_28524281	P5	28,524,281	T/C	7.99E-08	7.10	3.59E-04	0.3516	0.2206	0.3294
S05_28524343	P5	28,524,343	C/T	3.95E-09	8.40	5.42E-05	0.0549	0.2372	-0.4939
S05_28634197	**P5**	**28,634,197**	**G/A**	**5.65E-08**	**7.25**	**2.99E-04**	**0.0653**	**0.2225**	**0.4242**
S05_28670550	P5	28,670,550	T/C	7.50E-08	7.12	3.53E-04	0.2685	0.2209	0.3536
S05_28670630	P5	28,670,630	A/G	1.51E-08	7.82	1.28E-04	0.0935	0.2298	-0.3745
S05_28673591	P5	28,673,591	C/T	1.49E-09	8.83	2.52E-05	0.2018	0.2427	-0.3609
S05_28673596	P5	28,673,596	G/T	7.13E-10	9.15	1.51E-05	0.1944	0.2468	-0.4116
S05_28676908	P5	28,676,908	T/C	1.35E-08	7.87	1.20E-04	0.3828	0.2304	0.3400
S05_28676914	P5	28,676,914	G/T	1.12E-08	7.95	1.03E-04	0.3739	0.2314	-0.3364
S05_28676957	P5	28,676,957	T/C	9.34E-09	8.03	9.12E-05	0.3813	0.2324	0.3435
S05_28825990	P5	28,825,990	T/C	7.81E-08	7.11	3.57E-04	0.2255	0.2207	0.3100
S05_29768389	P5	29,768,389	G/A	1.84E-08	7.74	1.39E-04	0.0386	0.2287	0.5791
S05_206043102	P5	206,043,102	G/A	3.11E-08	7.51	2.02E-04	0.1261	0.2258	0.3289
S05_206565568	P5	206,565,568	C/T	5.48E-08	7.26	2.95E-04	0.1261	0.2227	-0.3204
S05_207549766	P5	207,549,766	G/A	6.50E-10	9.19	1.44E-05	0.1469	0.2474	0.3360
S05_208290460	P5	208,290,460	G/A	1.91E-10	9.72	6.48E-06	0.0608	0.2543	0.5171
S05_208664944	P5	208,664,944	A/G	3.47E-08	7.46	2.22E-04	0.0519	0.2252	-0.4819
S05_208711022	P5	208,711,022	A/G	2.70E-08	7.57	1.93E-04	0.0445	0.2266	-0.5330
S05_208711026	P5	208,711,026	A/C	2.85E-08	7.55	1.94E-04	0.0445	0.2263	-0.5324
S05_208711043	P5	208,711,043	G/A	1.81E-08	7.74	1.39E-04	0.0430	0.2288	0.5456
S05_208814172	P5	208,814,172	C/T	5.27E-08	7.28	2.91E-04	0.0504	0.2229	-0.4766
S05_208958802	P5	208,958,802	G/A	7.77E-08	7.11	3.57E-04	0.0460	0.2207	0.4919
S05_208963284	P5	208,963,284	G/A	4.47E-09	8.35	5.70E-05	0.0504	0.2365	0.5150
S05_210409729	P5	210,409,729	G/T	2.87E-08	7.54	1.94E-04	0.0371	0.2262	-0.5652
S05_212675435	P5	212,675,435	C/T	3.08E-08	7.51	2.02E-04	0.1306	0.2258	-0.3159
S06_171517874	P6	171,517,874	G/A	6.35E-08	7.20	3.22E-04	0.0712	0.2218	0.3978
S06_171572930	P6	171,572,930	G/A	6.01E-09	8.22	6.94E-05	0.0638	0.2349	0.4568
S07_25745130	P7	25,745,130	A/G	2.78E-08	7.56	1.94E-04	0.0356	0.2264	-0.5825
S07_90664059	P7	90,664,059	G/A	5.17E-08	7.29	2.91E-04	0.0950	0.2230	0.3758
S07_90774844	P7	90,774,844	G/A	5.52E-08	7.26	2.95E-04	0.1098	0.2226	0.3590
S07_90803701	P7	90,803,701	A/G	9.04E-08	7.04	3.98E-04	0.2478	0.2199	-0.3161
S07_90992151	P7	90,992,151	T/G	5.83E-08	7.23	3.02E-04	0.1083	0.2223	0.3601
S07_91181618	P7	91,181,618	A/T	3.50E-08	7.46	2.22E-04	0.0920	0.2251	-0.3890

S07_91530204	P7	91,530,204	G/T	7.32E-08	7.14	3.47E-04	0.1142	0.2211	-0.3594
S07_91628543	P7	91,628,543	A/G	4.71E-09	8.33	5.70E-05	0.0846	0.2362	-0.4223
S07_91631607	P7	91,631,607	C/A	4.93E-08	7.31	2.81E-04	0.1202	0.2232	0.3621
S07_92170698	P7	92,170,698	C/T	2.73E-08	7.56	1.93E-04	0.1083	0.2265	-0.3666
S07_92247467	P7	92,247,467	C/T	5.83E-08	7.23	3.02E-04	0.1128	0.2223	-0.3666
S07_92730273	P7	92,730,273	C/A	1.51E-08	7.82	1.28E-04	0.1202	0.2298	0.3683
S07_198979154	P7	198,979,154	A/G	6.19E-08	7.21	3.18E-04	0.1973	0.2220	-0.2889
S09_263032829	P9	263,032,829	G/A	1.73E-08	7.76	1.37E-04	0.0519	0.2290	0.4977
S10_8558399	P10	8,558,399	C/T	4.40E-08	7.36	2.54E-04	0.0519	0.2239	-0.4906
S10_8593630	P10	8,593,630	G/T	4.52E-09	8.35	5.70E-05	0.0608	0.2365	-0.4973
S10_8715796	P10	8,715,796	G/T	1.61E-08	7.79	1.32E-04	0.0593	0.2294	-0.4888
S10_8715862	P10	8,715,862	G/A	1.61E-08	7.79	1.32E-04	0.0593	0.2294	0.4888
S10_8767770	P10	8,767,770	C/T	1.70E-08	7.77	1.37E-04	0.0668	0.2291	-0.4688
S10_8851316	P10	8,851,316	A/G	7.00E-08	7.15	3.38E-04	0.0475	0.2213	-0.5002
S10_8851319	P10	8,851,319	C/T	7.00E-08	7.15	3.38E-04	0.0475	0.2213	-0.5002
S10_14299301	P10	14,299,301	G/A	8.90E-08	7.05	3.96E-04	0.1439	0.2200	0.2920
S10_70178306	P10	70,178,306	A/G	7.67E-08	7.12	3.57E-04	0.3887	0.2208	0.2894
S10_112433434	P10	112,433,434	T/G	1.15E-08	7.94	1.04E-04	0.3516	0.2313	-0.2897
S10_148308205	P10	148,308,205	C/G	3.97E-08	7.40	2.43E-04	0.2433	0.2244	0.2909
S10_148799589	P10	148,799,589	T/C	3.90E-08	7.41	2.43E-04	0.2507	0.2245	-0.2966
S10_148799665	P10	148,799,665	T/G	4.02E-08	7.40	2.43E-04	0.2493	0.2244	-0.2969
S10_196014162	P10	196,014,162	T/A	3.78E-10	9.42	1.09E-05	0.0430	0.2504	0.6209
S11_75334209	P11	75,334,209	C/A	8.67E-09	8.06	8.77E-05	0.1024	0.2328	0.3792
S11_75334214	P11	75,334,214	C/T	2.37E-08	7.62	1.75E-04	0.1068	0.2273	-0.3602
S11_75972407	P11	75,972,407	G/A	5.37E-08	7.27	2.93E-04	0.1053	0.2228	0.3609
S11_75972413	P11	75,972,413	A/T	4.37E-09	8.36	5.70E-05	0.0979	0.2367	-0.4041
S11_75972450	P11	75,972,450	G/A	6.76E-08	7.17	3.33E-04	0.1113	0.2215	0.3493
S11_90049314	P11	90,049,314	G/A	9.90E-08	7.00	4.29E-04	0.1113	0.2194	0.3446
S11_92434227	P11	92,434,227	C/T	9.09E-08	7.04	3.98E-04	0.1068	0.2199	-0.3533
S11_112878746	P11	112,878,746	G/A	8.62E-09	8.06	8.77E-05	0.0623	0.2329	0.4743
S12_34415370	P12	34,415,370	G/A	4.28E-08	7.37	2.52E-04	0.0623	0.2240	0.4499
S12_34415387	P12	34,415,387	C/A	7.25E-08	7.14	3.47E-04	0.0638	0.2211	0.4399
S12_34415405	P12	34,415,405	G/A	6.51E-08	7.19	3.24E-04	0.0445	0.2217	0.5157
S12_42543593	P12	42,543,593	A/G	6.40E-08	7.19	3.22E-04	0.3145	0.2218	0.2587
S12_57164515	P12	57,164,515	G/A	4.64E-09	8.33	5.70E-05	0.1365	0.2363	0.3546
S12_224191357	P12	224,191,357	A/T	7.95E-08	7.10	3.59E-04	0.0564	0.2206	-0.4523
S12_230507879	P12	230,507,879	C/A	2.34E-09	8.63	3.40E-05	0.0549	0.2401	0.5094
S12_230507903	P12	230,507,903	C/G	2.34E-09	8.63	3.40E-05	0.0549	0.2401	-0.5094
S12_230514706	P12	230,514,706	C/T	1.78E-08	7.75	1.39E-04	0.0593	0.2288	-0.4635
S12_230579178	P12	230,579,178	C/T	4.11E-08	7.39	2.45E-04	0.0593	0.2242	-0.4508
S12_230812962	P12	230,812,962	T/C	1.38E-08	7.86	1.21E-04	0.0579	0.2302	0.4852
S12_230887290	P12	230,887,290	C/G	3.02E-08	7.52	2.02E-04	0.0534	0.2259	-0.4727
S12_231022812	P12	231,022,812	G/T	5.21E-08	7.28	2.91E-04	0.0549	0.2229	-0.4674

*Position of detected SNPs in pepper reference genome version 1.6.

*Bold SNPs were commonly detected against KPC-7 and JHAI107 isolates.

Table I- 9. Haplotype blocks estimated by genotyping-by-sequencing of the GWAS core collection.

Chromosome	Haplotype	Average Size of LD Blocks (kb)	Average Number of SNPs Per LD Block
P1	3,267	76.4	15.6
P2	1,866	71.1	14.2
P3	2,723	83.0	15.7
P4	2,226	88.2	17.0
P5	2,321	84.5	17.2
P6	2,582	75.1	14.9
P7	2,548	81.2	16.9
P8	1,270	96.0	18.3
P9	2,623	84.3	16.7
P10	2,788	64.9	13.4
P11	2,953	72.6	16.8
P12	2,102	97.6	17.5
Total	29,269	81.2	16.2

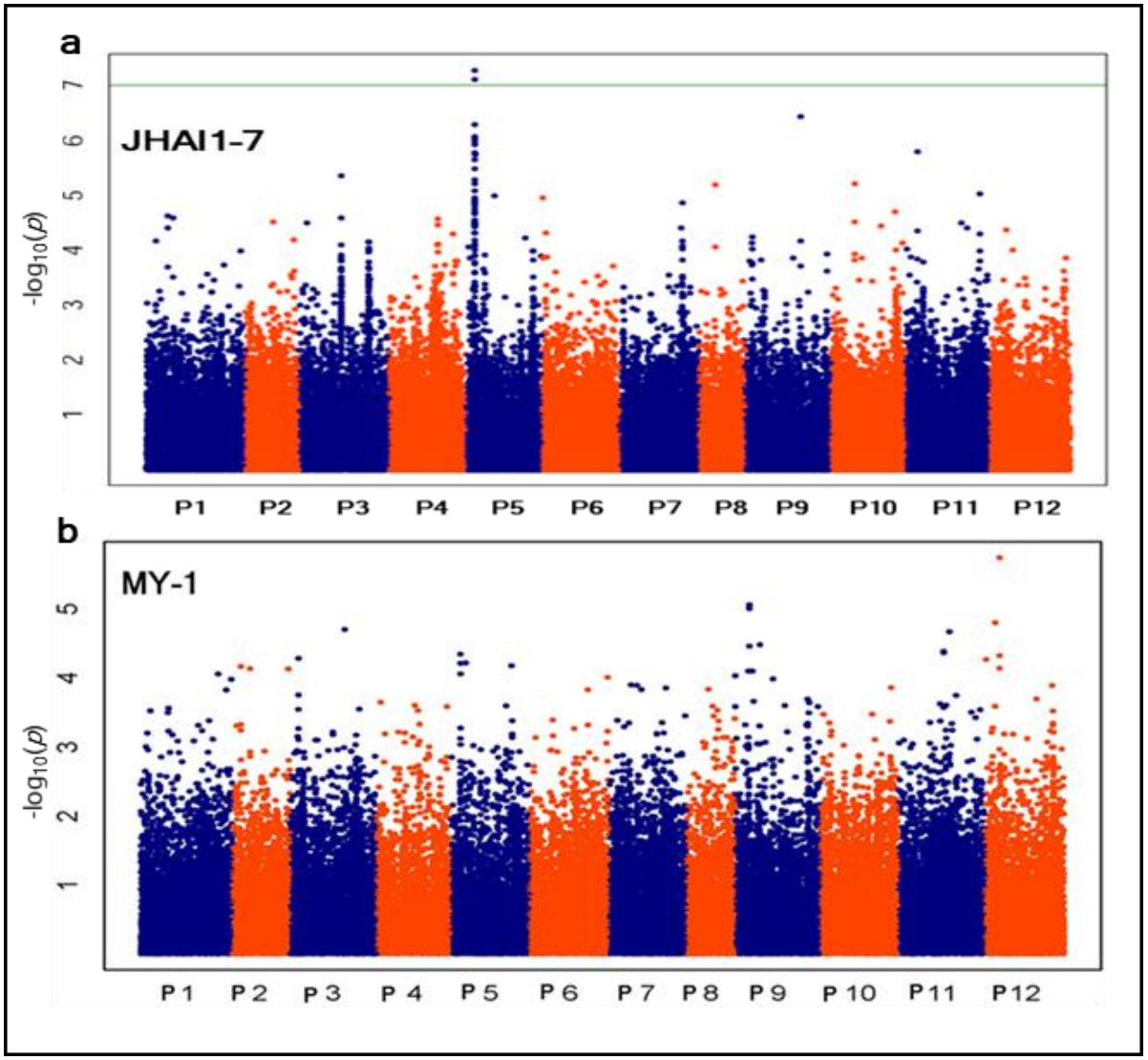


Figure I- 9. Manhattan plots based on GBS-GWAS for PcRR resistance. (a) Significant SNPs associated with PcRR isolate JHAI1-7 (b) against low-virulent isolate MY-1.

