Fall 2018

GENETIC DISSECTION OF NON-HOST RESISTANCE TO THE WHEAT STEM RUST PATHOGEN, USING AN INTERSPECIFIC BARBERRY HYBRID

Radhika Bartaula
University of New Hampshire, Durham

GENETIC DISSECTION OF NON-HOST RESISTANCE TO THE WHEAT STEM RUST PATHOGEN, USING AN INTERSPECIFIC BARBERRY HYBRID

By

Radhika Bartaula
B.Sc. in Agriculture, Tribhuwan University 2011, Nepal

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Genetics

September 2018

This dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Genetics by:

Dissertation Director: Dr. Iago Hale, Assistant Professor, Agriculture, Nutrition, and Food system (UNH)

Dr. Thomas M Davis, Professor, Agriculture, Nutrition, and Food system (UNH)

Dr. Matthew MacManes, Assistant Professor of Molecular, Cellular, and Biomedical Sciences (UNH)

Dr. Janet Sullivan, Adjunct Associate Professor, Biological Sciences (UNH)

Dr. Bryan Connolly, Assistant Professor, Biology (Framingham State University)

On 27 July 2018

Original approval signatures are on file with the University of New Hampshire Graduate School.

Dedication

To the memory of my father, Jagannath Bartaula,
for his unrelenting love and fostering in me the virtues of self-esteem;
and to my mother, Hari Maya Bartaula,
who raised me to be an independent and persistent woman.

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor Dr. Iago Hale for his fundamental role in my doctoral work. Iago provided me with every bit of guidance, expertise, and encouragement that I needed during my Ph.D., at the same time giving me freedom to grow as an independent research scientist. And during the most difficult times, he always gave me the moral support I needed to move on. I quite simply cannot imagine a better mentor, and it has been an honor to be his first Ph.D. student.

I would also like to thank my dissertation committee members for their guidance, suggestions, and insightful comments. In particular, I want to thank Dr. Thomas M. Davis for introducing philosophical aspect of a PhD and his generosity in the use of his lab equipment. I want to thank Dr. Matt MacManes for his advice on the bioinformatics part of my dissertation project, and Dr. Janet Sullivan for her insights into plant systematics. Special thanks to Dr. Bryan Connolly for sharing his taxonomic expertise so willingly and for always being so excited about my dissertation project.

My sincere thanks to Dr. Arthur Melo who provided tremendous help with bioinformatics work and made numerous contributions to my dissertation. Arthur has been a great mentor and a friend to me. Thank you for being so positive and supportive!

In addition, I want to thank Luke Hydock at the McFarlane Greenhouses for his continuous support in maintaining of the barberry plants in the greenhouse. I am also grateful to Alicia Cheevor for her help in the fieldworks.

I also want to thank all my friends and colleagues for their support and acts of kindness during my graduate studies. Specifically, I am indebted to Rekha Chhetri for her constant support and friendship.

Finally, I could not have succeeded in graduate school without the support of my family and Diwas Silwal. I am and always will be grateful to have such an amazing family and a best friend.

This work was supported by the Bill & Melinda Gates Foundation [OPPGD1389, OPP1133199]; and USDA NIFA Hatch Multistate Project NE009 [NH00611-R]. I was also supported by Summer TA fellowship (2015) provided by the University of New Hampshire.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

CHAPTER I

CHAPTER II

LIST OF FIGURES

ABSTRACT

GENETIC DISSECTION OF NON-HOST RESISTANCE TO THE WHEAT STEM RUST PATHOGEN, USING AN INTERSPECIFIC BARBERRY HYBRID

By

Radhika Bartaula

University of New Hampshire
September 2018

Stem rust, caused by the macrocyclic fungal pathogen *P. graminis* (*Pg*), is one of the most devastating diseases of wheat and other small grains globally; and the emergence of new stem rust races virulent on deployed resistance genes brings urgency to the discovery of more durable sources of genetic resistance. Given its intrinsic durability and effectiveness across a broad range of pathogens, non-host resistance (NHR) presents a compelling strategy for achieving long-term rust control in wheat. However, NHR to *Pg* (*Pg*-NHR) remains largely unexplored as a protection strategy in wheat, in part due to the challenge of developing a genetically tractable system in which *Pg*-NHR segregates. In this dissertation, an investigation of *Pg*-NHR is undertaken via the pathogen's alternate (sexual) host, barberry (*Berberis* spp.). Within the highly diverse *Berberis* genus, numerous species function as alternate hosts to *Pg* but others are non-hosts. European barberry (*B. vulgaris* L.), for example, is susceptible to *Pg* infection but Japanese barberry (*B. thunbergii* DC.) is a non-host. In this study, the nothospecies *B.* ×*ottawensis* C.K. Scheid, an inter-specific hybrid between *Pg*-susceptible *B. vulgaris* and *Pg*-resistant *B. thunbergii*, is explored as a possible means of mapping the gene(s) underlying the apparent *Pg*-NHR exhibited by *B. thunbergii*. The overall goal of this research is to contribute to the global search for novel sources of potentially durable stem rust resistance genes.

The first chapter describes a field study conducted in western Massachusetts, in which a natural population of *B.* ×*ottawensis* was characterized to determine if the hybrid can be used to genetically dissect the *Pg*-NHR exhibited by *B. thunbergii*. A population of 63 *B.* ×*ottawensis* individuals were clonally propagated, phenotyped for disease response to *Pg* via controlled inoculation using overwintered telia of *Pg* found on naturally infected *E. repens*, and genotyped using the *de novo* genotyping-by-sequencing (GBS) pipeline GBS-SNP-CROP. Controlled inoculation of a subset of 53 *B.* ×*ottawensis* accessions, verified via GBS to be true, first-generation hybrids, revealed 51% susceptible, 33% resistant, and 16% intermediate phenotypes. Although such variation in disease response within a natural population of F_1 hybrids could be explained by non-nuclear (cytoplasmic) inheritance of resistance, a similar pattern of segregation was observed in a population of *B.* ×*ottawensis* full-sibs, developed via controlled crosses. The results of this first chapter demonstrate not only that the *Pg*-NHR observed in *B. thunbergii* segregates among F_1 interspecific hybrids with *Pg*-susceptible *B. vulgaris* but that the resistance is likely nuclearly inherited. Therefore, at least in principle, the gene(s) underlying *Pg*-NHR in *B. thunbergii* should be mappable in an F_1 population derived from the controlled hybridization of the two parental species.

Building on the results of first chapter, the second chapter of this dissertation details the generation and use of a bi-parental *B.* ×*ottawensis* mapping population to develop genetic linkage maps for both parental species and begin mapping the gene(s) underlying *Pg*-NHR in *B. thunbergii*. Using 162 full-sib F_1 hybrids and a total of 15,411 sequence variants (SNPs and indels) identified between the parents via GBS, genetic linkage maps with 1,757 and 706 markers were constructed for *B. thunbergii* accession 'BtUCONN1' and *B. vulgaris* accession 'Wagon Hill', respectively. In each map, the markers segregated into 14 linkage groups, in

agreement with the 14 chromosomes present in these *Berberis* spp. The total lengths of the linkage maps were 1474 cM (*B. thunbergii*) and 1714 cM (*B. vulgaris*), with average distances between markers of 2.6 cM and 5.5 cM. QTL analysis for *Pg* resistance led to the identification of a single QTL, dubbed *QPgr-3S*, on the short arm of chromosome 3 of *B. thunbergii*. The peak LOD score of *QPgr-3S* is 28.2, and the QTL spans 13 cM, bounded by the distal SNP marker *M411* and proximal SNP marker *M969*. To gain further insight into the *QPgr-3S* region, a chromosome-level 1.2 Gb draft genome for *B. thunbergii* was assembled using long PacBio reads and Hi-C data. By anchoring the *B. thunbergii* linkage map to the draft genome, the 13 cM *QPgr-3S* region was found to correspond to ~3.4 Mbp, represented by 10 contigs. Using a 189.3 Mb transcriptome assembled from a multiple tissue library of RNA-seq data, the *QPgr-3S* region was found to contain 99 genes. To help narrow this list to candidate genes of highest priority for subsequent investigation, a combination of approaches was taken. Specifically, annotation of the QTL region and differential gene expression analysis led to the identification of 12 candidate genes within the region. Of those, two emerge as particularly noteworthy due to the fact that both appear to be differentially expressed in *B. thunbergii* during *Pg* inoculation and belong to families implicated in disease resistance in other plant-pathogen systems, namely leucine-rich repeat receptor-like kinases (gene *GG9708*) and zinc ion binding SSM4 proteins (gene *GG9868*).

The final chapter provides a summary of the research and offers recommendations for future studies, building on the results to date. The original vision of this research was to explore the possible use of a heretofore uncharacterized system in identifying novel sources of resistance to *Pg*, and much progress has been made toward this end. Significant follow up research is needed, however, and the objectives of that research should be to validate, further characterize,

and continue to dissect the *QPgr-3S* region; test candidate gene hypotheses; and seriously confront the difficult question of the relevance of gene(s) regulating *Pg*-NHR in *B. thunbergii* to durable strategies of stem rust resistance in wheat.

This dissertation demonstrates the feasibility of dissecting the genetics of a source of *Pg*-NHR in a novel way, using the interspecific hybrid barberry *B. ×ottawensis*. The genetic linkage maps developed in this work are the first available for any species within the ancient plant family Berberidaceae; and the draft genome is the first available for any plant within the order Ranunculales. Altogether, the resources developed in this study not only establish *B. ×ottawensis* as a viable system to dissect *Pg*-NHR but also makes available valuable resources for global rust surveillance work and ornamental horticulture breeding.

INTRODUCTION

1 Overview

Stem rust (casual organism *Puccinia graminis* - *Pg*) is one of the most globally important diseases of wheat and other cereal grains, responsible for severe epidemics and major recurring yield losses [1–3]. Historically, the ability of *Puccinia* spp. to rapidly evolve new forms and combinations of virulence has necessitated the continual development and deployment of improved wheat varieties with updated sources of rust resistance [4]. Due to a combination of concerted breeding efforts and the systematic eradication of the pathogen's alternate host from major wheat producing regions, wheat stem rust epidemics became less common over the course of the 20th century, to the extent that concern over the disease largely disappeared by the 1970's [5]. In 1999, however, that concern was rekindled by the emergence of Ug99, a new stem rust race radiating out of East Africa. Ug99 was noteworthy because it was virulent on *Sr31*, a stem rust resistance gene conferring protection to the vast majority of the world's currently grown varieties [6]. The rapid proliferation and increase in both the virulence and aggression of the Ug99 family of races [3] sounded the alarm for wheat scientists; and the search for new sources of resistance, particularly durable (i.e. race non-specific) resistance [7], continues to be a top priority in terms of wheat security.

To complement this global search for novel sources of potentially durable *Pg* resistance within the *Triticum* genepool, the current research was undertaken to investigate the mechanism(s) of rust resistance in *Pg*'s alternate host, barberries. Within the highly diverse *Berberis* genus, numerous species function as alternate (sexual) hosts to *Pg* [8,9]; but others do not. For example, while the common European barberry, *B. vulgaris*, is susceptible to *Pg*

infection, Japanese barberry (*B. thunbergii*) is identified as a non-host [10]. The molecular basis of *Pg* non-host resistance (NHR) in *B. thunbergii* is unknown, but it is of enormous interest given the historic durability of this resistance and the potential to transfer such resistance mechanism(s) by transgenesis to crop plants [11]. In his dream for tomorrow, Dr. Norman Borlaug envisioned the ability of biotechnology to transfer the rust immunity of rice to wheat, forever solving the rust problem in one of humanity's most important staple crops [12]. When simply inherited (e.g. maize *Rxo1* gene), the proof of concept for such a visionary transfer of NHR between species has already been shown [13]; but studies suggest that rust NHR in rice is far from simply inherited [11] . Dr. Borlaug's vision of looking beyond the *Triticum* genepool for novel sources of durable resistance remains compelling, however. The evolutionary relationship between barberry and *Pg* is thought to predate the pathogen's host jump to the cereals [14]; thus barberry presents itself as a unique system to explore the basis of potentially durable rust resistance. Furthermore, although the multi-million dollar commercial trade of ornamental barberries in the U.S. is tightly regulated due to this issue of rust susceptibility on some *Berberis* spp. [15], no molecular markers are available that can distinguish susceptible and resistant cultivars. This lack of genetic information also stymies current barberry surveillance work, occurring globally, as no molecular taxonomic tools are available to help wheat rust researchers identify alternate host species in the field. To address these gaps in knowledge, this research seeks to develop genetic and genomic resources for *Berberis* spp. and to identify the genetic mechanism(s) of NHR to stem rust pathogen in the alternate host.

2 Research objectives

The main goal of this research was to understand the genetic mechanism(s) of *Pg*-NHR exhibited by *B. thunbergii*, with the aim of discovering novel sources of potentially durable and transferable stem rust resistance genes. To meet this broad goal, the specific objectives of this dissertation were to:

1. Characterize a naturally occurring population of *Berberis* spp. to explore the viability of the "*Berberis*-*Pg*" pathosystem as a novel system for characterizing and mapping the gene(s) underlying the putative *Pg*-NHR in non-host species;
2. Develop genetic and genomic resources of relevant *Berberis* spp. to facilitate the above objective, including:
    - High-resolution genetic linkage maps of both *Pg* non-host *B. thunbergii* and *Pg* susceptible *B. vulgaris*;
    - Chromosome-level reference genome of *B. thunbergii*; and
    - Assembled transcriptome of *B. thunbergii*
3. Perform quantitative trait loci (QTL) analysis for identifying loci regulating *Pg*-NHR in *B. thunbergii*;
4. Identify a list of candidate genes governing *Pg*-NHR response in *B. thunbergii*; and
5. Develop a strategy for subsequent candidate gene validation and future studies

3 Background

3.1 The stem rust pathogen

Stem rust disease, caused by *Puccinia graminis* (*Pg*), is one of the most devastating diseases of small grain cereal crops, affecting wheat, barley, rye, oat, and triticale [1,2]. *Pg* is a fungal pathogen belonging to the phylum Basidiomycota, class Urediniomycetes, order

Uredinales, and family Pucciniaceae [1]. The host range of *Pg* is very broad, including at least 365 species of plants across 54 genera [16]. Based on its host range, the species is further divided into subgroups commonly referred to as formae speciales (f. sp.). The commonly known f. sp. of *Pg* and their respective cereal hosts are *P. graminis* f. sp. *tritici* (wheat), *P. graminis* f. sp. *hordei* (barley), *P. graminis* f. sp. *avenae* (oat), *P. graminis* f. sp. *secalis* (rye), *P. graminis* f. sp. *poae* (blue grass), and *P. graminis* f. sp. *lolii* (stiff brome). While many cereals and grasses (non-cereals) affected by *Pg* are valuable dietary and industrial crops, historically the threat posed by *P. graminis* f. sp. *tritici* (*Pgt*) to wheat production is the most feared.

3.1.1 Biology and life cycle of the wheat stem rust pathogen

Pg is an obligate parasite, meaning that it requires living host tissue for growth and reproduction. In the absence of living host tissue, however, *Pg* can survive in the form of dormant spores (teliospores) for years [17]. *Pg* is a heteroecious, macrocyclic fungus that requires two unrelated host plants to compete its life cycle, namely various species of cereals (Poaceae) and barberry (Berberidaceae). The asexual spores of *Pg* (urediniospores, teliospores, and basidiospores) are produced during its asexual cycle on its gramineous hosts, and the sexual spores (pycniospores and aeciospores) are produced during its sexual cycle on barberry. The various cereal hosts are often referred to as the primary hosts (i.e. the hosts of greatest concern), while barberry is referred to as the alternate host; but the term "primary" refers only to the agricultural importance of these species relative to barberry, not to their role as asexual hosts. For other rusts, for example the cedar-apple rust (causal organism *Gymnosporangium juniperi-virginianae*), apple (*Malus* spp.) is often referred to as the primary host, even though as the

sexual host it is more analogous to barberry in the life cycle of *Pg*. The complete life cycle of the stem rust fungus is shown in Fig.1.

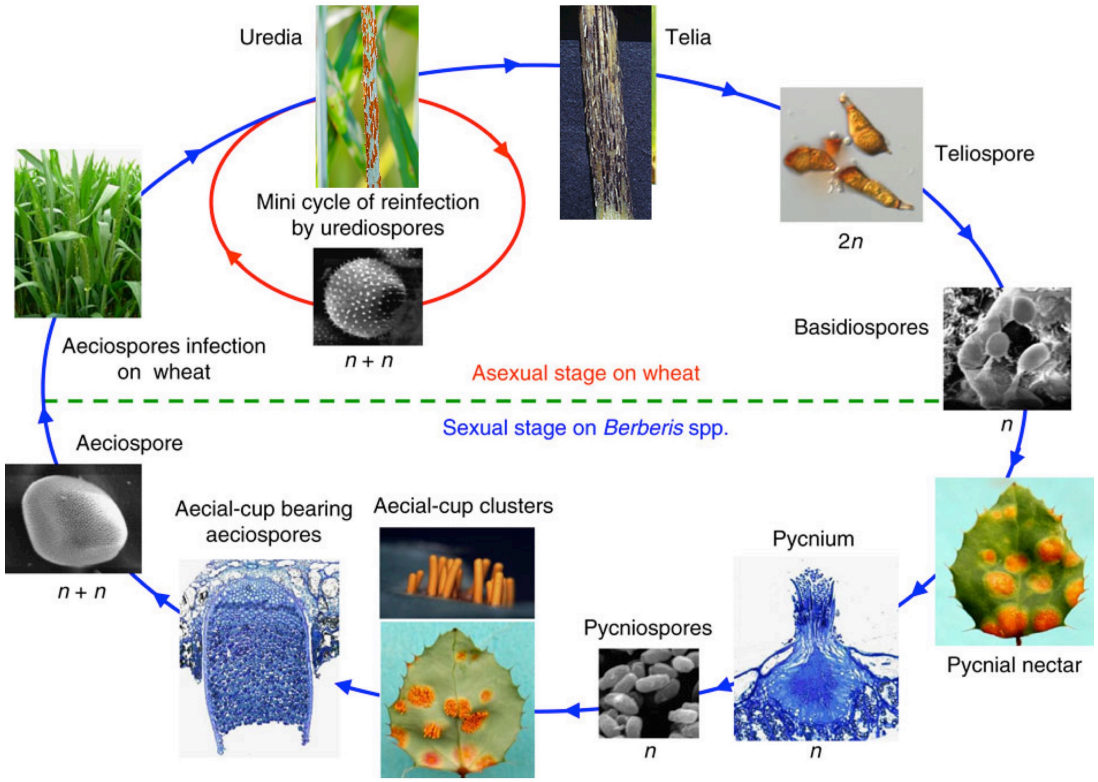


Fig. 1 Life cycle of heteroecious, macrocyclic *Puccinia graminis*, adapted from Zheng et.al [18]. To complete its complex life cycle, *Pg* alternates between cereal crops (Poaceae) as its primary host and barberry (Berberidaceae) as its alternate host.

Each season, the disease cycle of wheat stem rust can be thought to begin when a wheat crop is exposed to infectious *Pgt* spores under appropriate environmental conditions. The possible sources of inoculum for wheat are either aeciospores, produced during sexual recombination on barberry, or urediniospores, originating from infected volunteer grasses or wheat from distant regions. Once this primary inoculum lands on the surface of a wheat plant (usually the stem) and comes into contact with free moisture, the spores can germinate within 1-3

h [19]. A germinating spore produces a germ tube, which locates stomata and forms a specialized penetration structure, known as an appressorium [20]. Upon exposure to sunlight, this appressorium develops a penetration peg and subsequently forms a substomatal vesicle that gives rise to the primary infection hypha [20,21]. As the fungus grows, it produces specialized feeding structures, known as haustoria, which allow the pathogen to directly uptake nutrients from within the plant cell [19,20]. The haustoria then branch to produce secondary infection hyphae, drawing more nutrients from the plant cells. As the invasion progresses, the fungus produces a mass of urediniospores which erupts through the stem's surface, providing a new source of inoculum [20]. The process from initial infection to the production of new spores can take as little at 10 days [20]; and these urediniospores, genetically identical to the primary inoculum, spread to neighboring plants to cause a new round of infection. Under the right environmental conditions, epidemic levels of disease can quickly develop over the course of the growing season. Stem rust is found wherever wheat production occurs, in part because urediniospores can move extremely long distances by wind. Long-distance transport through prevailing winds commonly occurs across the North American Great Plains [22], but the rare long-distance dispersal of about 8000 km from southern Africa to Australia has also been reported [23].

When the growing season comes to an end and there is no longer any living host tissue for pathogen to draw nutrients from, the fungus produces thick-walled, two-celled overwintering spores called teliospores. The teliospores of *Pg* have two nuclei, one of + mating type and the other of - mating type [24]. During the process of entering dormancy, these two nuclei fuse to produce a single diploid nucleus, which quickly divides meiotically to produce four haploid nuclei. Dormant, overwintering teliospores can survive on dead wheat stubble until spring when

each haploid nucleus germinates to produce another spore type known as a basidiospore [16,25]. Basidiospores are not cable of infecting wheat crops; instead, they infect young leaves of its alternate host, for example *B. vulgaris*.

Upon infection on barberry, each haploid basidiospore spore produces pycniospores that act as gametes [24]. When + type pycniospores are fertilized with - type pycniospores from a nearby infection (pycnia), the fertilized dikaryotic structure leads to the eventual production of diploid aeciospores. Pycnia typically form on the adaxial surface of barberry leaves within 5-6 days of infection, and mature aeciospores appear on the abaxial surface 5-7 days after pycniospore fertilization. The lifecycle is completed when aeciospores, carrying novel recombinations of genetic material, germinate on the surface of a wheat plant to begin the next cycle of infection.

The life cycle of *Pgt* illustrates the important role of the alternate host on stem rust disease epidemiology. Aeciospores from barberry provide early-season of inoculum, thereby promoting the rapid development of stem rust epidemics [26]. Without barberry, the teliospores of *Pgt* represent a dead end in its life cycle, so barberry also serves as the bridge to carry the rust fungus from one crop season to the next. Finally, as the sexual host, barberry serves as the breeding ground for the pathogen and thus plays a significant role in the generation of new virulent races of the wheat stem rust pathogen [9].

3.1.2 Symptoms and signs of stem rust

Stem rust infection in cereals occurs mainly on the stems and leaf sheaths of susceptible varieties. The symptom of the disease in cereals appears as brick-red pustules (Fig. 2A), the result of powdery masses of urediniospores similar in appearance to rust spots on a weathered

iron surface. Later in the season, teliospores appear as a layer of black spores, causing the stems of heavily infected plants to appear blackened (Fig. 2B); hence stem rust is often commonly referred to as black rust. On barberry, pycniospore-containing pycnia usually appear on the adaxial leaf surfaces, while striking aeciospore-bearing aecial horns break through the abaxial surface. Pycniospores (Fig. 2C) appear as a bright orange-yellow in sticky honeydew and aeciospores (Fig. 2D) appear as bright yellow rusty powder in aecial cups.