Predicted candidate genes for PcRR.

I searched for candidate PcRR resistance genes in the 1 Mb upstream and downstream flanking regions of all significant GWAS-SNPs and bin markers linked to PcRR QTLs. A major and colocalized genomic region detected for isolate KPC-7 via GWAS and biparental QTL mapping (*QTL5.2*) associated with PcRR resistance encompassing a 2.5 Mb region (27.0–29.5 Mb) (Fig. I-7 and Fig. I-8a) contains 20 predicted genes, including genes for three receptor-like kinase (RLK) domain-containing proteins, two nucleotide-binding site-leucine-rich repeat (NBS-LRR) domain-containing proteins, and one SAR8.2 precursor protein known to be associated with disease resistance (Table I-10). Two NBS-LRR genes (*PHT81215.1* and *PHT81216.1*) are located 618 and 319 kb upstream of the most significant SNP (S05_27703815), respectively. The most closely located RLK gene (*PHT81221.1*) was identified 144 kb upstream of a highly significant GWAS SNP. Two RLK genes (*PHT81227.1* and *PHT81229.1*) were identified at 5.5 and 12.4 kb downstream of GWAS (SNP S05_28163627). One candidate gene, encoding the SAR8.2 precursor protein, was detected 148 kb upstream of GWAS SNP (S05_28825990) (Table I-10). These genes were detected by both QTL mapping and GWAS analysis, making them strong candidate genes for PcRR resistance.

I used the highly significant GWAS SNP (S05_27703815), with -log10(p) value of 11.23 (Table I-8), located close to the candidate genes to compare the resistance levels of the GWAS accessions. As shown in the box plots in Fig. I-10c, the homozygous genotype (S05_27703815) AA was associated with increased resistance compared to the alternative homozygous genotype GG among the three PcRR isolates examined. I used the tightly linked bin marker EC5-bin40 of QTL segment *QTL5.2* overlapping with candidate genes to compare the resistance levels of the ECRILs. As shown in the box plots, the homozygous resistant genotype BB is associated with enhanced resistance compared to the homozygous susceptible

genotype AA against all three isolates in both E1 and E2 (Fig. I-10a and 10b).

The major QTL region *QTL5.1*, spanning 0.7 Mb, which was commonly detected against two isolates, JHAI1-7 and MY-1, contains 15 predicted genes, including genes associated with disease resistance, such as genes for six RLK domain-containing proteins and four NBS-LRR domain-containing proteins (Table I-10). Finally, I used the tightly linked bin marker EC5-bin27 of QTL segment *QTL5.1* overlapping with candidate genes to compare the resistance levels of the ECRILs. As shown in the box plots, the homozygous resistant genotype BB is associated with enhanced resistance compared to the homozygous susceptible genotype AA against all three isolates in both E1 and E2 (Fig. I-11a and 12b).

Finally, *QTL5.3*, a QTL region against isolates KPC-7 and MY-1, was detected at 34.6–37 Mb on chromosome P5, with 28 predicted genes. Among these, three predicted genes are linked to disease resistance proteins, including two NBS-LRR domain-containing proteins and one Kunitz trypsin inhibitor 2-like protein (Table I-10). I used the tightly linked bin marker EC5-bin51 of QTL segment *QTL5.3* overlapping with the candidate genes to compare the resistance levels of the ECRILs. As shown in the box plots, the homozygous resistant genotype BB is associated with enhanced resistance compared to homozygous susceptible genotype AA against all three isolates in both E1 and E2 (Fig. I-11c and 11d).

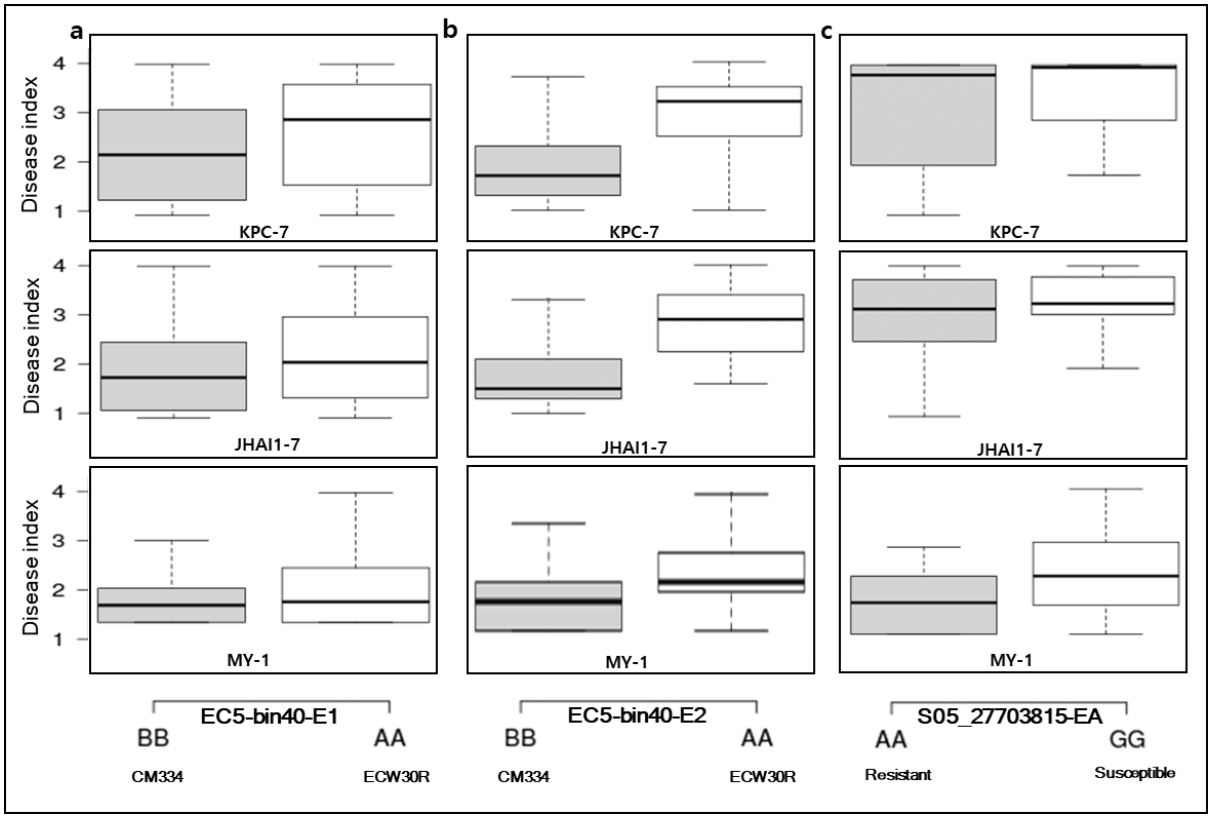


Figure I- 10. Box plots of tightly linked bins to QTLs in ECRILs and significantly associated GWAS-SNPs from chromosome P5. (a) ECRILs grouped based on the tightly linked bin to QTL 5.2 in environment E1. (b) ECRILs grouped based on the tightly linked bin to QTL 5.2 in environment E2. (3) Core collection GWAS (core collection) grouped based on the most significant SNP from QTL region 5.2.

Table I- 10. Candidate resistance genes associated with the significant GWAS regions and QTLs for PcRR isolate KPC-7 and their Gene Ontology (GO) descriptions.

Chr.	Mapping strategy (GWAS/QTL)	Location (bp)	Peptide ID[a]	GO Description
P5	GWAS-*QTL5.2*	27085614-27086321	PHT81215.1	PREDICTED: disease resistance protein RPP13-like
P5	GWAS-*QTL5.2*	27311924-27314491	PHT81216.1	PREDICTED: disease resistance protein RPP13-like
P5	GWAS-*QTL5.2*	27558828-27565702	PHT81221.1	Inactive leucine-rich repeat receptor-like serine/threonine-protein kinase
P5	GWAS-*QTL5.2*	28168701-28169191	PHT81227.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	GWAS-*QTL5.2*	28172631-28176106	PHT81228.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	GWAS-*QTL5.2*	28974813-28975852	PHT81231.1	SAR8.2 precursor
P5	*QTL 5.1*	18707644-18708354	PHT81101.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	*QTL 5.1*	18810592-18813153	PHT81104.1	PREDICTED: disease resistance protein RPP13-like
P5	*QTL 5.1*	18943734-18945083	PHT81107.1	PREDICTED: disease resistance protein RPH8A-like
P5	*QTL 5.1*	19074847-19077759	PHT81108.1	PREDICTED: disease resistance protein RPP13-like
P5	*QTL 5.1*	19094517-19097099	PHT81109.1	PREDICTED: disease resistance protein RPP13-like
P5	*QTL 5.1*	19153252-19169266	PHT81110.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	*QTL 5.1*	19288855-19292090	PHT81111.1	PREDICTED: LRR receptor-like serine/threonine-protein kinase EFR
P5	*QTL 5.1*	19292463-19293858	PHT81112.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	*QTL 5.1*	19294140-19297023	PHT81113.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	*QTL 5.1*	19344986-19347916	PHT81114.1	PREDICTED: probable LRR receptor-like serine/threonine-protein kinase
P5	*QTL 5.3*	35411316-35411978	PHT81306.1	Kunitz trypsin inhibitor 2-like
P5	*QTL 5.3*	36088454-36089143	PHT81317.1	putative late blight resistance protein homolog R1B-17
P5	*QTL 5.3*	36205447-36219154	PHT81318.1	PREDICTED: putative late blight resistance protein homolog R1C-3

[a]Peptide id given in the NCBI genome data base

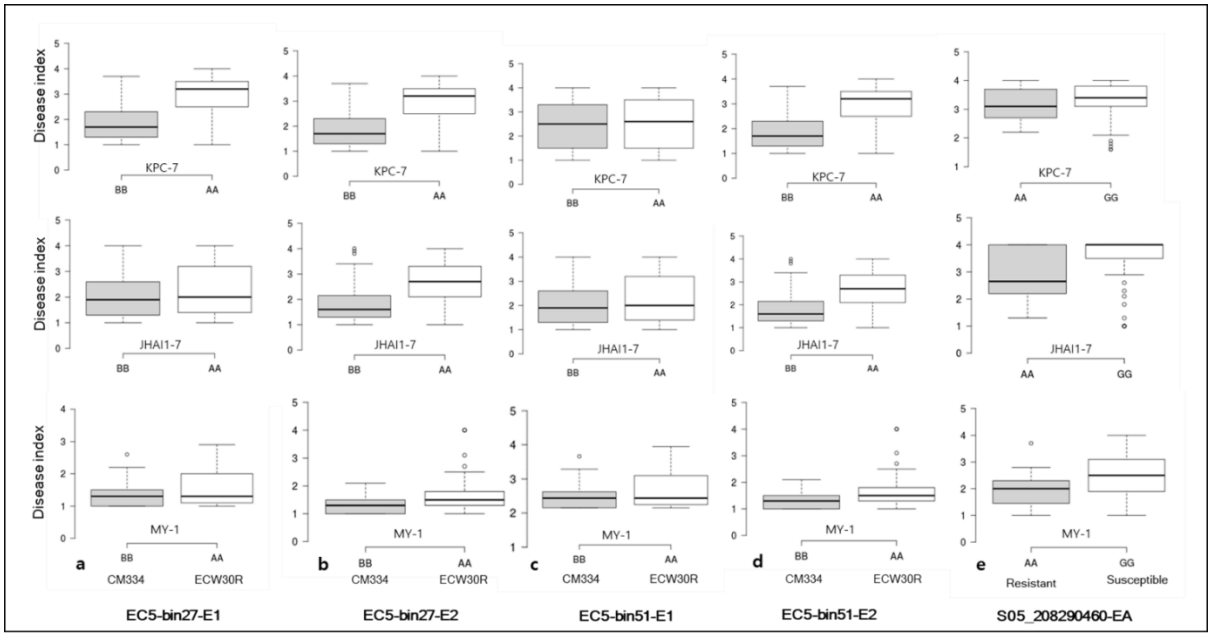


Figure I- 11. Box plots of the tightly linked bin to PcRR resistance QTLs and GWAS-SNPs associated with PcRR resistance against three isolates. (a and b) ECRILs grouped based on the tightly linked bin to QTL 5.1 in environments E1 and E2. (c and d) ECRILs grouped based on the tightly linked bin to QTL 5.3 in environments E1 and E2. (e) Core collection population grouped based on the most significant SNP on the lower arm of chromosome P5.

DISCUSSION

Rapid genetic evolution and diversity in the *Phytophthora* urges multiple isolates disease resistance screening to PcRR in current pepper breeding (Naegele et al. 2014; Rehrig et al. 2014). Many QTLs were earlier reported for PcRR (Kim et al. 2008; Mallard et al. 2013; Rehrig et al. 2014; Wang et al. 2016) but there is no any such study, which combined genetic investigation for PcRR resistance loci using GBS-GWAS approach and further validated by QTL mapping with multiple virulence isolates in multiple environments to accelerate pepper breeding.

Genetic diversity in the capsicum for PcRR resistance with different isolates

Differential response of capsicum (GWAS accessions and ECRILs) to three PcRR isolates provides valuable information about range and frequency distribution of symptom development and virulence level (Fig. I-2). The mean disease scores of GWAS accessions were lower than ECRILs (Table I-2) which indicates that higher allelic diversity to PcRR resistance in the genetically diverse GWAS accessions. Coefficients of variation values were consistently lower (ECRILs < GWAS accessions) for highly virulent isolate KPC-7 than other two isolates. Furthermore, results of very high heritability and genetic advancement for PcRR disease score emphasize the association of additive gene action and close by is an opportunity for improvement in this trait through selection, these results were in accordance with previous reports of Naegele et al. (2014) in the different population background.

GBS approach for QTLs and candidate genes uncovering to PcRR resistance

Further, CM334 is a prominent resistant source of PcRR. The inheritance of PcRR resistance in CM334 is complex (Liu et al. 2014; Truong et al. 2012) and genetic dissection of PcRR resistance can be of great use for improving pepper cultivars. For precise genetic mapping and dissection of the complex traits, high-density linkage maps are essential. In this study, I developed and used different mapping populations and *P. capsici* isolates for linkage and association mapping to detect the genome-wide novel loci conferring resistance to PcRR by using high throughput genome-wide SNP identified through GBS approach.