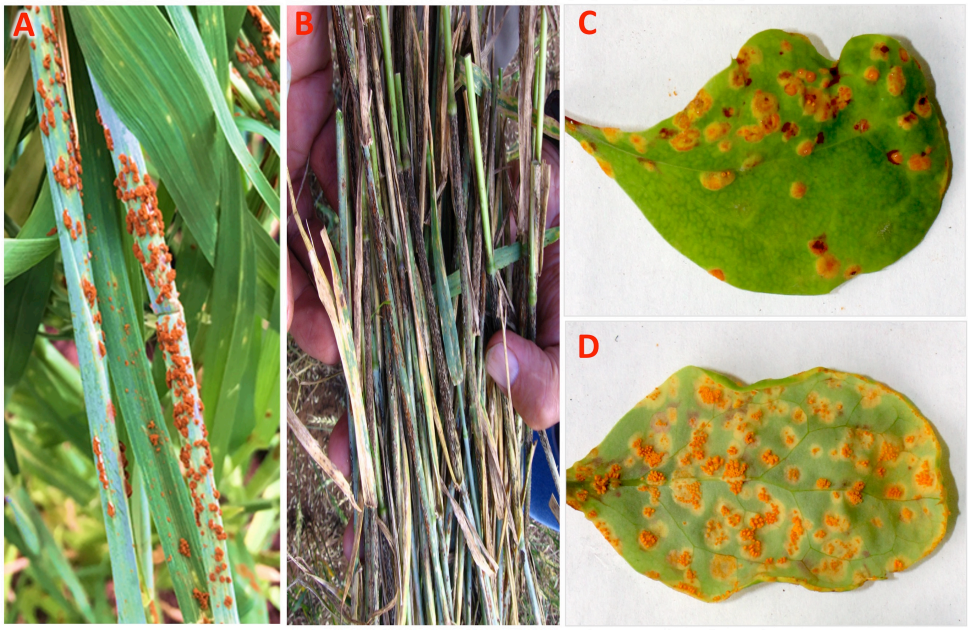


Fig. 2 Symptoms of infection by the stem rust pathogen on its primary and alternate hosts**. (A)** Masses of brick-red urediniospores erupting through the epidermis of a wheat stem. **(B)** Black masses of overwintering teliospores in wheat. **(C)** Bright orange-yellow pycniospores appear on upper surface of a barberry leaf. **(D)** Aecial horns emerge from the lower leaf surface of a barberry leaf. Image A photo credit: R. Park.

3.1.3 Historical significance of stem rust

With its global distribution and its ability to destroy an apparently healthy crop of susceptible cultivar in as little as 3 weeks [1,2,20], stem rust is one of the most feared agricultural diseases in the world. The scourge of rust is recorded as far back as 500 BC, as

evidenced in early Greek and Roman writings which describe the details of sacrificial ceremonies to appease Robigus, the rust god, in an attempt to prevent crop failure [27,28]. Severe stem rust epidemics have been recorded throughout Africa, the Middle East, Australia, New Zealand, Europe, the Americas, and Asia [29]. In the 20th century, stem rust epidemics in Europe resulted in yield losses of up to 33% [30]. In the United States, yield reductions exceeding 50% were noted during epidemic years (1918 to 1976) in major wheat growing areas [31].

3.1.5 Present day risk of stem rust

By the late 20th century, stem rust had been effectively controlled through the development and deployment of resistant wheat varieties, in combination with the systematic eradication of the alternate host *Berberis vulgaris* L. in some major wheat growing areas. In 1998, a new highly virulent race of *Pgt*, referred to as Ug99, was detected in Uganda [32] that overcame *Sr31*, one of the most widely used stem rust resistance gene in wheat cultivars around the world. In the years following the first detection of Ug99, 80% of wheat and 95% of barley cultivars worldwide were considered to be susceptible to Ug99 [3,33]. Since its first detection, thirteen different variants of Ug99 have emerged, combining various other virulences on important resistance genes to the *Sr31* virulence; and the race group has spread to more than 13 countries in Africa and the Middle East [3,32,34–36]. More recent outbreaks of stem rust in Ethiopia (2013-2014) and Kenya (2014) have broken additional resistance genes widely used in modern wheat varieties [37,38]. Additionally, recent outbreaks of the stem rust in United Kingdom and other parts of Europe, have increased the potential for wheat stem rust epidemics in Europe [39,40]. Thus, there is an

ongoing need to explore new sources of resistance to diversify the current pool of genetic resistance [7].

3.2 Barberry and its connection with stem rust

The genus *Berberis*, commonly known as barberries, represents the largest genus within the family Berberidaceae and consists of nearly 400 species [41]. *Berberis* is distributed nearly worldwide, with centers of diversity in southern Asia as well as Central and South America [41,42]. Many *Berberis* spp. are highly valued as ornamental shrubs because of their lush evergreen and/or vibrant, multicolored leaves. In the USA and Canada alone, ornamental barberry is a multi-million dollar industry [43]. Some *Berberis* spp. are also highly valued for their pharmaceutical and medicinal properties [44,45]. In addition, many have also been recognized for their use in the printing and dyeing industry, as well as for their culinary uses [46,47].

Beyond their economic importance, numerous species within the highly diverse *Berberis* are known to function as competent sexual hosts to the stem rust pathogen [8,9]. As the alternate host of *Pg*, susceptible barberry species play a key role in the sexual stage of the pathogen's life cycle and provide an early source of inoculum for the rapid onset of disease. In particular, the European, or common, barberry (*B. vulgaris*) has been implicated in the wheat-stem rust pathosystem for centuries, as evidenced by the existence of barberry eradication laws as far back as the 1600's. Barberry eradication programs undertaken as a means of controlling wheat stem rust epidemics in both the United Kingdom and the USA had significant positive effects on the control of stem rust epidemics [26,48]. Under a massive *B. vulgaris* eradication program undertaken by the USA government from 1918-1974, the largest plant eradication effort in

history, more than 500 million common barberry plants were destroyed throughout the wheat-growing areas of the North Central Plains [49]. This effort resulted not only in a significant reduction of early season inoculum but also a reduction in the genetic variation in the fungal population by effectively eliminating its sexual cycle [5]. Similarly, barberry eradication in the United Kingdom during the late nineteenth and early twentieth centuries resulted in the almost complete disappearance of wheat stem rust in those countries [48].

Over more recent decades, however, barberry has begun to re-establish in these historic eradication zones and their populations are increasing in major wheat growing areas [50]. Because of the re-emergence of barberry on many sites in Minnesota and recent changes observed in the regional *Pg* population, there is increasing concern over the possibility of stem rust epidemics once again occurring in the area [50]. Outside of the historic federal eradication zone, New England is another region of concern. In the northeast, *B. vulgaris* plants are so widespread that the species is considered to be invasive [51]; and a growing interest in revitalizing the small grain industry in the region raises a serious concern about the potential for the region functioning as an epidemiological hotspot for the emergence of new virulent races of *Pg*. Similar concerns exist in the United Kingdom, where purposeful replanting of *B. vulgaris* has been underway for decades, largely due to a habitat conservation program for the endangered barberry carpet moth [52]. The recent detection of the stem rust pathogen on barberry in the United Kingdom raises questions regarding the potential role of barberry in the rust epidemiology of Europe.

3.3 Breeding for stem rust resistance

Although cultural practices and fungicide application are widely used methods for disease control in many crops, such practices are not viable options for wheat producers due to the low profit margins of the crop. Thus, genetic resistance has long been viewed as the most economically appropriate and environmentally sustainable method for disease control in wheat. Race-specific resistance genes, often referred to as major resistance genes, are the most common types of disease resistance genes used in breeding programs. Such genes follow the "gene-for-gene" model, conferring complete resistance to specific races of the pathogen that carry a corresponding avirulence gene [53]. Major genes in plants are known to predominantly encode nucleotide-binding and leucine-rich repeat (NB-LRR) proteins [54]. To date, more than 50 major stem rust resistance genes have been described in wheat and its close relatives [55]. Of these, only 5 have been successfully cloned (Table 1) and they all encode NB-LRR proteins [53]. Although many R genes have been identified and utilized effectively in wheat cultivars, the resistance afforded by such genes often breaks down due to the evolution of the pathogen, as exemplified by the recent emergence of Ug99.

Table 1. List of major wheat stem rust resistance genes that have been cloned

Species	Gene	Gene product	Source
Triticum monococcum	*Sr22*	CC-NB-LRR[1]	Steuernagel et al. 2016 [56]
Aegilops tauschii	*Sr33*	CC-NB-LRR	Periyannan et al. 2013 [57]
Triticum monococcum	*Sr35*	CC-NB-LRR	Saintenac et al. 2013 [58]
Aegilops tauschii	*Sr45*	CC-NB-LRR	Steuernagel et al. 2016 [56]
Secale cerale	*Sr50*	CC-NB-LRR	Mago et al. 2015 [59]

[1] Coiled Coil Nucleotide Binding Leucine Rich Repeat protein

In contrast to major resistance, race non-specific resistance is another category of genetic resistance to the wheat stem rust. Generally speaking, race non-specific resistance is defined as broad spectrum, meaning the underlying genes confer resistance against all races of a pathogen

species. In many cases, the conferred resistance is partial, or quantitative, and is sometimes effective against multiple pathogens [53]. In wheat, many such partial resistance genes are referred to as adult plant resistance (APR) genes because they provide protection only at adult stages of wheat development. In contrast to NB-LRR resistance genes, some APR genes have proven themselves to be highly durable. The widely deployed stem rust resistance complex *Sr2*, for example, has remained effective for all races of stem rust for nearly 100 years [60].

Apart from race-specific and race non-specific genes, another far less explored option for durable resistance is the use of non-host resistance (NHR) genes. NHR, in which an entire plant species is resistant to all genetic variants of a pathogen species, is the most common type of resistance exhibited by plants [61]. Due to its durability and inherent race non-specificity, NHR has gained more attention recently as a potential strategy for improving crop resistance; but the genetic basis of NHR remains poorly understood. Currently, plant immune responses are thought to fall into two major classes, those triggered by pathogen-associated molecular patterns (PAMP-triggered immunity; PTI) and those triggered by pathogen effectors (effector-triggered immunity; ETI). Whereas PTI is thought to be associated with receptor kinases located in the plant plasma membrane, ETI is triggered by intracellular NB-LRR proteins [53]. Recent studies of NHR suggest that it may be a form of PTI, though its specific mode of action and its relationship to basal defense response remains unclear [62].

Summary of dissertation research and thesis chapter descriptions

This thesis explores NHR as a protection strategy against the wheat stem rust pathogen *Puccinia graminis*. Intrinsically durable and the most common form of genetic resistance, NHR holds promise as a strategy for disease protection. Despite the potential relevance of NHR for

disease resistance, however, the genetic basis of this form of resistance remains poorly understood, in part due to the inherent challenge of developing a genetically tractable system in which genes controlling NHR segregate. Because all of the individuals in a non-host plant species are, by definition, resistant to the pathogen, relevant genetic analyses are difficult to perform. The objective of this research was to overcome this fundamental obstacle by developing a genetically tractable system for *Pg*-NHR in the pathogen's alternate host, barberries. Such a system was realized in the form of an interspecific hybrid between *Pg* non-host *B. thunbergii* and *Pg* susceptible *B. vulgaris*; and foundational genetic and genomic resources were developed to facilitate its use. The development of this unique pathosystem allows for the first time the ability to associate a genomic region with *Pg*-NHR and identify underlying candidate genes.

Chapter 1*: "An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen Puccinia graminis"* is a published manuscript in Journal of Experimental Botany [63] that characterizes a natural population of *Berberis spp*. present in New England. By demonstrating segregation of *Pg*-NHR in an F_1 interspecific hybrid population (*Pg* susceptible *B. vulgaris* x *Pg* non-host *B. thunbergii*), this chapter establishes the "*Berberis-Pg*" pathosystem as a viable model for characterizing and mapping the gene(s) underlying the putative *Pg*-NHR of *B. thunbergii*. This chapter also rules out non-nuclear (maternal) mode of inheritance of NHR in *B. thunbergii*.

Chapter 2*: "Mapping non-host resistance to the stem rust pathogen in an interspecific barberry hybrid"* describes the development of foundational genetic and genomic resources for

Berberis spp. to enable QTL analysis and the positing of candidate genes for *Pg*-NHR. To conduct QTL analysis of the *Pg*-NHR exhibited by *B. thunbergii*, an F_1 bi-parental mapping population was developed via a controlled cross between *Pg* susceptible *B. vulgaris* and *Pg* non-host *B. thunbergii*. Using genotyping-by-sequencing data, genetic linkage maps were constructed for both parental species; and a single QTL associated with *Pg*-NHR was located on the short arm of chromosome 3 of *B. thunbergii*.

To gain further insight into the QTL region, a chromosome-level draft genome for *B. thunbergii* was assembled using long PacBio reads and Hi-C data. The QTL region was found to correspond to ~3.4 Mbp in the physical map and to contain 99 genes. Transcriptome based annotation of the QTL region and differential gene expression analysis led to the identification of 12 candidate genes within the region. Of those, two emerged as noteworthy due to their differential expression in *B. thunbergii* during *Pg* inoculation and both genes belong to families implicated in disease resistance in other plant-pathogen systems, namely leucine-rich repeat receptor-like kinases (gene *GG9708*) and zinc ion binding SSM4 proteins (gene *GG9868*).

Chapter 3 *"Summary and future directions"* provides a summary of the completed work and offers recommendations for future studies, building on the results to date. Future work should focus on the validation, further characterization, and dissection of the identified QTL, including testing of candidate gene hypotheses. Beyond this, now that the *Berberis*-*Pg* system has been shown to be a viable means of probing the mechanism of *Pg*-NHR in *B. thunbergii*, future work must also wrestle with the question of potential translatability of that resistance to wheat.

References

1. Leonard KJ, Szabo LJ. Stem rust of small grains and grasses caused by *Puccinia graminis*. Mol Plant Pathol. 2005; 6:99–111.

2. Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Njau P, Wanyera R, et al. Will stem rust destroy the world's wheat crop? Adv Agron. 2008;98:271–309.

3. Singh RP, Hodson DP, Jin Y, Lagudah ES, Ayliffe MA, Bhavani S, et al. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. Phytopathology. 2015;105:872–884.

4. Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, et al. The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. Science. 2013;341:786–788.

5. Peterson PD. Stem rust of wheat: from ancient enemy to modern foe. American Phytopathological Society (APS Press). 2001.

6. Hale IL, Mamuya I, Singh D. *Sr31*-virulent races (TTKSK, TTKST, and TTTSK) of the wheat stem rust pathogen *Puccinia graminis* f. sp. *tritici* are present in Tanzania. Plant Dis. 2013;97:557–557.

7. Ayliffe M, Singh R, Lagudah E. Durable resistance to wheat stem rust needed. Curr Opin Plant Biol. 2008;11:187–192.

8. Zhao J, Wang L, Wang Z, Chen X, Zhang H, Yao J, et al. Identification of eighteen *Berberis* species as alternate hosts of *Puccinia striiformis* f. sp. *tritici* and virulence variation in the pathogen isolates from natural infection of barberry plants in China. Phytopathology. 2013;103:927–934.

9. Jin Y. Role of *Berberis* spp. as alternate hosts in generating new races of *Puccinia graminis* and *P. striiformis*. Euphytica. 2011;179:105–108.

10. Department of Agriculture. Black Stem Rust; Identification Requirements and Addition of Rust Resistant Varieties. Federal Register, No. 36. 67.

11. **Ayliffe M, Devilla R, Mago R, White R, Talbot M, Pryor A, et al**. Nonhost resistance of rice to rust pathogens. Mol Plant Microbe Interact. 2011;24:1143–1155.

12. Borlaug NE. Ending world hunger. The promise of biotechnology and the threat of antiscience zealotry. Plant Physiol. 2000;124:487–490.

13. Zhao B, Lin X, Poland J, Trick H, Leach J, Hulbert S. A maize resistance gene functions against bacterial streak disease in rice. Proc Natl Acad Sci. 2005;102:15383–15388.

14. Bettgenhaeuser J, Gilbert B, Ayliffe M, Moscou MJ. Nonhost resistance to rust pathogens–a continuation of continua. Front Plant Sci. 2014;5:664.

15. Smith K, Draper M, Simmons K, Bennett R, Hebbar P, Royer M, et al. US preparations for potential introduction of Ug99 strains of wheat stem rust. Outlooks Pest Manag. 2009;20:148–152.

16. Bushnell W. The cereal rusts: origins, specificity, structure, and physiology [Internet]. Elsevier; 2012. Available from:

https://books.google.com/books?hl=en&lr=&id=jIClNPdFDMUC&oi=fnd&pg=PP1&dq=Evolution+at+the+center+of+origin.+Cereal+Rusts&ots=EYvfJ1XCX6&sig=nWWW6ZimixkKlUS7PFZc0BggVOc

17. Kislev ME. Stem rust of wheat 3300 years old found in Israel. Science. 1982;216:993–994.

18. Zheng W, Huang L, Huang J, Wang X, Chen X, Zhao J, et al. Erratum: High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. Nat Commun. 2014;5.

19. Staples RC, Macko V. Germination of urediospores and differentiation of infection structures. Cereal Rusts Orig Specif Struct Physiol. 1984;1.

20. Figueroa M, Upadhyaya NM, Sperschneider J, Park RF, Szabo LJ, Steffenson B, et al. Changing the game: using integrative genomics to probe virulence mechanisms of the stem rust pathogen *Puccinia graminis* f. sp. *tritici*. Front Plant Sci. 2016;7:205.

21. Yirgou D, Caldwell RM. Stomatal penetration of wheat seedlings by stem and leaf rusts in relation to effects of carbon dioxide light and stomatal aperture. Phytopathology. 1968;58:500.

22. Roelfs AP, Bushnell WR. Diseases, Distribution, Epidemiology, and Control. Academic Press; 2014.

23. Luig NH. Epidemiology in Australia and New Zealand. Diseases, Distribution, Epidemiology, and Control. Elsevier; 1985; p. 301–328.

24. Leonard KJ. Black stem rust biology and threat to wheat growers. Cent Plant Board Meet Febr. 2001; p. 5–8.

25. Roelfs AP, Groth JV. *Puccinia graminis* f. sp. *tritici*, black stem rust of *Triticum* spp. Adv Plant Pathol. 1988;6:345–361.

26. Roelfs AP. Effects of Barberry eradication. Plant Dis. 1982;66:177.

27. McIntosh RA, Wellings CR, Park RF. Wheat rusts: an atlas of resistance genes. CSIRO Publishing. 1995.

28. Chester KS. The nature and prevention of the cereal rusts as exemplified in the leaf rust of wheat. Chronica Botanica Company; USA; 1946.

29. Shank R. Wheat stem rust and drought effects on Bale agricultural production and future prospects. Rep Febr. 1994;17–28.

30. Zadoks JC, Bouwman JJ. Epidemiology in Europe. Dis Distrib Epidemiol Control. Elsevier. 1985; p. 329–369.

31. Roelfs AP. Estimated losses caused by rust in small grain cereals in the United States, 1918-76. Dept. of Agriculture, Agricultural Research Service: for sale by the Supt. of Docs., US Govt. Print. Off. 1978.

32. Pretorius ZA, Singh RP, Wagoire WW, Payne TS. Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis*. f. sp. *tritici* in Uganda. Plant Dis. 2000;84:203–203.

33. Steffenson BJ, Zhou H, Chai Y, Grando S. Vulnerability of cultivated and wild barley to African stem rust race TTKSK. Adv Barley Sci. Springer. 2013; p. 243–255.

34. Nazari K, Mafi M, Yahyaoui A, Singh RP, Park RF. Detection of wheat stem rust (*Puccinia graminis* f. sp. *tritici*) race TTKSK (Ug99) in Iran. Plant Dis. 2009;93:317–317.

35. Newcomb M, Olivera PD, Rouse MN, Szabo LJ, Johnson J, Gale S, et al. Kenyan isolates of *Puccinia graminis* f. sp. *tritici* from 2008 to 2014: Virulence to *SrTmp* in the Ug99 race group and implications for breeding programs. Phytopathology. 2016;106:729–736.

36. Fetch T, Zegeye T, Park RF, Hodson D, Wanyera R. Detection of wheat stem rust races TTHSK and PTKTK in the Ug99 race group in Kenya in 2014. Plant Dis. 2016;100:1495–1495.

37. Olivera P, Newcomb M, Szabo LJ, Rouse M, Johnson J, Gale S, et al. Phenotypic and genotypic characterization of race TKTTF of *Puccinia graminis* f. sp. *tritici* that caused a wheat stem rust epidemic in southern Ethiopia in 2013–14. Phytopathology. 2015;105:917–928.

38. Patpour M, Hovmøller MS, Justesen AF, Newcomb M, Olivera P, Jin Y, et al. Emergence of virulence to *SrTmp* in the Ug99 race group of wheat stem rust, *Puccinia graminis* f. sp. *tritici*, in Africa. Plant Dis. 2016;100:522.

39. Lewis CM, Persoons A, Bebber DP, Kigathi RN, Maintz J, Findlay K, et al. Potential for re-emergence of wheat stem rust in the United Kingdom. Commun Biol. 2018;1:13.

40. Bhattacharya S. Deadly new wheat disease threatens Europe's crops. Nature. 2017;542:145–146.

41. Rounsaville TJ, Ranney TG. Ploidy levels and genome sizes of *Berberis* L. and *Mahonia* Nutt. species, hybrids, and cultivars. HortScience. 2010;45:1029–1033.

42. Ahrendt LWA. *Berberis* and *Mahonia*: a taxonomic revision. Bot J Linn Soc. 1961;57:1–410.

43. Lubell JD, Brand MH, Lehrer JM, Holsinger KE. Detecting the influence of ornamental *Berberis thunbergii* var. *atropurpurea* in invasive populations of *Berberis thunbergii* (Berberidaceae) using AFLP1. Am J Bot. 2008;95:700–705.

44. Maliwichi-Nyirenda CP, Maliwichi LL, Franco M. Medicinal uses of *Berberis holstii* Engl. (Berberidaceae) in Malawi, the only African endemic barberry. J Med Plants Res. 2011;5:1367–1373.

45. **Javadzadeh SM, Ebrahimi A**. The traditional uses and pharmacological effects of different parts *Berberis vulgaris* (berberine) in Iran. Scientia. 2013;1:61–66.

46. Zhang Y-J, Meng A-P, Li J-Q, Dang H-S, Li X-W. Observation on meiotic behavior in three Mahonia species, with special reference to the intergeneric relationship of *Mahonia* and *Berberis*. Caryologia. 2006;59:305–311.

47. Rahimi-Madiseh M, Lorigoini Z, Zamani-gharaghoshi H, Rafieian-kopaei M. *Berberis vulgaris*: specifications and traditional uses. Iran J Basic Med Sci. 2017;20:569.

48. Stakman EC. Barberry eradication prevents black rust in Western Europe. 1923

49. Peterson Jr PD. The common barberry: the past and present situation in Minnesota and the risk of wheat stem rust epidemics. 2003.

50. Peterson PD. The common barberry: the past and present situation in Minnesota and the risk of wheat stem rust epidemics. 2003. Available from: https://repository.lib.ncsu.edu/handle/1840.16/3775

51. Hale IL, Connolly BA, Bartaula R. The Occurrence of Hybrid Barberry, *Berberis ×ottawensis* (Berberidaceae), in New Hampshire and Rhode Island. Rhodora. 2015;117:384–7.

52. Waring P. Successes in conserving the Barberry Carpet moth *Pareulype berberata* (D. & S.) (Geometridae) in England. J Insect Conserv. 2004;8:167–171.

53. Periyannan S, Milne RJ, Figueroa M, Lagudah ES, Dodds PN. An overview of genetic rust resistance: From broad to specific mechanisms. PLoS Pathog. 2017;13:e1006380.

54. Jones JD, Vance RE, Dangl JL. Intracellular innate immune surveillance devices in plants and animals. Science. 2016;354: aaf6395.

55. McIntosh RA, Wellings CR, Park RF. Wheat rusts: an atlas of resistance genes. CSIRO Publishing. 1995.

56. Steuernagel B, Periyannan SK, Hernández-Pinzón I, Witek K, Rouse MN, Yu G, et al. Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. Nat Biotechnol. 2016;34:652.

57. Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, et al. The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. Science. 2013;341:786–788.

58. Saintenac C, Zhang W, Salcedo A, Rouse MN, Trick HN, Akhunov E, et al. Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. Science. 2013;341:783–786.

59. Mago R, Zhang P, Vautrin S, Šimková H, Bansal U, Luo M-C, et al. The wheat *Sr50* gene reveals rich diversity at a cereal disease resistance locus. Nat Plants. 2015;1:15186.