Reliable markers tightly linked to resistant loci or genomic regions controlling disease resistance through quantitative trait loci (QTL) mapping and association studies are important for MAS. Several types of genetic markers have been developed and being utilized in pepper breeding programs (Rehrig et al. 2014; Wang et al. 2016; Xu et al. 2016), however, very few diagnostic markers are available for PcRR. Therefore, to make use of additional resistance loci more diagnostic markers are need to be developed. GBS is a simple and cost effective approach for SNP markers identification and high-resolution map construction. The genome-wide identification of SNPs for mapping of complex traits through GBS has been performed in several crops, including cabbage, chickpea, common bean and wheat (Gao et al. 2015; Hart and Griffiths 2015; Jaganathan et al. 2015; Li et al. 2015). I constructed a high-density bin map using 8,979 SNPs derived from GBS distributed on 12 pepper chromosomes covering a total genetic distance of 1,330 cM with an average interval of 0.6 cM. The density of SNP markers was consistent with previous reports of linkage maps constructed in pepper (Han et al. 2018).

Breeding for resistance to PcRR is quite challenging, even resistance incorporated into the commercial pepper lines could be readily overcome by highly virulent *P. capsici* isolates

(Foster and Hausbeck 2010). Therefore, I aimed at identifying QTLs for different *P. capsici* isolates categorized based on virulence. In our study, QTL mapping for PcRR resistance has revealed twelve significant QTLs, including four major effect QTLs on P5 and eight minor effect QTLs on P2, P4, P8, and P11 against medium and high virulent isolates of *P. capsici*. Among several QTLs identified for PcRR resistance, a major QTL on P5 has been steadily identified in several resistance sources irrespective of race, isolates and level of virulence (Rehrig et al. 2014). In agreement with this, in our study the two common loci *Kpc5.1 Kpc5.2*, *Jha5.1* and *Jha5.2* were consistently detected on P5 at 24.2 and 30.5 cM against medium (JHAI1-7) and high virulent (KPC-7) isolates of *P. capsici*. The locations of these QTLs were close to the previously identified *Pc5.1* and *Pc5.2* (Mallard et al. 2013; Rehrig et al. 2014). Previously, in the close genomic region of *Pc5.1*, dominant resistant locus, *CaPhyto* conferring race specific resistance was also mapped in a genetic interval of 3.3-cM region between two SSR markers, *ZL6726* and *ZL6970* in *C. annuum* PI201234 (Wang et al. 2016). Several minor effect QTLs have previously been reported on chromosomes P1, P3, P4, P6, P8, P9, P10, P11, and P12 (Bonnet et al. 2007; Kim et al. 2008; Ogundiwin et al. 2005; Rehrig et al. 2014; Thabuis et al. 2003; Truong et al. 2012). The number of QTLs and their positions conferring resistant to *P. capsici* in different studies vary depending on mapping populations used, disease screening methods, inoculum density and isolate aggressiveness. In our study, I detected some minor QTLs with isolate specificity on P2 (*Kpc2.1* and *Kpc2.1* for KPC-7) and P4 (for *Jha4.1* and *Jha8.1* for JHAI1-7) and with isolate non-specific on P11 (*Kpc11.1*, *Kpc11.2*, *Jha11.1* and *Jha11.2*). However, I could not compare the exact physical positions of these QTLs with previously detected QTLs because of the limited number of common markers and lack of sequence information.

Validation of PcRR loci with GWAS

Additionally, bi-parental QTL mapping was complimented with GWAS to validate known QTL regions and identify new PcRR resistance related loci. A total of 98 significant SNPs related to PcRR were identified across different chromosomes however, only SNPs identified on P2, P5, and P11 were found to be co-located with QTLs detected in the present study. This might be due to the use of highly diverse accessions with limited resistant sources in GWAS (Lee et al. 2016). In previous reports, resistant loci against PcRR has been mapped on chromosomes P1, P3, P4, P6, P8, P9, P10, P11, and P12 (Bonnet et al. 2007; Kim et al. 2008; Ogundiwin et al. 2005; Rehrig et al. 2014; Thabuis et al. 2003; Truong et al. 2012).The SNP markers on P5 at positions 28,444 Kb, 28,019 kb and 28,361 kb were the most significant markers and found to be co-localized with the QTL detected on P5 in previous studies (between 28 Mb and 32 Mb) (Mallard et al. 2013; Rehrig et al. 2014) (Fig. I-4c). The SNP marker located at 28,013 kb on P11 is co-localized with the previously identified QTL *Ph127-11* (*RAUDPC*) for PcRR resistance (Truong et al. 2012) (Fig. I-4c). Regions detected by GWAS with significant SNPs on P10 (between 24-26 and 196 Mb) found to be co-localize with *Ph051-10-3/4(RAUDPC)* (Truong et al. 2012) and *PhR10* (Xu et al. 2016), respectively (Fig. I-4c).

Candidate gene identification for PcRR resistance

To ward off the pathogen attack, plants depend on two classes of immune receptors; intracellular (NBS-LRRs) and extra-cellular cell surface (also called pattern recognition) receptors (PRRs) (Du et al. 2015; Witek et al. 2016). Plant genomes contain a large number of NBS-LRR class of *R* genes, which is often clustered in the specific chromosome because of tandem and segmental duplications (Huang et al. 2005; Hulbert et al. 2001; Leister 2004; McDowell and Simon 2006). Several dominant *R* genes, such as *R1* gene from potato (Ballvora

et al. 2002), *Ph-3* from tomato (Zhang et al. 2014), *RpsUN1, RpsUN2,* and *Rpg1-b* genes from soybean (Ashfield et al. 2003; Li et al. 2016) are belong to the NBS-LRR domain containing complex *R* gene clusters. In the present study, cluster of NBS-LRR domain containing *R* genes; five NBS-LRRs from *qtl5.1*, four NBS-LRRs from *qtl5.2* and *qtl11.1* and a single NBS-LRR from *qtl11.2* were identified. These candidate genes showed high sequence identity to the RPP13-like NBS-LRR protein genes and represent potential candidates for PcRR resistance in pepper. The *RPP13* gene belong to the CC-NBS-LRR domain containing *R* gene known to confer resistance to the oomycete pathogen, *Peronospora parasitica*, in *Arabidopsis* (Rose et al. 2004). It is reported that the *R* gene clusters could be hotspots for novel *R* genes and enhance the possibility of structural and copy number variation through DNA recombination and rearrangement mechanisms, including duplications, unequal crossing over, gene conversion, and diversifying selection and could account for gain or loss of resistance (Li et al. 2016; Michelmore and Meyers 1998; Nagy and Bennetzen 2008). The candidate *R* genes clusters identified in the present study could account for enhanced durability and resistance to PcRR in CM334 pepper.

In addition to the *R* gene clusters, I identified several RLK like genes; eight RLKs from *Kpc-5.1* and *Jha-5.1*, five RLKs from *Kpc-5.2* and *Jha-5.2*, eight RLK from *Kpc-11.1*and *Jha-11.1*, and a single RLK from *Kpc-11.2* and *Jha11.2*. In plants, the first line of defense against pathogens is triggered by PRRs, which involves the recognition of conserved pathogen-associated molecular patterns (PAMPs) (Du et al. 2015; Witek et al. 2016). RLKs as PRRs, play crucial roles in plant defense-related processes, including both host and non-host defense responses (Torii 2004). In the present study, I identified two *MLO*-like protein genes from the *Kpc-2.1* and *Jha-8.1*, which could be additional potential candidate genes for PcRR resistance in pepper. However, it remains to be determined whether MLO and RLK proteins are indeed

co-function in PcRR resistance pepper. Loss-of-function alleles of specific *Mildew resistance locus O* (*MLO*) like protein genes are known to be associated with powdery mildew resista nce in several plant species (Humphry et al. 2011; Pavan et al. 2011; Zheng et al. 2013). Intriguingly, in pepper *CaMLO2* was reported to be involved in the regulation of biotic stress responses, including *P. capsici* (Kim and Hwang 2012). It was suggested that CaMLO2 required for oomycetal hyphal growth and the formation of sporangiophores and spores (Kim and Hwang 2012). Furthermore, several lines of evidences support the concerted action of MLO and RLK proteins in same signal transduction pathways, for instance, pollen tube perception and fungal invasion process (Humphry et al. 2011; Kessler et al. 2010; Lalonde et al. 2010). However, to uncover the potential role of these co-localized intra- and extra-cellular immune receptors in disease resistance, further candidate gene approach involving the targeted knock down of candidate genes and subsequent molecular genetic analyses should be performed.

In the present study, I demonstrated the combine use of high-resolution bi-parental QTL mapping and GWAS for identification of closely linked markers and candidate genes for PcRR resistance in pepper. Based on these genetic map and phenotypic data, 12 significant QTLs were identified with PV ranged from 3.7% to 35.4%. Two of these QTLs showed major effects with PV of more than 60% together. GWAS identified 98 highly significant SNPs in association to PcRR across the genome. The additional loci detected through GWAS may enhance the resistance selection accuracy. The SNP markers associated with the candidate genes or QTLs will be used for MAS and genomic selection in pepper breeding programs for PcRR resistant. The predicted candidates for PcRR resistance lay a foundation for candidate gene approach for molecular dissection of PcRR resistance.

REFERENCES

Allard, R. W. *Principles of plant breeding* (John Wiley and Sons Inc.1960).

Ashfield, T. *et al*. Genetic and physical localization of the soybean *Rpg1-b* disease resistance gene reveals a complex locus containing several tightly linked families of NBS-LRR genes. *Mol. Plant Microbe Interact*. **16,** 817-826 (2003).

Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3,** e3376 (2008).

Ballvora, A. *et al*. The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. *Plant J*. **30,** 361-371 (2002).

Bonferroni, C. E. *Teoria statistica delle classi e calcolo delle probabilita*. Florence (Libreria Internazionale Seeber; 1936).

Bonnet, J. *et al*. Are the polygenic architectures of resistance to *Phytophthora capsici* and *P. parasitica* independent in pepper? *Theor. Appl. Genet*. **115,** 253-264 (2007).

Bosland, P. & Lindsey, D. A. seedling screen for Phytophthora root rot of pepper, *Capsicum annuum*. *Plant Dis*. **75,** 1048-1050 (1991).

Brachi, B. *et al.* Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet.* **6,** e1000940 (2010).

Brouwer, D. J. & St. Clair, D. A. Fine mapping of three quantitative trait loci for late blight resistance in tomato using near isogenic lines (NILs) and sub NILs. *Theor. Appl. Genet.* **108,** 628-638 (2004).

Chen, A. L. *et al*. Reassessment of QTLs for late blight resistance in the tomato accession L3708 using a restriction site associated DNA (RAD) linkage map and highly aggressive isolates of *Phytophthora infestans*. *PLoS ONE* **9,** e96417 (2014).

Consortium, T. T. G. The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* **485,** 635–641 (2012).

Cook, D. E. *et al*. Copy number variation of multiple genes at *Rhg1* mediates nematode resistance in soybean. *Science* **338,** 1206-1209 (2012).

Crowell, S. *et al*. Genome-wide association and high-resolution phenotyping link *Oryza sativa* panicle traits to numerous trait-specific QTL clusters. *Nat. Commun*. **7,** 10527 (2016).

Du, J. *et al*. Elicitin recognition confers enhanced resistance to *Phytophthora infestans* in potato. *Nat. Plants* **1,** 15034 (2015).

Elshire, R. J. *et al*. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6,** e19379 (2011).

Foster, J. & Hausbeck, M. Resistance of pepper to Phytophthora crown, root, and fruit rot is affected by isolate virulence. *Plant Dis*. **94,** 24-30 (2010).

Gao, H. & Bhattacharyya, M. K. The soybean-Phytophthora resistance locus *Rps1-k* encompasses coiled coil-nucleotide binding-leucine rich repeat-like genes and repetitive sequences. *BMC Plant Biol.* **8,** 29(2008).

Gao, H., Narayanan, N. N., Ellison, L. & Bhattacharyya, M. K. Two classes of highly similar coiled coil-nucleotide binding-leucine rich repeat genes isolated from the *Rps1-k* locus encode *Phytophthora* resistance in soybean. *Mol. Plant Microbe Interact*. **18,** 1035-1045 (2005).

Gao, L. *et al.* Development of genotyping by sequencing (GBS)-and array-derived SNP markers for stem rust resistance gene *Sr42*. *Mol. Breeding* **35,** 207 (2015).

Guo, S. *et al*. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat. Genet.* **45,** 51-58 (2013).

Han, K. *et al*. An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). *DNA Res*. **23,** 81-91 (2016).

Han, K. *et al*. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. *Plant Biotechnol. J.* **16,** 1546-1558 (2018).

Hart, J. P. & Griffiths, P.D. Genotyping-by-sequencing enabled mapping and marker development for the *by-2* potyvirus resistance allele in common bean. *Plant Genome* **8,** (2015).

Hausbeck, M. K. & Lamour, K. H. *Phytophthora capsici* on vegetable crops: research progress and management challenges. *Plant Dis*. **88,** 1292-1303 (2004).

Huang, S. *et al*. Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. *Plant J*. **42,** 251-261 (2005).

Hulbert, S. H., Webb, C. A., Smith, S. M. & Sun, Q. Resistance gene complexes: evolution and utilization. *Annu. Rev. Phytopathol*. **39,** 285-312 (2001).

Humphry, M., Reinstaedler, A., Ivanov, S., Bisseling, T. & Panstruga, R. Durable broad-spectrum powdery mildew resistance in pea *er1* plants is conferred by natural loss-of-function mutations in *PsMLO1*. *Mol. Plant. Pathol.* **12,** 866-878 (2011).

Jaganathan, D. *et al*. Genotyping-by-sequencing based intra-specific genetic map refines a ‘‘QTL-hotspot” region for drought tolerance in chickpea. *Mol. Genet. Genomics* **290,** 559-571 (2015)

Jo, J., Purushotham, P. M., Han, K., Lee, H-R., Nah, G. & Kang, B-C. Development of a genetic map for onion (*Allium cepa* L.) using reference-free genotyping-by-sequencing and SNP assays. *Front. Plant. Sci*. **8,** 1606 (2017).

Jo, S-J., Shim, S-A., Jang, K. S., Choi, Y. H., Kim, J-C. & Choi, G. J. Resistance of chili pepper cultivars to isolates of *Phytophthora capsici*. *Korean J. Hortic. Sci. Technol.* **32,** 66-76 (2014).