60. Ellis JG, Lagudah ES, Spielmeyer W, Dodds PN. The past, present and future of breeding rust resistant wheat. Front Plant Sci. 2014;5:641.

61. Lipka U, Fuchs R, Kuhns C, Petutschnig E, Lipka V. Live and let die–*Arabidopsis* nonhost resistance to powdery mildews. Eur J Cell Biol. 2010;89:194–199.

62. Senthil-Kumar M, Mysore KS. Nonhost resistance against bacterial pathogens: retrospectives and prospects. Annu Rev Phytopathol. 2013;51:407–427.

63. Bartaula R, Melo AT, Connolly BA, Jin Y, Hale I. An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen *Puccinia graminis*. J Exp Bot. 2018;69:2483–2493.

CHAPTER I

AN INTERSPECIFIC BARBERRY HYBRID ENABLES GENETIC DISSECTION OF NON-HOST RESISTANCE TO THE STEM RUST PATHOGEN *PUCCINIA GRAMINIS*

Radhika Bartaula[1], Arthur T. O. Melo[1], Bryan A. Connolly[2], Yue Jin[3], and Iago Hale[1,*]

[1] Department of Agriculture, Nutrition, and Food Systems, University of New Hampshire, Durham, NH 03824, USA

[2] Department of Biology, Framingham State University, Framingham, MA 01701, USA

[3] USDA-ARS Cereal Disease Laboratory, St. Paul, MN 55108, USA

Correspondence: iago.hale@unh.edu

Received: 28 September 2017

Editorial decision: 15 February 2018

Accepted: 15 February 2018

Editor: Katherine Denby, York University, UK

How to cite this article: Radhika Bartaula, Arthur T O Melo, Bryan A Connolly, Yue Jin, Iago Hale; An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen *Puccinia graminis*, *Journal of Experimental Botany*, Volume 69, Issue 10, 27 April 2018, Pages 2483–2493, https://doi.org/10.1093/jxb/ery066

ABSTRACT

Stem rust, caused by *Puccinia graminis* (*Pg*), remains a devastating disease of wheat, and the emergence of new *Pg* races virulent on deployed resistance genes fuels the ongoing search for sources of durable resistance. Despite its intrinsic durability, non-host resistance (NHR) is largely unexplored as a protection strategy against *Pg*, partly due to the inherent challenge of developing a genetically tractable system within which NHR segregates. Here, we demonstrate that *Pg's* far less studied ancestral host, barberry (*Berberis* spp.), provides such a unique pathosystem. Characterization of a natural population of *B.* ×*ottawensis*, an interspecific hybrid of *Pg*-susceptible *B. vulgaris* and *Pg*-resistant *B. thunbergii* (*Bt*), reveals that this uncommon nothospecies can be used to dissect the genetic mechanism(s) of *Pg*-NHR exhibited by *Bt*. Artificial inoculation of a natural population of *B.* ×*ottawensis* accessions, verified via genotyping by sequencing to be first-generation hybrids, revealed 51% susceptible, 33% resistant, and 16% intermediate phenotypes. Characterization of a *B.* ×*ottawensis* full sib family excluded the possibility of maternal inheritance of the resistance. By demonstrating segregation of *Pg*-NHR in a hybrid population, this study challenges the assumed irrelevance of *Bt* to *Pg* epidemiology and lays a novel foundation for the genetic dissection of NHR to one of agriculture's most studied pathogens.

Keywords: Barberry, genotyping by sequencing, non-host resistance, stem rust, wheat.

Introduction

Stem rust, caused by the fungal pathogen *Puccinia graminis* (*Pg*), is one of the most devastating diseases of wheat and other small grains, and is responsible for severe epidemics and major recur- ring yield losses worldwide (Leonard and Szabo, 2005). The threat of *Pg* to global food security is further enhanced by its ability rapidly to evolve new forms and combinations of virulence (Pretorius *et al.*, 2000). Since the pioneering work of Dr. Elvin Stakman nearly 90 years ago (see Christensen, 1984), tremendous effort has been made by a global community of researchers to identify and deploy genetic sources of *Pg* resistance in wheat cultivars. Despite these efforts, the protection conferred by the vast majority of resistance genes has been temporary, or non- durable, due to evolving virulence (Singh *et al.*, 2015). With concern rekindled over the emergence and spread of new virulent stem rust races, most notably the Ug99 family of races radiating out of East Africa, the search for novel sources of durable resistance has come to be considered essential to achieving long-term wheat security (Stokstad, 2007; Ayliffe *et al.*, 2008).

Non-host resistance (NHR), in which an entire plant species is resistant to all genetic variants of a pathogen species, is the most common type of resistance exhibited by plants (Lipka *et al.*, 2010). Given its intrinsic durability and efficacy across a broad range of pathogens (Thordal-Christensen, 2003; Mysore and Ryu, 2004), NHR presents a compelling strategy for achieving long-term rust control in wheat. In his dream for tomorrow, Dr. Norman Borlaug envisaged the eventual transfer of *Pg*-NHR from rice to wheat, forever solving via biotechnology the historic rust problem plaguing one of humanity's most important staple crops (Borlaug, 2000). When simply inherited, the proof of concept for such a visionary transfer of NHR between species has been demonstrated [e.g. maize gene *Rxo1* for bacterial streak disease of rice

(Zhao *et al.*, 2005)], but some studies suggest that *Pg*-NHR may not be simply inherited (Cheng *et al.*, 2012).

Considerable effort has been made to understand the response to rust pathogens using various non-host and intermediate host pathosystems, including *Uromyces vignae*, *Puccinia triticina*, and *P. striiformis* on the model plant *Arabidopsis thaliana* (Mellersh and Heath, 2003; Shafiei *et al.*, 2007; Cheng *et al.*, 2013); *P. graminis*, *P. triticina*, *P. striiformis*, *P. hordei*, *P. sorghi*, and *Melampsora lini* on rice (Ayliffe *et al.*, 2011); *P. striiformis* on broadbean (Cheng *et al.*, 2012); *P. hordei* and *U. fabae* on wheat (Prats *et al.*, 2007; Zhang *et al.*, 2011); and *P. striiformis*, *P. triticina*, *P. hordei*, *P. coronata*, *P. recondite*, *P. hordei-secalini*, and *P. persistens* on barley (Atienza *et al.*, 2004; Jafary *et al.*, 2008; Dawson *et al.*, 2016).While the specific mechanisms underlying NHR to rust fungi remain unknown, progress has been made in understanding the basal host defense responses of NHR under pathogen pressure. The current body of know- ledge, based largely on histological and cytological studies, suggests that NHR involves multiple mechanisms, including callose deposition, production of reactive oxygen species, phytoalexin synthesis, salicylic acid signaling, and jasmonic acid signaling (Perumalla and Health, 1989;Ayliffe *et al.*, 2011; Zhao *et al.*, 2016). Despite a growing understanding of such basal mechanisms of some forms of NHR, however, little is known about the genetics underlying such responses. Since all of the individuals in a non-host plant species are, by definition, resistant to the pathogen, relevant genetic analyses are difficult to perform. Simply stated, in order to study the genetics of this type of resistance, a genetically tractable system segregating for NHR is required.

Within the highly diverse *Berberis* genus, numerous species are known to function as competent alternate (or sexual) hosts to *Pg* (Jin, 2011; Zhao *et al.*, 2013), but others do not. For

example, European barberry (*B. vulgaris* L.) is susceptible to *Pg* infection, but Japanese barberry (*B. thunbergii* DC.) is identified as a non- host, with no infection observed either under natural conditions or through extensive laboratory testing (Levine and Cotter, 1932). The association of *B. vulgaris* with *Pg* has been implicated in the wheat stem rust pathosystem for centuries, as evidenced by the existence of *B. vulgaris* eradication laws as far back as the 1600s (Zadoks and Bouwman, 1985). From 1918 to 1974, a massive *B. vulgaris* eradication program was undertaken by the US government as a means of controlling wheat stem rust (Peterson, 2013). Under that program, the largest plant eradication effort in history, >500 million common barberry plants were destroyed throughout the North Central Plains of the USA (Peterson, 2003). In contrast, cultivars of *B. thunbergii* continue to this day to be sold as part of a multi-million dollar ornamental landscape industry (Lubell *et al.*, 2008), provided their resistance to *Pg* is confirmed by the USDA Cereal Disease Laboratory through its long-running barberry testing program (Silander and Klepeis, 1999). In the northeastern USA, outside the boundary of the 20th century federal barberry eradication zone, both *B. vulgaris* and *B. thunbergii* are found in great abundance, to the extent that both are considered invasive species (Silander and Klepeis, 1999; Mehrhoff *et al*., 2003). When the two species co-occur, they can hybridize to produce the relatively rare nothospecies *B.* ×*ottawensis*, and several natural populations of this interspecific hybrid have been documented in the region in recent years (Connolly *et al.*, 2013; Hale *et al.*, 2015; Van Splinter *et al.*, 2016).

Despite the evolutionary relationship between some barberry species and *Pg*, and despite wheat stem rust being one of the most intensively researched plant diseases; no attempt has been made to understand the genetic mechanism of *Pg*-NHR exhibited by some *Berberis* spp. It is hypothesized that the modern-day macrocyclic, heteroecious species of *Pg* evolved from a

progenitor that existed in aecial form, parasitizing dicot ancestors of the Berberidaceae prior to its host expansion to the grasses (Leppik, 1961; Wahl *et al.*, 1984). It is thus of interest to investigate the mechanism of NHR exhibited by descendants of the ancestral hosts of cereal rust pathogens. Since *B. vulgaris* is a competent host of *Pg* and *B. thunbergii* is not, their interspecific hybrid presents a unique system for studying the genetic mechanism(s) of the apparent *Pg*-NHR in *B. thunbergii*. For this study, we utilized a natural population of *B. ×ottawensis* to determine if indeed the nothospecies can be used toward this end, thereby providing insight into mechanisms of NHR that may inspire novel strategies of stem rust resistance in wheat.

Materials and methods

Study taxa and field survey

Naturally occurring individuals of three barberry taxa, *B. vulgaris* L. (common or European barberry), *B. thunbergii* DC. (Japanese barberry), and *B. ×ottawensis* C.K. Scheid, were collected from Mass Audubon's Lime Kiln Farm Wildlife Sanctuary in Sheffield, MA, for use in this study. *Berberis vulgaris* was first introduced to North America by European settlers during the 17th century (Grieve, 1931; Gleason and Cronquist, 1963) and is now considered an invasive species throughout many regions of the USA, including New England (Mehrhoff *et al*., 2003). These upright, perennial shrubs grow up to 3 m tall, display 2–5 cm long obovate to obovate–oblong leaves with highly serrated margins (>50 serrations), and have 5–8 cm long pendant racemes of bright yellow flowers (Gleason and Cronquist, 1963; Mehrhoff *et al*., 2003). *Berberis thunbergii*, first introduced to North America as an ornamental plant from Japan in 1875 (Steffey, 1985), is also now considered an invasive species throughout New England, the Midwest, and eastern states. It is a relatively smaller shrub, ranging from 0.5 m to 2.5 m tall, that

displays 1.3–3.8 cm long entire leaves and 1–2 cm long inflorescences with few umbellate but mostly solitary flowers (Gleason and Cronquist, 1963; Mehrhoff *et al*., 2003; Haines, 2011). The third taxon, *B.* ×*ottawensis*, is the nothospecies resulting from the interspecific hybridization of *B. vulgaris* and *B. thunbergii* (Rehder, 1949). This hybrid is intermediate in height and leaf size be- tween the two parental species, with either entire or moderately serrated leaf margins and truncated pendant racemes bearing 5–12 bright yellow flowers (Mehrhoff *et al*., 2003; Connolly *et al.*, 2013). Representative images of the leaf morphologies and inflorescence types of the three taxa are shown in Figs 1 and 2, respectively.