Kamoun, S. *et al*. The top 10 oomycete pathogens in molecular plant pathology. *Mol. Plant. Pathol* **16,** 413-434 (2015).

Kessler, S. A. *et al*. Conserved molecular components for pollen tube reception and fungal invasion. *Science* **330,** 968-971 (2010).

Kim, B. S., Ahn, J. H., Lee, J. M., Park, D. G. & Kim, H. Y. Restorer genotype for male sterile cytoplasm of genetic resources moderately resistant to *Phytophthora capsici* in Capsicum pepper. *Korean J. Hortic. Sci. Technol*. **30,** 73-78 (2012).

Kim, D. S. & Hwang, B. K. The pepper *MLO* gene, *CaMLO2*, is involved in the susceptibility cell-death response and bacterial and oomycete proliferation. *Plant J*. **72,** 843-855 (2012).

Kim, H. J. *et al*. BAC-derived markers converted from RFLP linked to *Phytophthora capsici* resistance in pepper (*Capsicum annuum* L.). *Theor. Appl. Genet*. **118,** 15-27 (2008).

Korte, A. & Farlow, A. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9,** 29 (2013).

Lalonde, S. *et al*. A membrane protein/signaling protein interaction network for *Arabidopsis* version *AMPv2*. *Front. Physiol*. **1,** 24 (2010).

Lamour, K. & Hausbeck, M. Mefenoxam insensitivity and the sexual stage of *Phytophthora capsici* in Michigan cucurbit fields. *Phytopathology* **90,** 396-400 (2000).

Lee, H. Y. *et al*. Genetic diversity and population structure analysis to construct a core collection from a large Capsicum germplasm. *BMC Genet.* **17,** 142 (2016).

Leister, D. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends. Genet.* **20,** 116-122 (2004).

Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26,** 589-595 (2010).

Li, H. *et al*. A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. *BMC Genomics* **16,** 216 (2015).

Li, L. *et al*. Fine mapping and candidate gene analysis of two loci conferring resistance to *Phytophthora sojae* in soybean. *Theor. Appl. Genet.* **129,** 2379-2386 (2016).

Lipka, A. E. *et al*. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28,** 2397-2399 (2012).

Liu, W. Y. *et al*. Combined use of bulked segregant analysis and microarrays reveals SNP markers pinpointing a major QTL for resistance to *Phytophthora capsici* in pepper. *Theor. Appl. Genet.* **127,** 2503-2513 (2014).

Lush, J. L. *Animal breeding plans* 473 (IOWA State Uni. Press Ames., Iowa, 1949).

Mallard, S., Cantet, M., Massire, A., Bachellez, A., Ewert, S. & Lefebvre, V. A. key QTL cluster is conserved among accessions and exhibits broad-spectrum resistance to *Phytophthora capsici*: a valuable locus for pepper breeding. *Mol. Breeding* **32,** 349-364 (2013).

McDowell, J. M. & Simon, S. A. Recent insights into *R* gene evolution. *Mol. Plant. Pathol.* **7,** 437-448 (2006).

Michelmore, R. W. & Meyers, B. C. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8,** 1113-1130 (1998).

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A. & Johnson, E. A. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res*. **17,** 240-248 (2007).

Minamiyama, Y., Tsuro, M., Kubo, T. & Hirai, M. QTL analysis for resistance to *Phytophthora capsici* in pepper using a high density SSR-based map. *Breed. Sci*. **57,** 129-134 (2007).

Naegele R. P. *et al.* QTL mapping of fruit rot resistance to the plant pathogen *Phytophthora capsici* in a recombinant inbred line *Capsicum annuum* population. *Phytopathology*, **104**, 479-483 (2014).

Nagy, E. D. & Bennetzen, J. L. Pathogen corruption and site-directed recombination at a plant disease resistance gene cluster. *Genome Res*. **18**, 1918-1923 (2008).

Ogundiwin, E. A. *et al*. Construction of 2 intraspecific linkage maps and identification of resistance QTLs for *Phytophthora capsici* root-rot and foliar-blight diseases of pepper (*Capsicum annuum L.*). *Genome* **48,** 698-711 (2005).

Pavan, S. *et al*. Pea powdery mildew *er1* resistance is associated to loss-of-function mutations at a *MLO* homologous locus. *Theor. Appl. Genet.* **123,** 1425-1431 (2011).

Quirin, E. *et al*. Development of sequence characterized amplified region (SCAR) primers for the detection of Phyto. 5.2, a major QTL for resistance to *Phytophthora capsici* Leon. in pepper. *Theor. Appl. Genet.* **110,** 605-612 (2005).

Rehrig, W. Z., Ashrafi, H., Hill, T., Prince, J. & Van Deynze, A. *CaDMR1* cosegregates with QTL *Pc5.1* for resistance to *Phytophthora capsici* in pepper (*Capsicum annuum*). *Plant Genome* **7,** (2014).

Rose, L. E., Bittner-Eddy, P. D., Langley, C. H., Holub, E. B., Michelmore, R. W. & Beynon, J. L. The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics* **166,** 1517-1527 (2004).

Sallam, A. Arbaoui, M. El-Esawi, M.,Abshire, N. & Martsch, R. Identification and verification of QTL associated with frost tolerance using linkage mapping and GWAS in winter faba bean. *Front. Plant. Sci.* **7,** 1098 (2016).

Siddique, M. I. et *al*. Resistance to *Phytophthora capsici*, restorer-of-fertility genotype for cytoplasmic male sterility and chemical quality components of breeding lines developed for improvement of subicho, a land race of pepper in Yeongyang. *Hort. Sci. Technol.* **35,** 758-768 (2017).

Sonah, H. O'Donoughue, L. Cober, E. Rajcan, I. & Belzile, F. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. *Plant Biotechnol. J.* **13,** 211-221 (2015).

Sugita, T. *et al*. QTL analysis for resistance to Phytophthora blight (*Phytophthora capsici* Leon.) using an intraspecific doubled-haploid population of *Capsicum annuum*. *Breed. Sci*. **56,** 137-145 (2006).

Thabuis, A. *et al*. Phenotypic and molecular evaluation of a recurrent selection program for a polygenic resistance to *Phytophthora capsici* in pepper. *Theor. Appl. Genet*. **109,** 342-351 (2004).

Thabuis, A., Palloix, A., Pflieger, S., Daubeze, A-M., Caranta, C. & Lefebvre, V. Comparative mapping of Phytophthora resistance loci in pepper germplasm: evidence for conserved resistance loci across Solanaceae and for a large genetic diversity. *Theor. Appl. Genet*. **106,** 1473-1485 (2003).

Tian, F. *et al*. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* **43,** 159 (2011).

Torii, K. U. Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *Int. Rev. Cytol.* **234,** 1-46 (2004).

Truong, H. *et al*. Identification of isolate-specific resistance QTLs to phytophthora root rot using an intraspecific recombinant inbred line population of pepper (*Capsicum annuum*). *Plant Pathol.* **61,** 48-56 (2012).

Vleeshouwers. V. G. A. A. *et al*. Understanding and exploiting late blight resistance in the age of effectors. *Annu. Rev. Phytopathol*. **49,** 507-31 (2011).

Wang, P. Wang, L. Guo, J. Yang, W. & Shen, H. Molecular mapping of a gene conferring resistance to Phytophthora capsici Leonian race 2 in pepper line PI201234 (*Capsicum annuum L*.). *Mol. Breeding* **36,** 66 (2016).

Wang, S., Basten, C. & Zeng, Z. Windows QTL cartographer 2.5. Raleigh, NC: Department of Statistics, North Carolina State University, USA. (2011).

Witek, K., Jupe, F., Witek, A. I., Baker, D., Clark, M. D. & Jones, J. D. Accelerated cloning of a potato late blight–resistance gene using RenSeq and SMRT sequencing. *Nat. Biotechnol*. **34,** 656 (2016).

Xu, X. *et al*. Mapping of a novel race specific resistance gene to phytophthora root rot of pepper (*Capsicum annuum*) using bulked segregant analysis combined with specific length amplified fragment sequencing strategy. *PloS ONE* **11,** e0151401 (2016).

Yang, H. B. *et al*. Development and validation of L allele-specific markers in *Capsicum*. *Mol. Breeding* **30,** 819-829 (2012).

Yang, L. *et al*. A new resistance gene against potato late blight originating from *Solanum pinnatisectum* located on potato chromosome 7. *Front. Plant Sci*. **8,** 1729 (2017).

Yu, F., Zhang, X., Peng, G., Falk, K. C., Strelkov, S. E. & Gossen, B. D. Genotyping-by-sequencing reveals three QTL for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. *Sci. Rep.* **7,** 4516 (2017).

Zhang, C. *et al*. The *Ph-3* gene from *Solanum pimpinellifolium* encodes CC-NBS-LRR protein conferring resistance to *Phytophthora infestans*. *Theor. Appl. Genet.* **127,** 1353-1364 (2014).

Zheng, Z. *et al*. Loss of function in Mlo orthologs reduces susceptibility of pepper and tomato to powdery mildew disease caused by *Leveillula taurica*. *PloS ONE* **8,** e70723 (2013).

CHAPTER II

Comparative Genomic Analysis of Three Isolates of *Phytophthora capsici*

ABSTRACT

Phytophthora capsici is an oomycete pathogen responsible for damping off, root rot, fruit rot and foliar blight of popular vegetable and legume families, which causes yield and quality damage resulting in huge economical losses. The reference genome of *P. capsici* partially inbred isolate, LT1534 was previously sequenced using 454 FLX and Titanium technology supplemented by Sanger reads (5X). In the present study, I report the comparative genome assemblies of three *P. capsici* isolates showing a striking variation in pathogen virulence. The sequencing of three *P. capsici* genomes were performed using Illumina HiSeq and single-molecule real-time (SMRT) sequencing technologies. Genome sequencing and characterization was performed of the three *P. capsici* isolates representing distinct virulence profiles. An average number of 514 contigs with 50.96% of GC contents, 698,937 bp of N50 and 16,398 predicted genes were obtained. Genomic analysis of three *P. capsici* isolates revealed a large number of genes encoding potential secreted effectors including an average of 60 RxLR domain containing effectors, 42 Crinklers (CRN), 536 CAZymes, and several apoplastic effectors, such as cytochrome P450, phytotoxins (PcF proteins), NPP1 families, LRR kinase as well as virulence and necrosis inducing proteins. The comparative genomic analysis and GO term enrichment analysis revealed the numerous groups of genes which. In addition, the characterization of the virulence profiles of the isolates in laboratory and field experiments were assessed which showed a striking variation in virulence of three isolates. This study provides a genomic landscape of three *P. capsici* isolates for the comparative genomic analysis, which may help to identify the genes under positive and negative selections, homology analysis with other *Phytophtora* spp. and effector assisted breeding. This characterization and genome information elucidates virulence and genome structural variations for *P. capsici* isolates, offering opportunities to better understand the virulence profiles of this

oomycete pathogen, and presents possibilities for more detailed studies on genomics and virulence that may lead to future help in resistance incorporation and practical breeding applications.

INTRODUCTION

The oomycetes are a varied class of eukaryotic microbes that comprise pathogens of animals, plants and fungi (Kamoun, 2003). The phylogenetic analysis well represents the oomycetes as plant disease casual agents, with above 60% of well-known oomycetes delineated as plant pathogens (Thines and Kamoun, 2010). Among them, *Phytophthora* genus is accountable for some of the most destructive as well as culturally and economically imperative diseases, including late blight of potato by *P. infestans*, soybean stem rot caused by *P. sojae*, Sudden Oak Death by *P. ramorum* and crown rot of strawberry by *P. cactorum* (Kamoun et al. 2015).

The hemibiotrophic oomycete pathogen *P. capsici* was detected as the causal agent of pepper foliar and root rot disease in 1922, and is now considered a major pathogen of pepper in sub-tropical and temperate regions, leading to yield losses of up to 40% (Hausbeck and Lamour, 2004). *P. capsici* is homothallic and heterothallic oomycete producing sporangia and oospores (resting spores) in the infected plant tissue. These can persist for several years in soil and serve as a central source of contagion in open field production systems (Foster and Hausbeck, 2010). *P. capsici* is also a limiting factor in the propagation of pepper plants, risking prompt spread of the disease infection upon outbreak in the protected cultivation. Chemical control through soil fumigation with fungicides such as metalaxyl, mefenoxam and methyl bromide have been proved ineffective to control the pathogen (Castro-Rocha et al. 2014). This has intensified the significance of integrating disease resistance into current breeding programs. The durability of plant disease resistance is influenced by many factors including environmental cues, host resistance mechanism, pathogen-encoded effector proteins that can reprogram the plant cellular processes in order to overcome the plant defense system and facilitate infection. The sequenced genome of *Phytophthora* spp. pathogens and their functional

depiction of putative effector candidates from predicted gene models provide an opportunity and framework to study the *Phytophthora* infection mechanisms (Thines and Kamoun, 2010).

Genome sequences of the hemibiotrophic oomycetes *P. ramorum* (Tyler et al. 2006), *P. sojae* (Tyler et al. 2006), *P. infestans* (Haas et al. 2009), *P. capsici* (Lamour et al. 2012), and *P. litchii* (Ye et al. 2016) have been completed and analyzed. In parallel, the genomes of plant pathogenic species in other oomycete genera, for example, the obligate biotrophic *Hyaloperonospora arabidopsidis* (Baxter et al. 2010) and *Albugo* spp. (Kemen et al. 2011; Links et al. 2011), and the necrotrophic *Pythium* spp. (Levesque et al. 2010; Adhikari et al. 2013) have uncovered the prominent variations in genome size and composition, a set of pathogenesis-related genes, and adaptation associated to trophic (necrotropic and biotrophic) modes, thus contributing to the enhanced understanding of host–pathogen interactions (Jiang and Tyler, 2012; Judelson, 2012). Comparative genome analysis of different organisms is ideal to detect and classify the genes influenced by evolutionary selection. For instance, genes with increased or reduced substitution rates, which could be the consequence of negative selection (purifying selection) or positive selection (adaptive selection) could be detected (Petersen et al. 2007; Ye et al. 2016). For example, the genes classified as effector genes provide clues for adaptive fluctuations in functions. They might facilitate infection (toxins and virulence factors), activate defense responses (elicitors and avirulence factors), or both (Hogenhout et al. 2009). As a consequence, effector genes could serve as an important sign of natural selection during co-evolution process of host and pathogen (McCann and Guttman, 2008; Ma and Guttman, 2008; Stukenbrock et al. 2011; Rech et al. 2014).