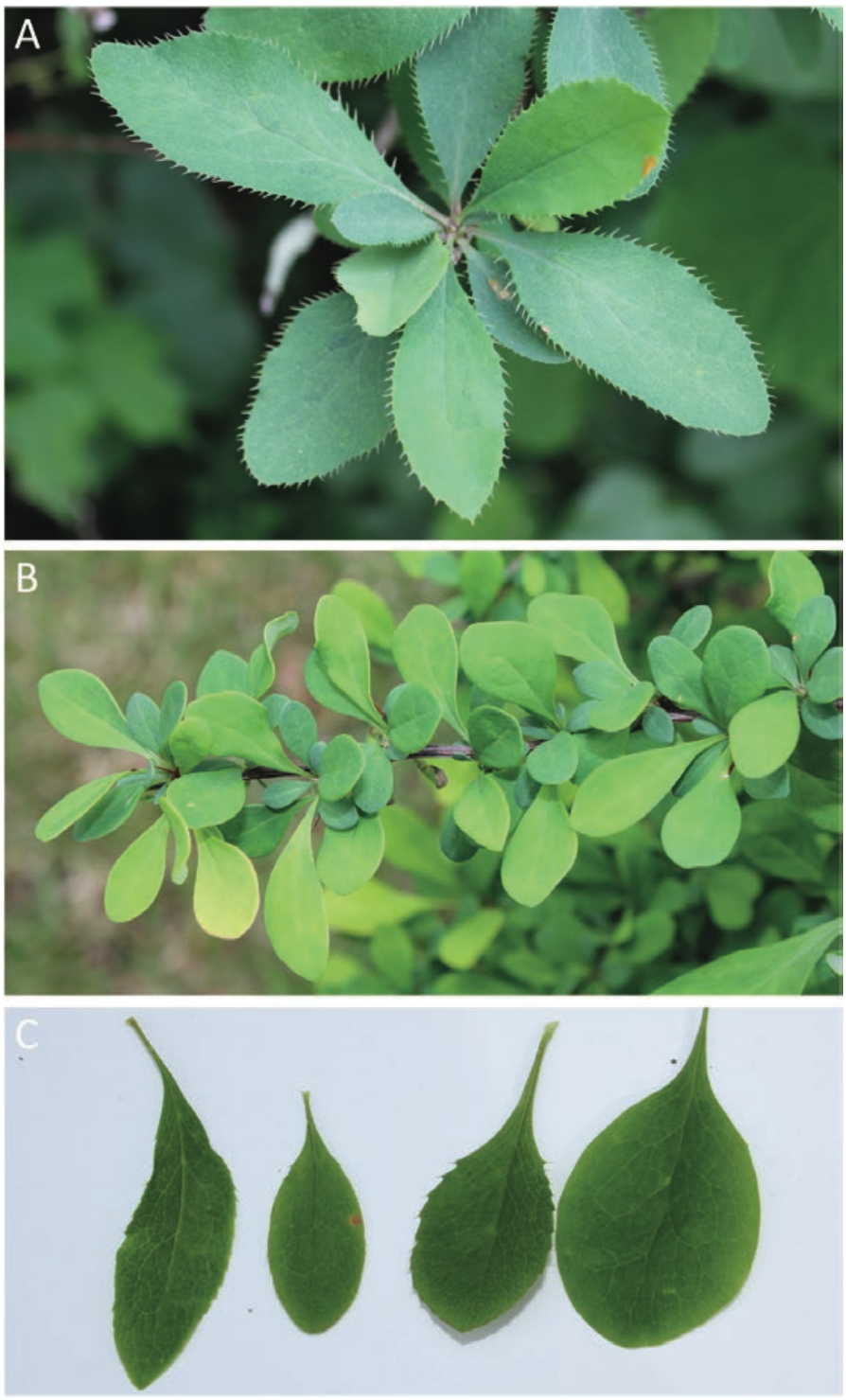


Fig. 1. Leaf morphologies of the three *Berberis* taxa in the study. (A) 2–5 cm long obovate to obovate–oblong leaves with highly serrated margins (>50 serrations) typical of *B*. *vulgaris*; (B) 1–4 cm long, entire leaves typical of *B*. *thunbergii*; and (C) variations in leaf shape and size observed among *B*. ×*ottawensis* hybrids. Among hybrid accessions, leaves vary in shape (ovate, oblong, or obovate), size (2–6 cm long and 1–4 cm wide), and margins (entire to >30 serrations).

Fig. 2. Inflorescence types of the three *Berberis* taxa in the study. (A) 5–8 cm long pendant racemes with bright yellow flowers typical of *B*. *vulgaris*; (B) 1–2 cm long inflorescences with mostly solitary flowers typical of *B*. *thunbergii*; and (C) truncated pendant racemes with 5–12 bright yellow flowers typical of *B*. ×*ottawensis*.

In June 2014, a field survey was conducted to document and characterize the natural populations of *B. vulgaris*, *B. thunbergii*, and *B. ×ottawensis* growing within an ~7 ha area section of the Lime Kiln Farm Wildlife Sanctuary. All *Berberis* spp. plants located within the study area were keyed to species based on the following morphological parameters: plant height, growth habit, leaf morphology, and inflorescence/flower morphology. Of the nearly 1000 plants keyed to species, 190 were selected for subsequent propagation, genotyping, and disease phenotyping, comprising 22 *B. vulgaris*, 27 *B. thunbergii*, and 141 putative *B. ×ottawensis* accessions. Prior to sampling, these 190 plants were assigned unique IDs, labeled with metal tags, and geo-referenced using a Garmin eTrex Vista HCx GPS unit.

In August 2014, 1-year-old semi-hardwood cuttings were taken from each of the 190 selected accessions for clonal propagation via rooting at the MacFarlane Research Greenhouses at the University of New Hampshire (UNH) in Durham, NH. To propagate, the proximal end of a 10–16 cm stem cutting was dipped in dry Hormodin-1 (0.1% indole-3-butyric acid) rooting hormone powder and inserted into moistened vermiculite. The set cuttings were maintained at 30–35 ºC, and light misting was provided every 5–6 min to maintain high relative humidity throughout the propagation period. Once rooted, cuttings were trans- planted into black plastic pots (11.5 cm diameter × 6.5 cm tall) filled with PRO-MIX HP growth media and maintained in the greenhouse for tissue sampling and disease phenotyping.

DNA isolation and genotyping by sequencing

Based on relative propagation success, a random subset of 80 collected accessions were selected for genotypic characterization, including 8 *B. vulgaris*, 9 *B. thunbergii*, and 63 putative *B. ×ottawensis* genotypes (see Supplementary Fig. S1 at *JXB* online). To increase the sampled

genetic diversity and thus the confidence in calling true species-specific variants (i.e. variants polymorphic between *B. vulgaris* and *B. thunbergii* but monomorphic within each species), three additional *B. vulgaris* and four additional *B. thunbergii* accessions were collected from other sites (Randall Road, Lee, NH; Adams Point Road, Durham, NH; Piscataqua Road, Durham, NH; and Agronomy Road, Storrs, CT) and included for genotyping (Supplementary Table S1.). For all 87 accessions, genomic DNA (gDNA) was extracted from ~100 mg of lyophilized leaf tissue using a modified cetyltrimethylammonium bromide (CTAB) method (Doyle, 1987). Prior to genotyping by sequencing (GBS) library preparation, the isolated gDNA was purified using Zymo Research's Genomic DNA Clean & Concentrator^{TM} (Catalog # D4011), following the manufacturer's protocol.

Reduced representation libraries were constructed using the two- enzyme (*Pst*I–*Msp*I) GBS protocol described by Poland *et al.* (2012) and sequenced via 150 bp paired-end (PE) reads on an Illumina HiSeq 2500 at the Hubbard Center for Genome Studies, UNH. Raw FASTQ files were generated using CASAVA 1.8.3 and analyzed using the *de novo* (i.e. reference-free) bioinformatics pipeline GBS-SNP-CROP (Melo *et al.*, 2016). In the first stage of the pipeline, all raw reads were parsed, quality trimmed, and demultiplexed into individual read pairs per genotype. A mock reference was constructed using the high-quality PE reads from all 63 *B. ×ottawensis* accessions, and putative variants, both single nucleotide polymorphisms (SNPs) and bi-allelic indels, were identified by aligning the high-quality PE reads of all 63 *B. ×ottawensis* accessions to the mock reference, following the pipeline's recommended parameters. Using SAMTools (Li *et al.*, 2009), only those read pairs possessing high mapping quality (q>30), exhibiting proper pair orientation, and showing no supplementary alignments were retained for variant calling. The parameters used for variant calling followed the pipeline recommendations

for diploid species, except that the -mnAvgDepth parameter was increased to 7 to enhance accuracy, at the expense of total marker number. Complete details of the GBS-SNP-CROP command lines used in this study, including all specified pipeline parameters, are provided in Supplementary Text S1.

The 11 *B. vulgaris* and 13 *B. thunbergii* accessions included in the GBS library were genotyped by mapping their high-quality PE reads to the *B. ×ottawensis* mock reference and calling alleles only for those marker positions identified as segregating among the natural population of *B. ×ottawensis* lines. For all downstream genetic diversity and pedigree analyses, only those variants located within centroids (i.e. consensus GBS fragments) containing single markers (hereafter referred to as simplex markers) were retained. All parsed, high-quality PE reads are deposited in the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA). Individual barcode assignments, species information, detailed sequencing statistics, and assigned SRA numbers for the 87 genotyped accessions can be found in Supplementary Table S1.

Characterization of genetic diversity

To characterize the genetic diversity both within and among the populations of the three *Berberis* taxa present in the Lime Kiln Farm Wildlife Sanctuary, variants (or markers) were first characterized based on average read depth (D), the percentage of homozygotes (Homo), the percentage of heterozygotes (Hetero), and the proportion of missing genotype calls (NA). The software GenAIEx 6.5.01 (Peakall and Smouse, 2006) was then used to generate descriptive population genetic parameters, such as the numbers of effective alleles (*N*E), the minor allele frequencies (MAF), the observed heterozygosities (*H*O), the unbiased expected heterozygosities

(*H*E), the inbreeding coefficients (*F*IS), and the measure of interspecific genetic structure (*F*ST). Interspecific genetic structure was also assessed via principal component analysis (PCA) through the dudi.pca() function in R (package 'adegenet') (Jombart and Ahmed, 2011).

Pedigree inference for* B. ×ottawensis *individuals

To infer the pedigrees (e.g. F_1, BC_1, etc.) of the 63 genotyped *B. ×ottawen- sis* hybrid individuals, species-specific markers (i.e. variants polymorphic between *B. vulgaris* and *B. thunbergii* but monomorphic within each species) were identified. To ensure the informativeness of such markers across the population, we excluded all SNPs and indels for which there were >25% missing data. Further filtering was applied to retain only 459 high confidence, species-specific markers that exist in alternate homozygous states between all genotyped accessions of the two parental species, including the additional seven accessions from sites other than Lime Kiln (Supplementary Table S1). For those 459 high confidence, species-specific markers, alleles specific to *B. thunbergii* are referred to as 'Bt' alleles and alleles specific to *B. vulgaris* are referred to as 'Bv' alleles.

Pedigree inferences for the 63 *B. ×ottawensis* accessions were made based on the percentage compositions of Bt homozygous loci, Bv homozygous loci, and Bv/Bt heterozygous loci within each individual. Because the pedigree analysis involves only species-specific markers, the theoretical expectation for true F_1 hybrids is 100% heterozygosity across all loci (i.e. 100% Bv/Bt). Due to both sampling bias and inherent genotyping error, however, 100% heterozygosity is not expected in empirical data; thus in this study, *B. ×ottawensis* individuals were considered F_1 hybrids if the observed percentage of Bv/Bt heterozygosity across all 459 high-confidence species-specific markers was >95% and the combined percentage of Bv

homozygous and Bt homozygous loci was <5%. More complex pedigrees were similarly inferred, under the simplifying assumption of independent segregation. A reference table of expected proportions of Bt homozygous, Bv homozygous, and Bv/Bt heterozygous loci for possible hybrid pedigrees, up to four generations, is provided in Supplementary Table S2.