Several effector coding genes evolve at faster rates compared to the primary genome of pathogens. These genes frequently exhibit great levels of positive selection, with remarkably larger rates of non-synonymous to synonymous nucleotide substitutions (Ka/Ks or dN/dS ratios

greater than 1) (Win and Kamoun, 2008; Wang et al. 2011; Dong et al. 2012). For example, in *P. sojae*, the sequenced genomes of different isolates have been compared and analyzed. Data analysis revealed the high level of polymorphisms and noticeable evidence for positive selection from several CRN, RxLR and NLP effectors genes (Ye et al. 2016). *P. capsici* genomes has been sequenced using the partially inbred strains (Lamour et al. 2012). However, there has been no studies on employing the long-read whole genome sequencing and comparative genomic analysis of *P. capsici* isolates, to identify the genes associated with structural variations, variations in virulence as well as the genes under positive selections.

The objective of the present study was to compare the genome of three isolates exhibiting different level of virulence to reveal the variation at genome level. The *P. capsici* isolates were categorized in vivo and vitro for the mycelium growth and aggressiveness on host plants. Using the complementary sequencing technologies of llumina HiSeq 4000 system and Pacific Biosciences Sequel system were employed to generate de novo hybrid assemblies of *P. capsici* isolates.

MATERIALS AND METHODS

Pathogen isolates and DNA extraction

P. capsici isolates MY-1, JHAI1-7, and KPC-7 (Table I-1) used in the current study were kindly provided by Dr. Choi (Korea Research Institute of Chemical and Technology, Korea). MY-1, JHAI1-7, and KPC-7 isolates were reported to have low, medium, and high virulence, respectively (Jo et al. 2014; Liu et al. 2014). *P. capsici* isolates were reisolated from the diseased plant tissues before they were used for virulence studies. The virulence of *P. capsici* isolates was reconfirmed as described below.

P. capsici isolates were cultured in V8-agar medium (V8 juice; 200 g/L, $CaCO_3$; 3g/L agar; 20 g/L). For DNA extraction, small round blocks (1 cm) harvested with cork borer of actively growing cultures were used to inoculate Erlenmeyer flasks (250 ml) having 200 ml of autoclaved potato dextrose broth. The cultures were placed at room temperature at 27 ± 2°C. After 5–10 days, depending on the growth of each isolate, mycelia were harvested by filtration through cheesecloth, blotted dry with sterile paper towels followed by freeze dried in liquid nitrogen and stored in -80ºC. High molecular weight DNA was isolated and purified using modified fungal CTAB methods as described in detail by (Zelaya-Molina et al. 2010). The purity and quality of the isolated DNA samples were evaluated by electrophoresis on 0.8% agarose gel and NanoDrop 2000 Spectrophotometer.

Growth and virulence evaluation in laboratory and field

To evaluate the virulence variation among the isolates, control plants were inoculated. The inoculum was prepared and adjusted as describe in the chapter I. A set of control plants containing 20 replicates for each including susceptible controls *C. annuum* ‘Tean’ and resistant control *C. annuum* ‘Criollo de Morelos 334’ were inoculated with 5 mL of inoculum using the

soil drenching method and the experiment was repeated three times. The disease reactions were observed according to the symptoms appeared on stem collars. In complementary to this, the mycelium growth of *P. capsici* isolates were tested on two different growing mediums at the same temperature and humidity to compare the isolate growth rate in correlation to virulence. The small round blocks (1 cm) of actively growing mycelium were placed on the center of petri plates filled with solid media of V8-agar and potato dextrose agar. The experiment was conducted in complete randomized block design with three replications per treatment. The mycelium length was measured and compared the averages of each plates after every 24 h until the 90 mm diameter plates fully covered with mycelia.

To test the virulence of isolates on foliage of plants, detached leaf assay was performed. Leaves of the susceptible cultivar 'Tean' were harvested at the mature stage and used for inoculation. The leaves were placed on a sterile moisture filter papers inside the petri dishes to maintain the humidity and kept at 27 ± 2°C. The inoculum density was adjusted 40,000 spores/mL and 20 µl of inoculum was dispensed on the leaf center using micro pipette. The sterile distilled water was used as mock control. The experiment was conducted in complete randomized block design with three replications per treatment. The infection severity in the form of mycelium growth on leaf surface was observed after 3 and 6 days post inoculation.

Whole genome sequencing and assembly

Genome sequencing of the three *P. capsici* isolates were performed using both Illumina HiSeq 4000 system and PacBio Sequel system (Macrogen Inc., Seoul, South Korea). Assemblies with 50x coverage of 101 bp reads were generated by Illumina paired-end sequencing (Table II-1). Additionally, 100x Sequel read was obtained with average length of 9 kb which is enough to accomplish *de novo* assembly. For hybrid assembly, MaSuRCA genome

assembly package version 3.2.9 (Zimin et al. 2013) was used with the default parameters. After completing the genome assembly, HaploMerger2 program was used to compute the haploid state from diploid assembled genome (Huang et al. 2017). The assembled genomes were validated for the accuracy and completeness by BUSCO version 3.1.0, with the option of eukaryote lineage dataset (Waterhouse et al. 2017).

Gene prediction and functional annotation

Gene prediction and functional annotation was performed using repeat sequences of assembled genome. RepeatMasker (ver 1.332) was used to perform the annotation of repeat sequences in *de novo* assembled *P. capsici* genome using custom repeat library generated by RepeatModeler (ver. 1.0.11). In next step, the structural genome annotation was performed by using MAKER 2.31.10 (Holt and Yandell, 2011). Functional genome annotation was done in two rounds. At first step, evidence-driven gene prediction was executed by importing the EST and protein sequences of *P. capsici* (Lamour et al. 2012) into MAKER2 pipeline. In next step, annotated files from MAKER2 pipelines were trained using the several *ab-initio* gene prediction programs including, Augustus (Stanke and Morgenstern, 2005), SNAP (Johnson et al. 2008) and GeneMark-ES (Ter-Hovhannisyan et al. 2008) employing the *de novo* assembled genome to annotate gene families and function precisely.

Identification of candidate effectors and orthology analysis

InterProScan (ver5.34-73.0) (Quevillon et al. 2005) was used to analyze the gene function and to obtain functional and Gene Ontology IDs. Genes coding for Carbohydrate-active enzymes (CAZymes) were predicted through scanning using HMMER 3.0, DIAMOND, and Hotpep in web-based tool, dbCAN2 (http://bcb.unl.edu/dbCAN2). Genes coding for

CAZymes protein were considered for further analysis after confirmation by two programs and included as a final set of CAZymes.

Different methods were employed to predict virulence-associated factors in *de novo* assembled genome. RxLR effectors were predicted by IPR domain number from InterProScan results. CRN effectors prediction was done by BLASTp (E-value cutoff < $10e^{-5}$). Protein sequences from de novo assembled sequence were used as a query to search *Phytophthora* CRN effectors database (reference). Most of the other proteins were predicted by domain search using HMMER3.0 against Pfam domain database.

Orthology analyses were performed using orthoMCL program (ver.2.0.9) following the default options for each *P. capsici* isolate including previously reported reference genome (LT1534) and Venn diagram was plotted using online web tools.

Genomic comparison for intergenic polymorphism

To find factors responsible for the virulence discrepancy among three *P. capsici* isolates, resequencing was performed using Illumina reads only. Read alignment was carried out by Burrows-Wheeler Aligner (BWA) program (ver. 0.7.17) against published *P. capsici* genome. Alignment file was processed to sort alignment information, duplicate removal, and grouping read using Picard tools (ver. 2.18.16). Variants calling for the three isolates was done using GATK (ver. 3.8) and snpEff (ver. 4.11) to annotate SNPs into gene function. After variant annotation, GO (Gene Ontology) enrichment analysis was performed by gene sets which have more than 10 variants in one gene among three isolates. AgriGO toolkit was used for GO enrichment test and ReviGO was used to visualize GO enrichment test result.

RESULTS

Characterization of *P. capsici* isolates

Three *P. capsici* isolates showed a significant difference in terms of symptom development and virulence. For the highly virulent isolate KPC-7, disease symptoms including wilting and the appearance of water-soaked lesions at stem collars were observed within 72 h of inoculation in the susceptible control ‘Tean’ and the plants were completely wilted and died at 7 to 10 days post inoculation (DPI) (Fig. II-1). By contrast, susceptible plants infected with the moderately virulent isolate JHAI1-7, disease symptoms were at 5 to 7 DPI, and the plants were completely wilted and died at two weeks post inoculation (Fig. II-1). In case of the isolate MY-1, with low virulence, symptoms were observed on the susceptible plants at 10 to 14 DPI, but complete plant death was not detected up to three weeks post inoculation (Fig. II-1). The resistant control CM334 infected these isolates remained symptomless even after three weeks post inoculation (Fig. II-1).

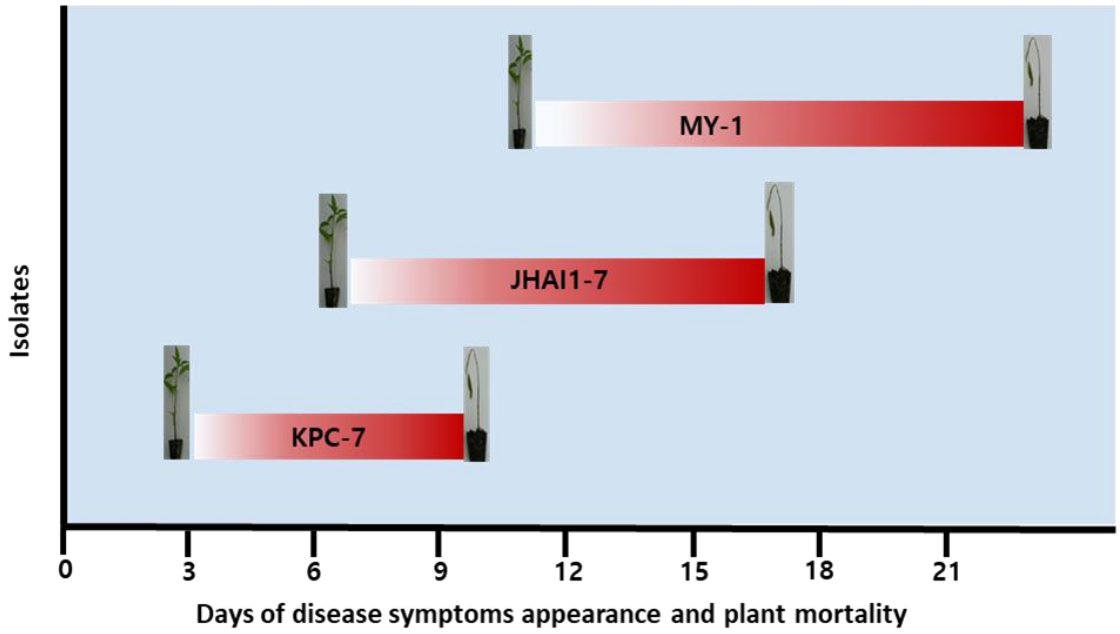


Figure II- 1. Disease symptoms appearance and plant mortality of 'Tean' inoculation with three different *P. capsici* isolates KPC-7, JHAI-7 and MY-1. The observations were made from 0-21 days.

In order to examine the correlation of *P. capsici* isolates virulence with mycelium growth in different growth media, isolates were cultured in V-8 agar and potato dextrose agar media and mycelial growth were observed and compared. Mycelial growth measurement data revealed isolates followed the similar trends of virulence as disease appearance on host plants in case of low and high virulence isolates. After 48 hours of culturing, isolates KPC-7, JHAI1-7 and MY-1 were grown up to 36.7, 41.3 and 22.2 mm on V8-agar medium (Fig. II-2a, b, c and g). The final growth was measured after 120 hours of culturing which showed the 85.5 mm in mycelium length for the KPC-7 (high virulent) and JHAI1-7 (medium virulent), whereas 64.3 mm for low virulent isolate MY-1 (Fig. II-2a, b, c and g). The mycelium growths on potato dextrose agar also displayed the similar trend of growth with a little exception. The measurement results showed, 24.7 and 25.3 mm length of mycelia for KPC-7 and JHAI-7 whereas 14.2 mm for MY-1 after 48 hours of culture (Fig. II-2d, e, g and h). The mycelium was spread throughout the petri dishes in 90 mm of diameter after 144 hours of culture for KPC-7 and JHAI1-7 while 63 mm length wad recorded for low virulence isolate MY-1 (Fig. II-2d, e, g and h). The above results indicate the correlation of virulence with the mycelium growth of isolates on growing medias, where isolates with high virulence showed fast mycelial growth compared with isolates having low virulence.

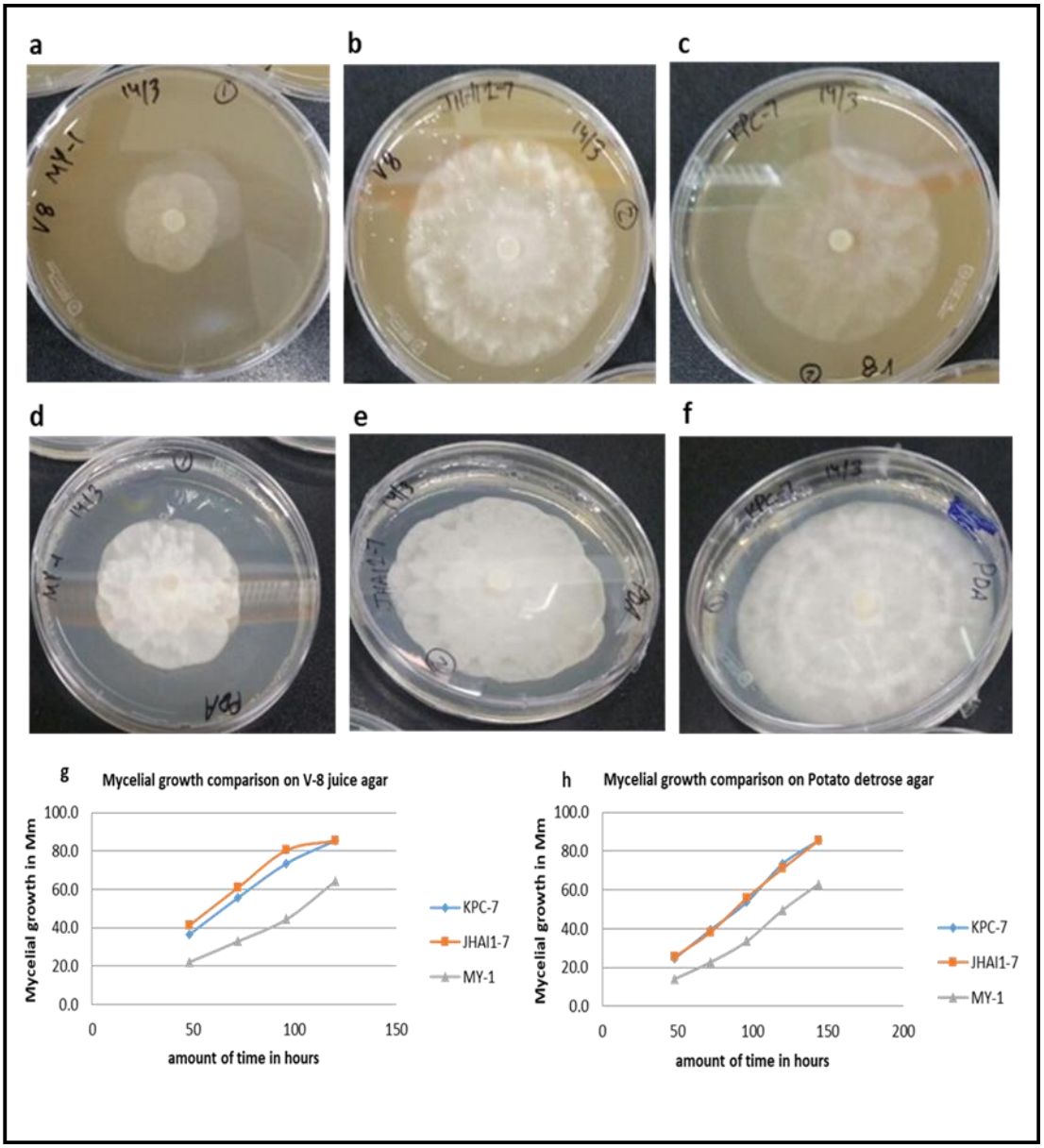


Figure II- 2. Mycelium growth of three isolates on two different grown media. (a, b, c) Isolates growth on V8-agar media. (d, e, f) Isolates growth on potato dextrose agar media. (g and h) the growth comparison of three different virulent isolates on two me.

In another study, infection on leaf surface was observed to examine the virulence of *P. capsici* isolates. The detached leaf assay results revealed that the isolates followed the similar trend of pathogen virulence as that of root rot infection and mycelia growth on different media. The leaves inoculated with low virulent isolate MY-1 showed the lowest infection (yellowing and lesion formation) after 3 and 6 days of post inoculation (Fig. II-3). The lesion diameter was observed less than the half of the leaf width (Fig. II-3). The medium virulent isolate JHAI1-7 showed the intermediate level of infection on leaf surface after 3 days of inoculation but the 6 days of post infection the whole leaves surfaces were covered with mycelial infection (Fig. II-3). The KPC-7 showed high virulence and the whole leaf infection was observed after 3 days of inoculation. The highest and severe infection was observed at the 6 DPI in KPC-7 (Fig. II-3). These results confirmed the low, medium and high virulence of MY-1, JHAI1-7, and KPC-7 isolates on plant leaves infection.

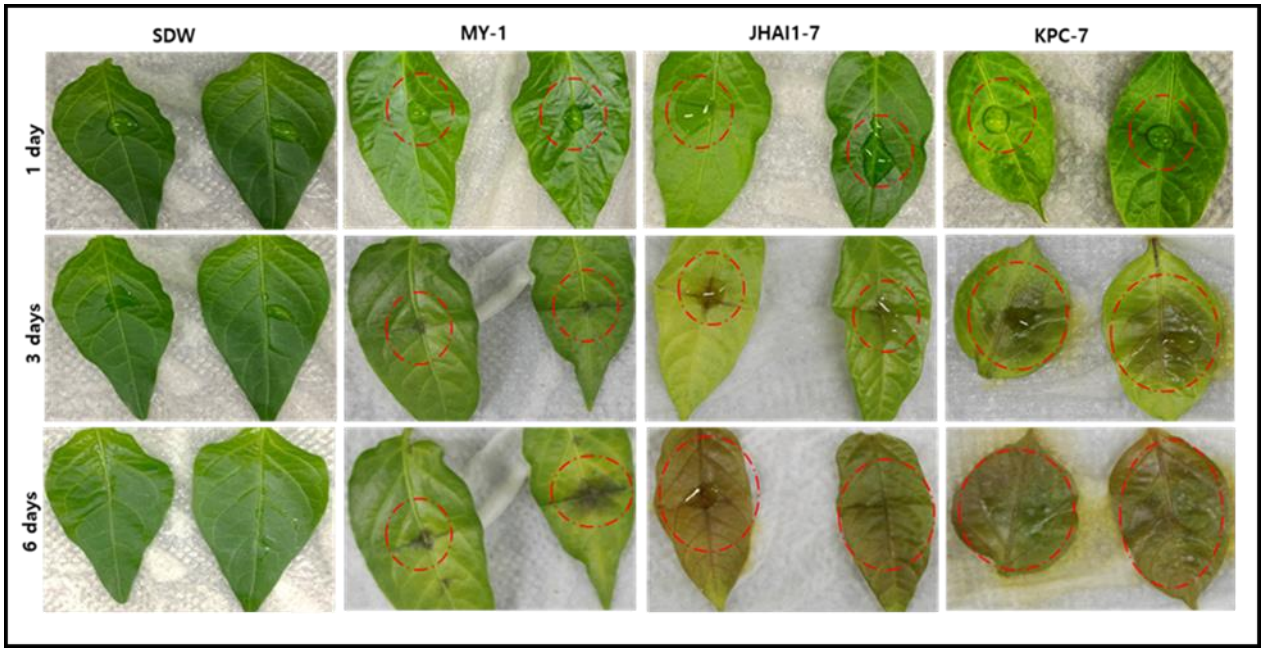


Figure II- 3. Leaf inoculation with three *P. capsici* isolates for virulence testing. Sterile distilled water was used as mock control.

Pathogen genome assembly

Assemblies with 50x coverage of 101 bp reads were generated by Illumina paired-end sequencing. Additionally, 100x Pacbio Sequel read was obtained with average length 9 kb enough to accomplish *de novo* assembly (Table II-1). *De novo* hybrid genome assembly using MaSuRCA genome assembly package (version 3.2.9) generated 74.9 Mb, 76.8 Mb and 76.6 Mb assemblies for KPC-7, JHAI1-7, and MY-1, respectively (Table II-2). The genome assembly process resulted in 521,472 and 549 number of scaffolds with the GC contents of 50.9%, 51.0%, and 50.9%, while N50 of 6.63 Mb, 7.98 Mb, and 6.35 Mb for KPC-7, JHAI1-7 and MY-1, respectively (Table II- 2). The longest and shortest scaffold length of 2.7 Mb and 657 bp for KPC-7, 2.5 Mb and 1,229 bp for JHAI-7 and 2.7 Mb and 723 bp for MY-1 were recorded (Table II- 2). The number of total genes were predicted to be 15,910, 16,261 and 17,024 for KPC-7, JHAI1-7, and MY-1, respectively (Table II-2). BUSCO completeness analysis revealed above 90% accuracy for the assembled genomes (Table II-2). Precision of the genome assemblies were further tested by the comparing with reference genome. The dot plot analysis showed the high accuracy and collinearity of genome assemblies with the reference genome (Figure II- 4).

Table II- 1. Summary of reads generated in present study

Sequencing Approach	MY-1	JHA1-7	KPC-7
No. of total HiSeq reads	28,853,066	28,240,964	27,025,390
No. of total Sequel reads	1,356,081	1,248,011	1,351,163

Table II-2. Summary of *P. capsici* isolates genome assemblies

	MY-1	**JHA1-7**	**KPC-7**	**LT1534 v11.0**
# of scaffolds	521	472	549	917
GC content (%)	50.96	51.01	50.91	44.14
Total size (bp)	76,624,583	76,839,304	74,955,867	64,023,748
Longest scaffold (bp)	2,720,928	2,514,355	2,715,842	2,170,955
Shortest scaffold length (bp)	723	1,229	657	1,001
Number of N's	1,409 (0.002%)	1,067 (0.001%)	1,504 (0.002%)	7,981,741 (12.467%)
Number of gaps	23	28	19	9,840
N50 of scaffolds (bp)	635,553	798,104	663,153	705,730
Number of genes	17,024	16,261	15,910	19,805
BUSCO completeness	91.40%	93.10%	93.70%	-
Remarks	This study	This study	This study	Lamour et al. 2012

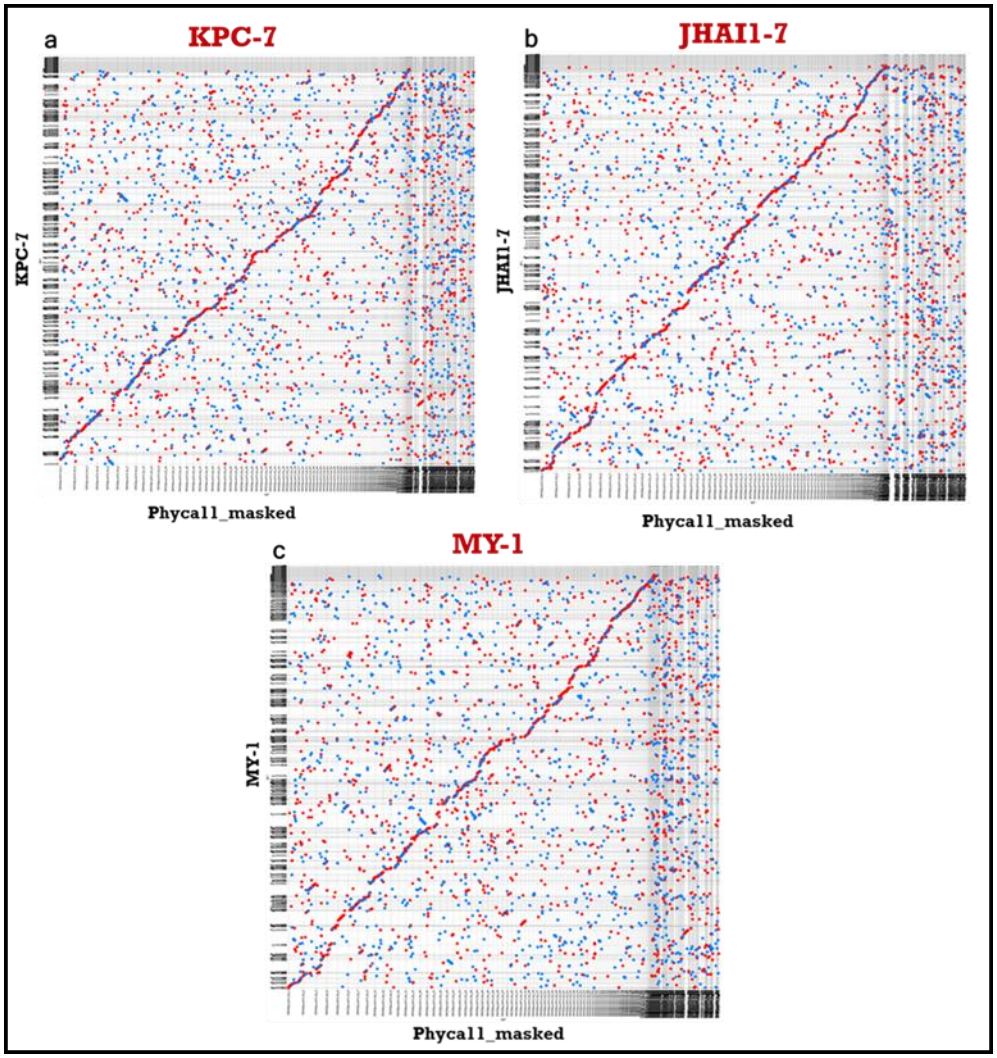


Figure II- 5. Dot plot analysis between published *P. capsici* genome and *de novo* assembled genomes in this study. X-axis represents *P. capsici* strain LT1534 genome version 11.0. Y-axis represents isolates, KPC-7, JHAI1-7 and MY-1.

Orthology analysis and functional annotation

Clustering of predicted proteins from the three *P. capsici* isolates in this study and the previously reported isolate using orthoMCL program revealed 13,308 orthogroups containing 60,612 proteins (15,557 for LT1534; 14,993 for KPC-7; 15,257 for JHA11-7; 14,805 for MY-1) (Fig. II-5). Several unique protein clusters specific to each isolate identified (Fig. II-5). In detail, the previously sequenced isolate LT1534 has 4,248 unique clusters, whereas KPC-7, JHAI1-7, and MY-1 isolates showed 917, 1,004 and 2,219 clusters, respectively (Fig. II-5). A total of 13,308 orthologous groups were commonly existed among four isolates (Fig. II-5).

Carbohydrate active enzymes (CAZymes) show a direct role in pathogenicity and virulence instigating to plant cell wall degradation. A total of 480, 573 and 557 transcripts encoding CAZymes were detected in assembled genomes of KPC-7, JHAI1-7 and MY-1 isolates, respectively (table II-3). These secreted CAZyemes include glycosyl hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), auxiliary activities (AAs), carbohydrate-binding modules (CBMs), carbohydrate esterases (CEs), and CAZymes including signal peptides families (Table II-3). The isolate MY-1 detected with 232 GHs, 136 GTs, 39 PLs, 34 AAs, 9 CBMs, 30 CEs and 152 signal peptides families (Table II-3). The isolate JHAI1-7 detected with 280 GHs, 166 GTs, 31 PLs, 38 AAs, 10 CBMs, 48 CEs and 158 signal peptides families (Table II-3). Whereas, the isolate KPC-7 detected with 264 GHs, 174 GTs, 28 PLs, 37 AAs, 7 CBMs, 47 CEs and 150 signal peptides families (Table II-3).