Disease phenotyping

In April and May 2015, rooted cuttings of the 87 genotyped accessions were phenotyped for their responses to *Pg* at the USDA-ARS Cereal Disease Laboratory (CDL) in St. Paul, MN, using the standard protocol (Cotter, 1932) for testing barberry varieties for release to regulated states, as per CDL's contract with the USDA Animal and Plant Health Inspection Service. If scored as resistant at the CDL, the disease reactions of the 63 genotyped *B.* ×*ottawensis* individuals were verified via an independent round of testing at UNH in April and May 2016. For all disease phenotyping, newly emerged leaves were inoculated with germinated *Pg* basidiospores by suspending overwintered, telia-laden straw of naturally infected *Elymus repens* (L.) Gould over barberry plants in an inoculation chamber cycling between 18 °C (light, 14 h) and 16 °C (dark, 10 h). The *Pg*-infected stems of *E. repens* were collected in 2013 in St. Charles, MN, where a population of *Pg* is alternating between *B. vulgaris* and *E. repens* (Y. Jin, unpublished) and are part of CDL's source inoculum for barberry testing. In that region, *E. repens* is commonly infected by two *Puccinia* spp. (*P. graminis*, the causal organism of stem rust, and *P. coronata*, the causal organism of crown rust), but the two are easily distinguished on the basis of telia and teliospore morphologies (Cummins, 1971). Also, given the fact that the alternate host for *P. coronata* is *Rhamnus* spp. rather than *Berberis* spp., the natural inoculum used in this study was certainly *Pg*.

Two days post-inoculation, plants were moved to a growth chamber or greenhouse cycling between 20 ± 2 °C (light, 14 h) and 18 ± 2 °C (dark, 10 h) for further incubation. Infections were scored visually 14 d after inoculation, when mature aecia developed on the susceptible *B. vulgaris* check. Individual plants were scored as *Pg* susceptible if more than five pycnia were seen to develop on the upper surfaces of individual leaves and mature aecia were seen to develop on the lower surfaces. Individuals were scored as *Pg* resistant if the inoculation failed to produce visual symptoms, apart from minor flecking. Individuals were scored as inter- mediate if any of the following were observed: hypersensitive reactions, including chlorosis and/or necrosis; leaf reddening; or fewer than five pycnia with no associated aecia. If the susceptible checks failed to exhibit clear infection, all accessions in that group were subjected to an- other round of inoculation and rescored. Because of the inherent spatial heterogeneity of basidiospore ejection during inoculation, *Pg* susceptibility is called with higher confidence than *Pg* resistance. Therefore, all *B.* ×*ottawensis* individuals scored as resistant were screened a second time to reduce the chance of false negatives (Supplementary Table S3).

Testing maternal inheritance of* Pg-*NHR

To test whether or not the *Pg*-NHR of *B. thunbergii* observed segregating in the population of Lime Kiln hybrids is maternally inherited, an independent population of 129 F_1 *B.* ×*ottawensis* full sibs was developed via a controlled cross between *B. vulgaris* accession 'Wagon Hill' (female parent) and *B. thunbergii* accession 'UCONN1' (pollen parent). The hybrid status of all progeny was validated via GBS, and their reactions to *Pg* were evaluated at UNH in April and May 2016, following the method described above.

Results

Relative abundance and phenotypic characterization of Berberis *taxa at the study site*

The field survey conducted to characterize the natural populations of *B. vulgaris*, *B. thunbergii*, and *B. ×ottawensis* growing within the Lime Kiln Farm Wildlife Sanctuary in Sheffield, MA, revealed wide distribution of all three barberry taxa within the study area (Supplementary Fig. S1). The combined population size of the three taxa was estimated to be ~1000 individuals, spread over an area of ~7 ha. *Berberis thunbergii* (~600 individuals) was observed to outnumber the other two taxa, although *B. ×ottawensis* and *B. vulgaris* plants were also observed in significant numbers (~200 each). While morphological variation was observed within the populations of all three taxa, the most pronounced variation within a taxon was observed among the *B. ×ottawensis* hybrids, especially in the diagnostic characteristics of plant height (0.5–3 m tall), leaf size (2–6 cm long and 1–4 cm wide), and leaf margin (entire to >30 serrations). While some *B. ×ottawensis* individuals were as tall as or even taller than mature *B. vulgaris* plants, their leaves were often smaller than those of *B. vulgaris* and/or their margins had far fewer serrations. In contrast, other hybrids largely resembled *B. thunbergii*, in the sense of being relatively short (0.5–1 m tall) and having non-serrated leaves, but bore 5–12 bright yellow flowers on truncated pendant racemes typical of the nothospecies.

Variant calling

GBS generated a total of 381 million 150 bp PE reads for the 87 accessions selected at random from the study site. After quality parsing and demultiplexing, the average numbers of high-quality reads per genotype retained by the GBS-SNP-CROP pipeline were 2.8, 3.5, and 4.8 million PE reads for *B. thunbergii*, *B. vulgaris*, and *B. ×ottawensis*, respectively (Supplementary

Table S1). Using high-quality PE reads from all 63 *B.* ×*ottawensis* individuals, we generated a mock reference comprised of 143 213 centroids (i.e. consensus GBS fragments), with a total length of ~25 Mbp. In total, the pipeline identified 23091 putative variants, including 20799 SNPs (average depth DSNPs=26.71) and 2292 bi-allelic indels (Dindels=25.22), and the percentage of missing data was low (4%). After filtering, the final set of 2369 simplex markers (i.e. SNPs or indels located within centroids containing a single polymorphic site) exhibited higher depth (D=66.67) but a similar pattern of heterozygosity, homozygosity, and missing data characteristic of the entire data set (Table 1).

Table 1. Summary data characterizing the variants called for 63 accessions of *B.* ×*ottawensis* collected from the Lime Kiln Farm Wildlife Sanctuary in Sheffield, MA

Type[a]	N[b]	D[c]	%Het[d]	%Hom[e]	%NA[f]
All markers					
SNP	20,799	48.74	55.89	40.22	3.89
Indel	2,292	46.45	54.79	40.69	4.52
Both	23,091	48.51	55.78	40.26	3.96
Simplex markers only					
SNP	2,164	67.55	50.03	46.15	3.81
Indel	205	57.42	55.02	39.20	5.77
Both	2,369	66.67	50.46	45.55	3.98

[a] The type of variant called by the *de novo* GBS-SNP-CROP pipeline (v.2.1), either SNPs or indels
[b] The number of variants, by type, called after imposing all recommended genotyping criteria for diploid species
[c] The average read depth, by variant type
[d] The percentage of heterozygous loci throughout the population
[e] The percentage of homozygous loci throughout the population
[f] The percentage of missing cells (i.e. no genotype assigned for a given variant-accession combination)

Assessment of genetic diversity

Genetic diversity analyses were performed within and among the three populations of barberry taxa at Lime Kiln using the 2369 simplex markers described above (see Table 2). In terms of intraspecific genetic diversity, the percentage of polymorphic loci was nearly 100% within the *B.* ×*ottawensis* subpopulation, a result which reflects the hybrid nature of these

individuals and supports the high level of unbiased expected heterozygosity (*H*E=0.375) estimated for the nothospecies. The percentages of polymorphic loci were relatively lower for both of the parental species (21.7% for *B. vulgaris* and 41.0% for *B. thunbergii*), and the lowest value of *H*E was found for *B. vulgaris* (*H*E=0.088). Like this low value of *H*E, the highly negative value of *F*IS (–0.24) for *B. vulgaris* is unsurprising in light of the severe genetic bottleneck presumed for this species during its colonial introduction from Europe to North America.

Table 2. Population parameters characterizing the genetic diversity among and within the three sampled sub-populations of barberry taxa, based on 2,369 simplex markers

Taxon	N[a]	%POL[b]	H_O[c]	H_E[d]	F_{IS}[e]
B. thunbergii	9	40.95	0.152	0.153	-0.065
B. vulgaris	8	21.74	0.104	0.088	-0.243
B. ×ottawensis	63	99.96	0.462	0.375	-0.085
Means		54.18	0.239	0.205	-0.101

[a] The number of genotypes sampled, by species
[b] The percentage of polymorphic loci within each species
[c] The observed heterozygosity
[d] The unbiased expected heterozygosity
[e] The inbreeding coefficient

In general, the low level of inbreeding observed for all three species (average *F*IS= –0.131) is an understandable consequence of their outcrossing physiologies (Lebuhn and Anderson, 1994). The high value of the interspecific genetic structure between *B. vulgaris* and *B. thunbergii* (*F*ST=0.738) indicates a robust population structure in spite of co-location, a structure that is probably maintained due to flowering time differences between the parental species, with *B. thunbergii* flowering 2–4 weeks earlier than *B. vulgaris* in the region (Connolly *et al.*, 2013). Both the overall taxa-based population structure and the relative genetic diversity within taxa are

well captured by a PCA (Supplementary Fig. S2), in which the first two axes account for >89% of the genetic diversity.

Inferred pedigrees of the B. ×ottawensis *individuals*

From a total of 2369 simplex markers, 459 high-confidence, species-specific variants (i.e. SNPs or indels polymorphic between the two parental species but monomorphic within) were retained and used for pedigree analysis. Using this reduced set of parental species-informative markers, we inferred the generic pedigrees of the 63 genotyped *B.* ×*ottawensis* individuals based on the percentage compositions of Bt homozygous, Bv homozygous, and Bt/Bv heterozygous loci within each individual, where 'Bt' designates an allele specific to *B. thunbergii* and 'Bv' designates an allele specific to *B. vulgaris*. Of the 63 genotyped *B.* ×*ottawensis* individuals, 53 (84%) were found to be heterozygous (Bt/Bv) at ≥95% of the 459 loci; hence the vast majority of *B.* ×*ottawensis* individuals appear to be true, first-generation (F_1) hybrids at the study site (Table 3). Comparatively smaller numbers of *B.* ×*ottawensis* individuals appear to be later generation hybrids (e.g. backcrosses to parental species, etc.; see Table 3). In addition to the expected proportions of Bt homozygous, Bv homozygous, and Bt/Bv heterozygous loci for various possible pedigrees (Supplementary Table S2), the full data set with actual proportions and inferred pedigrees is presented in Supplementary Table S3.

Table 3. Inferred pedigrees and observed reaction types of the 63 genotyped *B.* ×*ottawensis* accessions, based on 459 species-specific markers

Inferred pedigree[a]	S[b]	I[c]	R[d]	F[e]
Bt/Bv = F_1	26	8	17	2
F_1/Bv	1	0	0	0
F_1/Bt//Bv	1	1	0	0
F_1/Bv//Bv	1	0	0	0
F_1/Bv//Bt	1	0	0	0
F_1/Bt//Bv///Bt	0	2	0	0
F_1/Bv//Bt///Bv	1	0	0	0
F_1/Bt//Bt///Bt////Bv	2	0	0	0
Total	33	11	17	2

[a] Accession pedigrees are inferred based on observed proportions of homozygous and heterozygous loci, considering only species-specific markers and assuming independent segregation (see S2 Table). Bt designates a *B. thunbergii* parent, Bv designates a *B. vulgaris* parent, and F_1 designates the *B.* ×*ottawensis* F_1 hybrid. Within each pedigree, one slash ("/") indicates the first cross, two slashes ("//") indicate the second cross, and so forth. For example, pedigree "A/B//C///D" indicates that A was first crossed with B, their offspring was crossed with C, and that offspring was crossed with D.