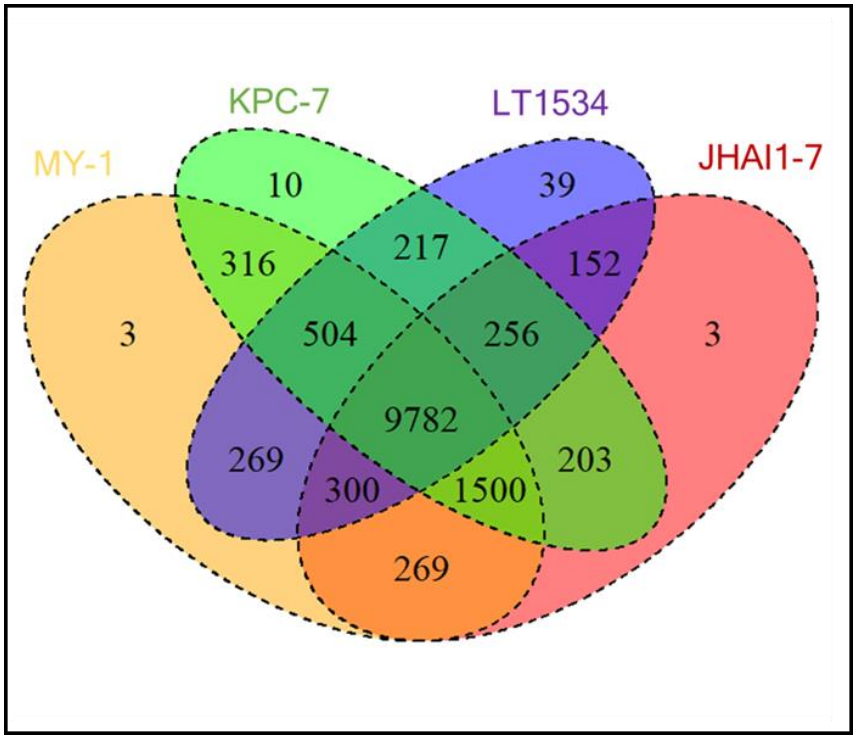


Figure II- 6. Venn diagram of the common and unique families of genes in four *P. capsici* isolates.

Table II- 3. Number of predicted CAZymes in *P. capsici* isolates

Category	MY-1	JHAI1-7	KPC-7
CAZymes annotated	480	573	557
Glycosyl hydrolases (GHs)	232	280	264
Glycosyl transferases (GTs)	136	166	174
Polysaccharide lyases (PLs)	39	31	28
Auxiliary Activities (AAs)	34	38	37
Carbohydrate-binding modules (CBMs)	9	10	7
Carbohydrate esterases (CEs)	30	48	47
CAZymes including signal peptides	152	158	150

Crinkler (CRN) and RxLR effectors considered as major contributor in pathogenicity and virulence determinants of *Phytophthora* species. The functional and structural annotation for RxLR effectors protein identification in assembled genomes revealed 59 RxLR genes for both MY-1 and JHAI1-7 and 60 for the KPC-7 (Table II-4). The identification of CRN effectors proteins showed 42, 37 and 42 genes for MY-1, JHAI1-7 and KPC-7, respectively (Table II-4). Several virulence associated factors have been reported and characterized in *Phytophthora* spp. including Cytochrome P450, NPP1 family, ABC transporters, PcF family and LRR kinase. The identification of virulence related factors in low virulent isolates MY-1 revealed 22 genes for Cytochrome P450, 55 for NPP1 family, 156 for ABC transporters, 1 for PcF family and 14 for LRR kinase (Table II-4). The results for medium virulent isolate JHAI1-7 detected with 33, 2, 158, 5 and 20 coding genes for Cytochrome P450, NPP1 family, ABC transporters, PcF family and LRR kinase, respectively (Table II-4). Whereas the high virulent isolate KPC-7 was identified with 31, 10, 142, 3 and 24 coding genes for Cytochrome P450, NPP1 family, ABC transporters, PcF family and LRR kinase, respectively (Table II-4).

Table II- 4. Composition of cytoplasmic and apoplastic effectors in genomes

Category	MY-1	JHA1-7	KPC-7
Cytoplasmic effectors			
RxLR effector	59	59	60
CRN effector	42	37	42
Virulence associated factors			
cytochrome P450	22	33	31
NPP1 family	55	2	10
ABC transporters	156	158	142
PcF family	1	5	3
LRR kinase	14	20	24

Whole-genome comparison of *P. capsici* isolates

The genome assemblies of three *P. capsici* isolates were evaluated and compared with previously published reference genome for the presence of SNPs in the genes involved in biological processes and molecular functions. *P. capsici* reference genome have shown one of the highest SNPs densities among *Phytophthora* spp. (at least one SNP for every 40 bp). The resequencing and SNP calling of three *P. capsici* isolates against reference genome detected with total 63,373 SNPs with an average including 23,809 missenses, 39,254 synonymous and 310 nonsense variants (Fig. II-6a).

GO term enrichment analysis for the genes involved in biological process revealed three genes families which contain 10 SNPs in each CDS sequence among three isolates including transposition, DNA recombination and lysosome organization (Fig. II-6b and Table II-5). In addition, the three isolates genomes were identified with 10 SNPs in each CDS sequence for the six genes families those are contributed in molecular functions comprising ATPase activity, hydrolase activity, transposase activity, DNA binding, protein dimerization activity and primary active transmembrane transporter activity (Fig. II-6c and Table II-5). These results suggested the SNPs among the important genes may leads to the functioning in the virulence discrepancy of the three isolates.

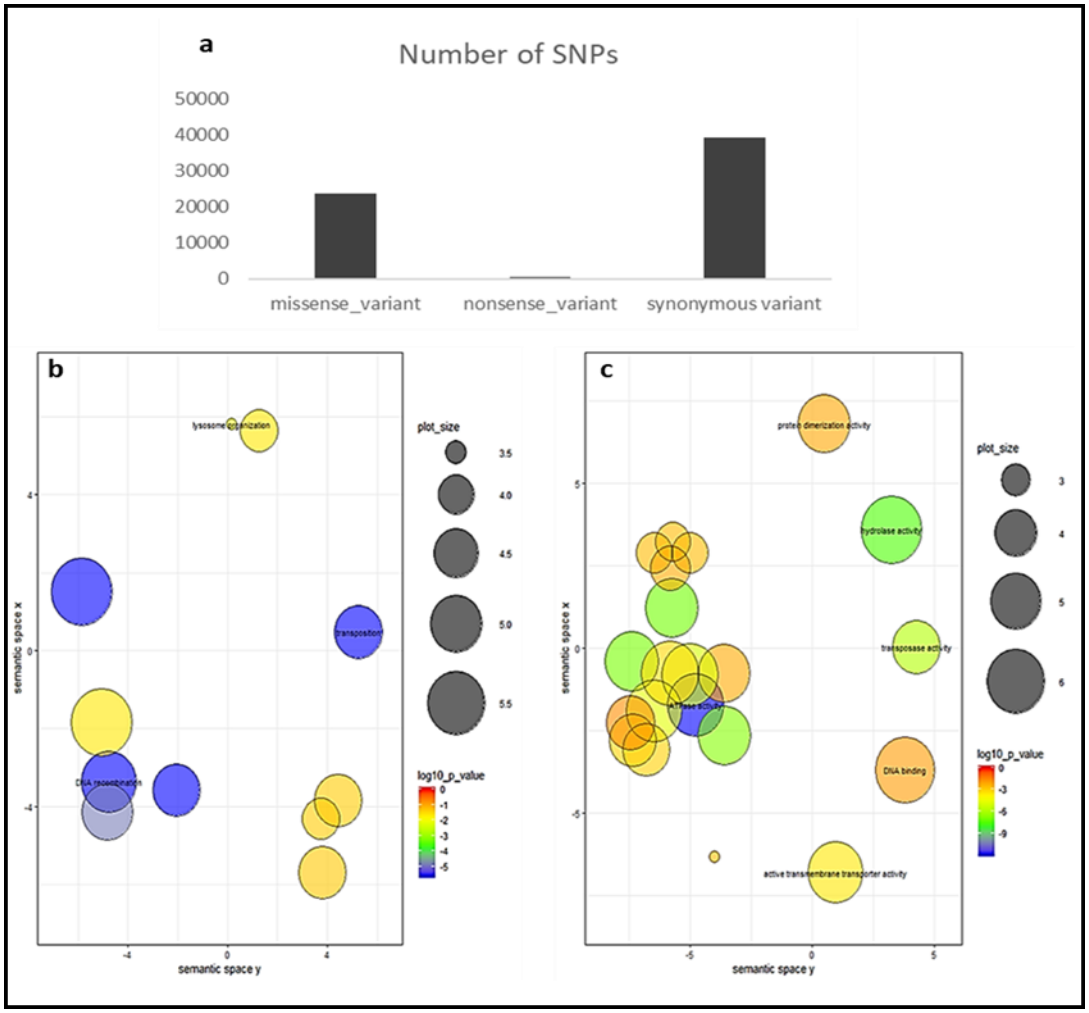


Figure II- 7. Summary of resequencing and GO term analysis results of three *P. capsici* isolates. (a) Number of SNPs in *P. capsici* isolates. (b) GO term enrichment analysis for the genes involved in biological process. (c) GO term enrichment analysis for the genes involved in molecular functions

Table II- 5. GO term enrichment analysis of resequencing data for polymorphism

GO-accession	Term-type	Term	Query-item	Query-total	Ref-item	Ref-total	P value	FDR
GO:0032196	P	transposition	31	2084	48	8543	5.00E-09	2.60E-06
GO:0006310	P	DNA recombination	40	2084	70	8543	6.00E-09	2.60E-06
GO:0007040	P	lysosome organization	10	2084	12	8543	3.00E-05	0.0086
GO:0016887	F	ATPase activity	181	2084	422	8543	4.50E-16	1.30E-12
GO:0016787	F	hydrolase activity	608	2084	1908	8543	2.00E-11	2.90E-08
GO:0015399	F	primary active transmembrane transporter activity	84	2084	186	8543	8.80E-10	2.80E-07
GO:0004803	F	transposase activity	31	2084	48	8543	5.00E-09	1.50E-06
GO:0046983	F	protein dimerization activity	46	2084	112	8543	8.40E-05	0.008
GO:0003677	F	DNA binding	212	2084	689	8543	0.00016	0.014

DISCUSSION

Recent discoveries have considerably expanded the repertoire of molecular determinants known to underlie microbial virulence at the cellular level. The sequencing of the *P. capsici* genome as a model pathosystem composed of a perennial plant and hemibiotrophic oomecete has opened new avenues for the study of comparative genomics (Lamour et al. 2012). I sequenced, assembled and annotated the genomes of three diverse *P. capsici* isolates using *de novo* hybrid assembly approach using Illumina HiSeq and Pacbio sequencedata. Our results revealed an average assembly size of 76 Mb with 16,368 predicted transcripts for three isolates. The assemblies were fragmented in 524 contigs. However, BUSCO statistics indicated highly-complete genome assembly and detection of 93% of completeness. Resequencing and SNP calling of three isolates detected with 10 polymorphisms per CDS in several gene families involved in biological processes and molecular functions. The annotation results also showed dissimilarities in genomes, for example, the expansion of some orthologs and few unique gene families. This study also provided evidence for differences in virulence of *P. capsici*. The in vivo and in vitro experiments showed the distinctive variations in the symptom development and virulence levels among three *P. capsici* isolates. The isolate MY-1 characterized as low virulent, whereas JHAI1-7 and KPC-7 designated as medium and high virulent isolates, respectively.

Phytophthora pathogens utilize a diverse range of secreted apoplastic and cytoplasmic effectors to aid infection. This work characterized the *P. capsici* genomes, identifying both apoplastic and cytoplasmic effector candidates as well as non-effectors that are typical of MAMP elicitors of host defense. This study unveiled the diversity of effectors in the *P. capsici* genomes, showing that those effectors identified during development infection with those that

may be specifically expressed during infection and the transition to necrotrophy. This study identified greater numbers of CRN, elicitins, GH, PL and RxLR candidates than previously identified in the *P. capsici* transcriptome (Lamour et al. 2012). Equal or greater numbers of genes encoding NLPs, protease inhibitors, cutinases and PcF domain-carrying proteins were identified, however some of the candidates were discarded due to possession of a transmembrane domains or a GPI anchor.

Although it is possible to construct a high-quality assembly for pathogen species with large genomes through hierarchical shotgun sequencing, the utilization of a sequencing platform capable of generating long reads that can span regions of complex repeats is likely to be more cost-effective and less labour-intensive/time-consuming. Recent development and applications of long-read sequencing technologies have shown substantial improvement in genome assemblies, primarily by joining contigs and scaffolds and spanning gaps around repetitive regions (Chaisson et al. 2015; Conte et al. 2015). The PacBio sequencing technology offers kilobase-sized reads without GC-bias or systematic errors; however, the raw data generated from this system is inherently error-prone, with an average accuracy of ~82% (Eid et al. 2009). Availability of hybrid assembly software allows biologists to combine PacBio long-read data with complementary, high-accuracy short-read data from other platforms (English et al. 2012). Nowak et al. (2018) demonstrated that the incorporation of additional 8×coverage of PacBio data to the existing short-read-only assembly of the *Primula veris* genome resulted in 21% of the gaps closed and 38% of the ambiguous positions in the gaps filled, in addition to the 40% improvement in N50 contig size (Nowak et al. 2018). Similarly, a set of 16.5×PacBio long reads was used to fill in 68% of the gaps existed in the Illumina-only assembly of the African cichlid (Conte et al. 2015).

Pathogenicity genes are often clustered and rapidly evolving in plant pathogens, less-

conserved, gene-poor genomic regions enriched in repeated elements or transposons (Hass et al. 2009; Raffaele et al. 2010; Rouxe et al. 2011; Bopp et al. 2013; Dutheil et al. 2016;). These regions generally contain effectors involved in host or ecological adaptation (Hatta et al. 2002; de Jonge et al. 2013). Rapid evolution of these regions can create lineage-specific or divergent non-core regions that vary between populations. The *de novo* assembled genome of *P. capsici* displays a distinct evolutionary trajectory compared to the reference genome. It is also enriched in genes associated with plant-pathogen interactions and TE-like sequences, some of which are in tandem with effector genes. It is therefore likely to play an important role in the evolution and adaptive potential of the lineages. There was an extensive loss of effector loci in the three isolates of *P. capsici* likely caused by differences in host selection pressures. Selection pressure on effectors may be especially strong since a mismatch in effectors or host recognition is expected to have a high fitness cost (Giraud et al. 2010). The evidence for positive selection and the evolutionary history of effector gene components provide further evidence of rapid evolution. Field and laboratory virulence tests showed that the KPC-7 and JHAI1-7 isolates are adaptively different from MY-1 and has higher virulence on plant roots and leaves. It is possible that the MY-1 source population coevolved with different primary hosts than the other two isolates, making some effectors obsolete in MY-1.

The expansion of gene families was frequently reported to directly or indirectly involve in virulence of fungal pathogens (Spanu et al. 2010; Morales-Cruz et al. 2015; Schwartze et al. 2014). Several gene families of *P. capsici* had undergone expansion and contraction. The number of expanded gene families in *P. capsici* was the largest among *Phytophthora* species, whereas the number of contracted gene families was relatively small. The expanded genes in *P. capsici* were enriched in membrane, cytoskeleton, transport, carbohydrate metabolism, nucleotide binding, transporter activity, transferase activity, and

hydrolase activity. The contracted genes in *P. capsici* were enriched in cellular catabolism, DNA metabolism, chromosome, nucleic acid binding, and nucleotide binding. The expansion of gene families is likely the result of the transposons elements (Schwartze et al. 2014). The expansion of gene families in *P. capsici* may be due to the large number of transposable elements in genome. However, more analysis should be performed to elucidate the relationship of transposons and gene family expansion.

Comparative genomics is a powerful tool to identify virulence-related candidates, as the present study describes. From the perspective of population biology, it is well-known that structure and demography can affect all loci equally. To identify loci under selection, a population genomics approach is required to take into account demographic history. A genomics study is ongoing to identify loci related to virulence. Several mechanisms can underlie genome evolution and a better knowledge of the structural rearrangements occurring in the *P. capsici* genome will help to determine their impact on virulence evolution. Together, these genomic analyses will accelerate functional studies by pinpointing numerous sites of sequence variation that may have key implications at the structural level for the function of effectors and other virulence genes.

REFERENCES

Adhikari, B. N., Hamilton, J. P., Zerillo, M. M., Tisserat, N., Levesque, C. A., & Buell, C. R. Comparative genomics reveals insight into virulence strategies of plant pathogenic oomycetes. *PLoS ONE* **8,** e75072 (2013).

Baxter, L. *et al.* Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science* **330,** 1549-1551(2010).

Bopp, S. E. R. *et al.* Mitotic evolution of *Plasmodium falciparum* shows a stable core genome but recombination in antigen families. *PLoS Genet* **9,** e1003293 (2013).

Castro-Rocha A. *et al.* Traditional and molecular studies of the plant pathogen *Phytophthora capsici*: A Review. *J Plant Pathol Microb* **5,** 245 (2014).

Chaisson, M. J., Wilson, R. K. & Eichler, E. E. Genetic variation and the de novo assembly of human genomes. *Nat Rev Genet* **16,** 627-640 (2015).

Conte, M. A. & Kocher, T. D. An improved genome reference for the African cichlid, *Metriaclima zebra*. *BMC Genom.* **16,** 864 (2015).

de Jonge, R. *et al*. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Res* **23,** 1271-1282. (2013).

Dong, S. *et al.* The NLP toxin family in *Phytophthora sojae* includes rapidly evolving groups that lack necrosis-inducing activity. *Mol. Plant Microbe Interact*. **25,** 896-909 (2012).

Dutheil, J. Y. *et al.* A tale of genome compartmentalization: the evolution of virulence clusters in smut fungi. *Genome Biol Evol* **8,** 681-704 (2016).

Eid, J. *et al.* Real-time DNA tequencing from single polymerase molecules. *Science* **323,** 133-138 (2009).

English, A. C. *et al.* Mind the gap: upgrading genomes with Pacifc Biosciences RS long-read sequencing technology. *PLoS ONE* **7,** e47768 (2012).

Foster, J. & Hausbeck, M. Resistance of pepper to Phytophthora crown, root, and fruit rot is affected by isolate virulence. *Plant Dis.* **94,** 24-30 (2010).

Giraud, T., Gladieux, P. & Gavrilets S. Linking the emergence of fungal plant diseases with ecological speciation. *Trends Ecol Evol* **25,** 387-395 (2010).

Haas, B. J. *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature **461,** 393-398 (2009).

Hatta, R. *et al.* A conditionally dispensable chromosome controls host-specific pathogenicity

in the fungal plant pathogen *Alternaria alternata*. *Genetics* **161,** 59-70 (2002).

Hausbeck, M. K. & K. H. Lamour. *Phytophthora capsici* on vegetable crops: Research progress and management challenges. *Plant Dis.* **88,** 1292-1303 (2004).

Hogenhout, S. A., Van Der Hoorn, R. A., Terauchi, R., & Kamoun, S. Emerging concepts in effector biology of plant-associated organisms. *Mol. Plant Microbe Interact*. **22,** 115-122 (2009).

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

Huang, S., Kang, M. & Xu, A. HaploMerger2: rebuilding both haploid sub-assemblies from high-heterozygosity diploid genome assembly. *Bioinformatics* **33,** 2577-2579 (2017).

Jiang, R. H., & Tyler, B. M. Mechanisms and evolution of virulence in oomycetes. *Annu. Rev. Phytopathol.* **50,** 295-318 (2012).

Jo, S-J. *et al.* Resistance of chili pepper cultivars to isolates of *Phytophthora capsici*. *Korean J. Hortic. Sci. Technol.* **32,** 66-76 (2014).

Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J. & de Bakker, P.I. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24,** 2938-2939 (2008).

Judelson, H. S. Dynamics and innovations within oomycete genomes: insights into biology, pathology, and evolution. *Eukaryot. Cell* **11,** 1304-1312 (2012).

Kamoun, S. Molecular genetics of pathogenic oomycetes. *Eukaryot. Cell* **2,** 191-199 (2003).

Kamoun, S. *et al.* The Top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol* **16,** 413-434 (2015).

Kemen, E. *et al.* Gene gain and loss during evolution of obligate parasitism in the white rust pathogen of *Arabidopsis thaliana*. *PLoS Biol.* **9,** e1001094 (2011).

Lamour, K. *et al.* Genome sequencing and mapping reveal loss of heterozygosity as a mechanism for rapid adaptation in the vegetable pathogen *Phytophthora capsici*. *Mol. Plant Microbe Interact*. **25,** 1350-1360 (2012).

Levesque, C. *et al.* Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire. *Genome Biol.* **11,** R73 (2010).

Links, M. G. *et al.* De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. *BMC Genomics* **12,** 503 (2011).

Liu, W. Y. *et al.* Combined use of bulked segregant analysis and microarrays reveals SNP

markers pinpointing a major QTL for resistance to Phytophthora capsici in pepper. Theor. Appl. Genet. **127,** 2503-2513 (2014).

Ma, W. & Guttman, D. S. Evolution of prokaryotic and eukaryotic virulence effectors. *Curr. Opin. Plant Biol.* **11,** 412-419 (2008).

McCann, H. C. & Guttman, D. S. Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytol.* **177,** 33-47 (2008).

Morales-Cruz, A. *et al.* Distinctive expansion of gene families associated with plant cell wall degradation, secondary metabolism, and nutrient uptake in the genomes of grapevine trunk pathogens. *BMC Genomics* **16,** 469 (2015).

Nowak, M. D. *et al.* Te draf genome of *Primula veris* yields insights into the molecular basis of heterostyly. Genome Biol **16,** 12 (2015).

Petersen, L., Bollback, J. P., Dimmic, M., Hubisz, M., & Nielsen, R. Genes under positive selection in *Escherichia coli*. *Genome Res.* **17,** 1336-1343 (2007).

Quevillon, E. *et al.* InterProScan: protein domains identifier. *Nucleic Acids Res.* **33,** 116-120 (2005).

Raffaele, S. *et al.* Genome evolution following host jumps in the Irish potato famine pathogen lineage. *Science* **330,** 1540-1543 (2010).

Rech, G. E., Sanz-Martin, J. M., Anisimova, M., Sukno, S. A., & Thon, M. R. Natural selection on coding and noncoding DNA sequences is associated with virulence genes in a plant pathogenic fungus. *Genome Biol. Evol*. **6,** 2368-2379 (2014).

Rouxel, T. *et al.* Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-induced point mutations. *Nat Commun* **2,** 202 (2011).

Schwartze, V. U. *et al.* Gene expansion shapes genome architecture in the human pathogen *Lichtheimia corymbifera*: an evolutionary genomics analysis in the ancient terrestrial mucorales (Mucoromycotina). *PLoS genet.* **10,** (2014).

Spanu, P. D. *et al.* Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330,** 1543-1546 (2010).

Stanke, M. & Morgenstern, B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Res.* **33,** (2005).

Stukenbrock, E. H. *et al.* The making of a new pathogen: insights from comparative population genomics of the domesticated wheat pathogen *Mycosphaerella graminicola* and its wild sister species. *Genome Res.* **21,** 2157-2166 (2011).

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O. & Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* **18,** 1979-1990 (2008).

Thines, M., & Kamoun, S. Oomycete-plant coevolution: recent advances and future prospects. *Curr Opin Plant Biol* **13,** 427-433 (2010).

Tyler, B. M. et al. Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313,** 1261-1266 (2006).

Wang, Q. *et al.* Transcriptional programming and functional interactions within the *Phytophthora sojae* RXLR effector repertoire. *Plant Cell* **23,** 2064-2086 (2011).

Waterhouse, R. M. *et al.* BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol.* **35,** 543-548 (2017).

Win, J. & Kamoun, S. Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes. *Plant Signal. Behav.* **3,** 251-253 (2008).

Ye, W., Wang, Y., Tyler, B.M. & Wang, Y. Comparative genomic analysis among four representative isolates of *Phytophthora sojae* reveals genes under evolutionary selection. *Front. Microbiol.* **7,** 1547 (2016).

Ye, W. *et al.* Sequencing of the litchi downy blight pathogen reveals it is a *Phytophthora* species with downy mildew-like characteristics. *Mol. Plant Microbe Interact.* **29,** 573-583 (2016).

Zelaya-Molina, L. X., Ortega, M. A. & Dorrance, A. E. Easy and efficient protocol for oomycete DNA extraction suitable for population genetic analysis. *Biotechnol Lett*. **33,** 715-720. (2010).

Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L. & Yorke, J. A. The MaSuRCA genome assembler. *Bioinformatics*. **29,** 2669-2677 (2013).

ABSTRACT IN KOREAN

*Phytophthora capsici*는 고추 뿌리썩음병을 유발하는 난균류의 병원균이다. *P. capsici* 에 대한 표준유전체, 고밀도 유전자 지도 및 단일염기다형성 분자표지 등이 최근에 개발되기 시작하였다. *Capsicum annuum* cv. CM334는 저항성 유전자원으로써 널리 알려져 있다. 저항성 개체인 CM334와 이병성 개체인 ECW30R을 교배한 188개의 recombinant inbred line (RIL) 집단을 이용하여 고밀도의 연관 지도를 작성하였다. 이 집단을 이용해 genotyping-by-sequencing (GBS)를 수행함으로써 13,021개의 단일염기다형성 분자표지를 얻을 수 있었고, 이 분자표지로 두 개의 다른 환경에서 *P. capsici*의 다양한 병원성을 갖는 균주에 대한 저항성 유전자 지도를 작성할 수 있었다. *P. capsici*의 뿌리썩음병에 대한 양적형질 유전분석을 수행한 결과, 다양한 역병 균주에 대해 넓은 범위의 저항성을 보이는 세 개의 주동유전자좌 (*QTL5.1*, *5.2*, *5.3*) 가 염색체 5번에 위치하였다. 또한 미동유전자좌로써 염색체 4번, 7번, 8번, 11번이 특정 균주 및 환경에서만 확인되었다. 양적형질 유전분석을 보충하기 위하여 352개의 고추 핵심집단에서 고추역병 저항성에 대한 전장유전체 연관분석을 수행하였다. 두 개의 GBS library를 통하여 총 507,713 개의 단일염기다형성이 확인되었으며, 이 단일염기다형성이 전장유전체 연관분석에 이용되었다. 전장유전체 연관분석을 이용하여 고추 역병 저항성과 연관된 117개의 유의미한 단일염기다형성을 탐색할 수 있었고, 그 중 염색체 5번, 7번, 11번은 이전 연구에서 확인된 QTL과 같은 위치에 위치하는 것을 확인할 수 있었다. 양적형질 유전분석과 전장유전체 연관분석을 통하여 병 저항성 기작과 연관된 유전

자좌에 nucleotide-binding site leucine-rich repeat (NBS-LRR)과 receptor-like kinases (RLKs) 단백질과 같은 후보 유전자들이 유전자군을 형성하고 있는 것을 확인하였다. 이 유전자군과 연관된 유의미한 단일염기다형성 마커들은 균주 특이적 혹은 비특이적으로 고추 역병균 저항성 품종을 육성하는데 중요한 역할을 수행할 수 있을 것이다.

두번째 주제에서는 각각 다른 병원성을 가지는 세 개의 *P. capsici* 균주의 유전체 서열을 분석하고 확인하였다. 세 균주를 Illumina HiSeq과 single-molecule real-time (SMRT) 염기서열 분석방법을 이용한 *de novo* hybrid assembly를 수행하였다. 평균적으로 514개의 contig, 50.96%의 GC contents, 698.937 bp의 N50, 16,398개의 유전자가 예측되었다. 유전체 분석은 유전체 서열 내에 많은 수의 effector 단백질을 예측하였는데, 세 개의 *P. capsici* 균주에서 60개의 RxLR effector, 42개의 Crinklers (CRN) effector, 536개의 CAZymes, cytochrome P450, 독소 단백질과 같은 아포플라스트 관련 effector, NPP1 family, LRR kinase와 같은 병원성에 관련된 단백질 등을 예측하였다. 비교 유전체 분석 및 다형성을 이용한 GO term enrichment 분석을 통해서 하나의 CDS 서열 내에서 10개 이상의 다형성을 갖는 다수의 유전자를 확인하였다. 뿐만 아니라, 각 균주의 실험실 및 노지에서의 병원성 검정을 통해서 세 균주 간의 병원성에 극명한 차이를 확인하였다. 이 연구는 *P. capsici* 균주의 비교유전체 분석을 통해 유전체 수준에서 다양한 병원성의 원인을 확인하였으며, 다른 Phytophthora 속과의 비교를 통해 유전자의 진화 및 유사도 연구를 수행하는데 도움을 줄 수 있을 것이다. 본 연구 내에서 비교한 병원성 정도와 유전체 서열 정보는 *P. capsici* 균주 간에 병원성 정도와 유전체 구조

적 변이를 확인함으로써 난균류 병원균의 병원성을 이해하는데 도움을 줄 수 있을 것이고, 고추역병 저항성 품종 육성에 도움을 줄 수 있을 것이다.

학번: 2015-30766