[b] The number of *Pg*-susceptible genotypes

[c] The number of intermediate genotypes

[d] The number of *Pg*-resistant genotypes

[e] The number of failed inoculations (i.e. no disease phenotype)

Reaction to Puccinia graminis *inoculation*

To determine disease responses to *Pg*, individual propagated accessions were inoculated using overwintered telia of *Pg* found on naturally infected *E. repens*. Of the 190 individuals collected from the field, inoculation was successful for 122 accessions, with the other 68 accessions dying during either shipment, propagation, or handling in the greenhouse. One week after inoculation, pycniospores began to develop on the upper surfaces of young leaves of both *B. vulgaris* and susceptible *B.* ×*ottawensis* individuals, whereas resistant *B.* ×*ottawensis* individuals showed varying responses, ranging from no visual symptoms (similar to *B. thunbergii*) to the development of sparse brown flecking. Two weeks after inoculation, accessions of *B. vulgaris* exhibited a clear susceptible reaction, with well-developed mature aecia

visible on the lower leaf surfaces (Fig. 3A). In contrast, accessions of *B. thunbergii* developed either no symptoms at all or small, sparse flecking (Fig. 3B), and accessions of *B.* ×*ottawensis* showed varying responses, ranging from *B. vulgaris*-like full susceptibility (Fig. 3C) to *B. thunbergii* like resistance (Fig. 3H). Various intermediate disease responses of some *B.* ×*ottawensis* accessions included aecial development in the presence of red necrotic islands (Fig. 3D), reddish or brown necrotic lesions with no aecial development (Fig. 3E, F), and flecking (Fig. 3G). As summarized in Table 4, all *B. vulgaris* accessions exhibited clear susceptible reactions and all *B. thunbergii* accessions exhibited clear resistant reactions. Of the 105 *B.* ×*ottawensis* accessions successfully tested for disease response to *Pg*, 54 (52%) exhibited clear susceptible reactions with well-developed pycnia and mature aecia, and 37 (35%) accessions exhibited clear resistant reactions with either no visual symptoms or sparse, small flecking. The remaining 14 accessions (13%) exhibited various intermediate responses, usually involving <5 pycnia, no aecial development, and associated hypersensitive-like reactions, including chlorosis, necrosis, and leaf reddening.

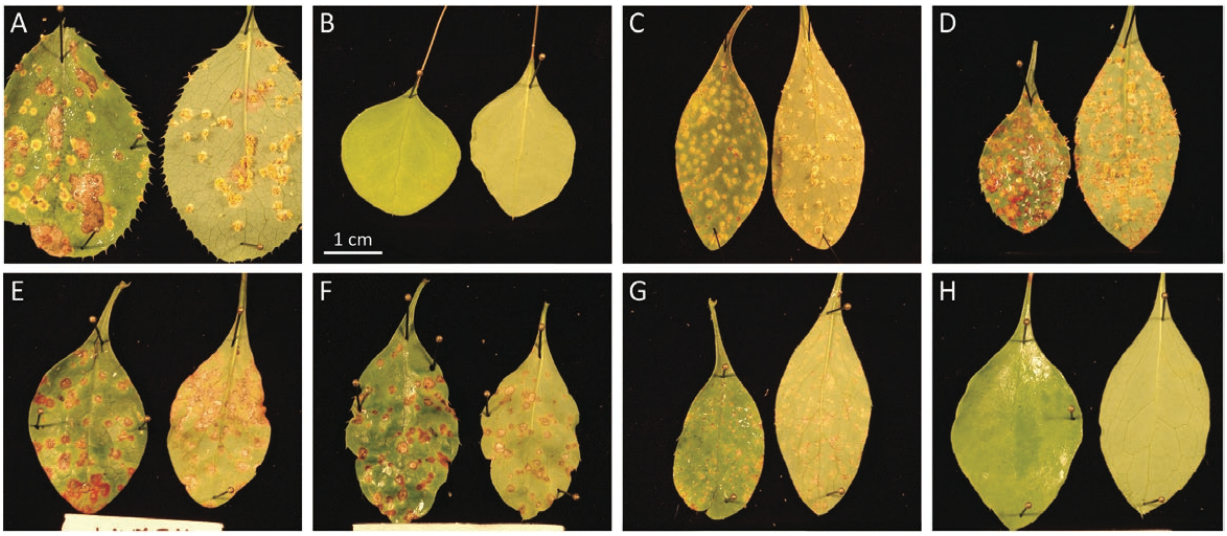


Fig. 3. Representative responses of *B. vulgaris*, *B. thunbergii*, and *B. ×ottawensis* accessions to *P. graminis* inoculation, using overwintered teliospores of naturally infected *E. repens* under controlled conditions. All photos were taken 14 d post-inoculation. (A) Susceptible reaction of *B. vulgaris* accession 'LK-070', showing pycnia on the upper surfaces of leaves and well-developed aecia on the lower surfaces of leaves. (B) Resistant reaction of *B. thunbergii* accession 'LK-107', showing no visual symptoms. (C, D) Susceptible reactions of *B. ×ottawensis* accessions 'LK-165' and 'LK-160', showing well-developed pycnia and aecia. (E–G) Intermediate reactions on *B. ×ottawensis* accessions 'LK-074', 'LK-121', and 'LK-137', showing sites of red or brown necrosis. (H) Resistant reaction of *B. ×ottawensis* accession 'LK-015', with no visual symptoms. White scale bar=1 cm.

The population of F_1 *B. ×ottawensis* full sibs derived from the cross between *B. vulgaris* accession 'Wagon Hill' and *B. thunbergii* accession 'UCONN1' was similarly tested for response to *Pg.* As within the natural population of *B. ×ottawensis* hybrids, varying responses ranging from apparent immunity to severe susceptibility were observed to segregate among the full sibs. Of the 129 full sib lines tested for disease response to *Pg*, 81 (63%) exhibited a clear susceptible reaction, 23 (18%) exhibited an intermediate response, and 25 (19%) exhibited a resistant reaction (Table 4).

Table 4. Summary of reactions to *P. graminis* by the 122 Lime Kiln accessions of three barberry taxa as well as the derived population of 129 *B.* ×*ottawensis* full sibs

	S[a]	I[b]	R[c]
Taxa collected from Lime Kiln			
B. thunbergii	0 (0%)	0 (0%)	13 (100%)
B. vulgaris	4 (100%)	0 (0%)	0 (0%)
B. ×*ottawensis*	54 (52%)	14 (13%)	37 (35%)
'Wagon Hill'/'UCONN1' full sibs			
B. ×*ottawensis*	83 (63%)	23 (18%)	25 (19%)

[a]The number and percentage of Pg-susceptible genotypes.

[b]The number and percentage of intermediate genotypes.

[c]The number and percentage of Pg-resistant genotypes.

Discussion

As the most common form of genetic resistance, NHR has the potential to provide broad-spectrum, durable resistance to many plant pathogens, including the causal organism of wheat stem rust. Unfortunately, the genetic mechanisms underlying *Pg*-NHR are poorly understood, in part due to the inherent challenge of developing a genetically tractable system in which genes controlling *Pg*-NHR segregate. In this study, we investigated the disease response of an interspecific hybrid between *Pg*-resistant *B. thunbergii* and *Pg*-susceptible *B. vulgaris*, and demonstrated the viability of this unique pathosystem to begin characterizing and mapping the gene(s) underlying the putative *Pg*-NHR of *B. thunbergii*.

The apparent non-host resistance of B. thunbergii *to* P. graminis *segregates in the interspecific hybrid* B. ×ottawensis

The natural population of *B.* ×*ottawensis* hybrids screened for disease response in this study was found to segregate for resistance to *Pg*. Specifically, 52% of the successfully screened hybrid accessions exhibited *B. vulgaris*-like susceptibility, 35% showed clear *B. thunbergii*-like

resistance, and the remaining 13% showed varying intermediate reactions. GBS analysis performed on a random subset of 63 of these phenotyped *B.* ×*ottawensis* accessions showed that 53 were true, first- generation (F_1) hybrids, among which a similar proportion of susceptible (51%), resistant (33%), and intermediate (16%) reactions were observed (Table 3). These results demonstrate that the *Pg*-NHR observed in *B. thunbergii* segregates in a population of first-generation (F_1) interspecific hybrids with *Pg*-susceptible *B. vulgaris*. Therefore, the gene(s) underlying *Pg*-NHR in *B. thunbergii* are, in theory, mappable in an F_1 population derived from the controlled hybridization of *Pg*- susceptible *B. vulgaris* and *Pg*-resistant *B. thunbergii*.

Over the past decade, in light of growing global concern about the wheat rusts, a number of efforts have been mounted to understand NHR to rust pathogens using various model and non-model plants. Many plant species, including *A. thaliana*, *Brachypodium distachyon*, rice, barley, and cowpea (Ayliffe *et al.*, 2011; Cheng *et al.*, 2012, 2013; An *et al.*, 2016; Dawson *et al.*, 2016; Li *et al.*, 2016; Zhao *et al.*, 2016), have been used to study NHR to *P. striiformis* f. sp. *tritici* (*Pst*), the causal organism of wheat stripe rust. In contrast, *Pg*-NHR has thus far been studied only in rice (Ayliffe *et al.*, 2011), as distinct from the studies of intermediate *Pg* resistance conducted in barley and *B. distachyon* (Figueroa *et al.*, 2013; Dracatos *et al.*, 2014). Given the evolutionary relationship between the *Berberis* genus and *Pg* prior to its host expansion to the grasses, the findings of this study suggest that the interspecific hybrid *B.* ×*ottawensis* may have value as an alternative, novel system for mapping and gaining insight into the genetic mechanism(s) underlying resistance to this complex pathogen.

Pg-*NHR in* B. thunbergii *probably involves more than one nuclear gene*

From the practical standpoint of breeding for improved resistance to biotic factors, the central questions regarding NHR concern the nature and modes of inheritance of the under- lying genes. In the case of maize NHR to the rice bacterial streak pathogen *Xanthomonas oryzae* pv. *oryzicola*, Zhao *et al.* (2004) reported that the resistance is mediated by a single gene, designated *Rxo1*. If a single gene governs *Pg*-NHR in *B. thunbergii*, the segregation of resistance among F_1 hybrids documented in this study suggests that the underlying resistance gene may exhibit dominance and exist in a heterozygous state within *B. thunbergii*. In such a case, however, independent assortment during meiosis would invariably result in homozygous *Pg*-susceptible *B. thunbergii* progeny. To date, no accession of *B. thunbergii* tested by the CDL has shown susceptibility to *Pg*; thus a single gene governing *Pg*-NHR in *B. thunbergii* is unlikely. The range of disease responses observed in this study, including a complete lack of visual symptoms (i.e. immunity), various intermediate-level reactions, and full susceptibility (Fig. 3), also suggests that the *Pg*-NHR of *B. thunbergii* is probably governed by more than a single gene.

Segregation of resistance within a natural population of F_1 hybrids could also be explained if the inheritance of NHR is non-nuclear. If the *Pg*-NHR of *B. thunbergii* is transmitted via the cytoplasm, all offspring should exhibit a disease response similar to that of the maternal plant. Under natural conditions, such as those at the Lime Kiln Farm Wildlife Sanctuary, the relatively restricted gene flow between *B. thunbergii* and *B. vulgaris* is assumed to be bi-directional, meaning both species have an equal chance to serve as the female parent of *B.* ×*ottawensis* hybrids. If *Pg*-NHR is cytoplasmically inherited, all hybrid progeny obtained from a *B. thunbergii* mother plant are expected to be resistant, and all hybrids obtained from a *B. vulgaris* mother plant are expected to be susceptible. In this study, we found that ~50% of the Lime Kiln F_1 hybrids exhibit a susceptible reaction, suggesting that the *Pg*-NHR of *B. thunbergii*

may indeed be cytoplasmically inherited. To test this hypothesis, a population of 129 F_1 *B.* ×*ottawensis* full sibs was developed via a controlled cross between *B. vulgaris* (female parent) and *B. thunbergii* (pollen parent) and screened for disease response to *Pg*. Under this scenario, all F_1 hybrids would be expected to exhibit *B. vulgaris*-like susceptibility to *Pg*; yet clear segregation in disease response was observed (Table 4), indicating that the *Pg*-NHR of *B. thunbergii* is not transmitted via the cytoplasm.

Pg-*resistant barberry species may be epidemiologically relevant to* Pg *evolution by virtue of their hybrid progeny*

It is well established that naturalized populations of *B. thunbergii* and *B. vulgaris* are widespread throughout New England (Connolly *et al.*, 2013; Hale *et al.*, 2015), to the extent that both are considered invasive species.This study shows that *B.* ×*ottawensis* is present throughout the Lime Kiln Farm Wildlife Sanctuary, and other recent studies report that this interspecific hybrid, assumed previously to be quite rare, is commonly found where the two parental species co-occur (Connolly *et al.*, 2013; Hale *et al.*, 2015). Once they are confirmed to be resistant to *Pg* by the USDA, *B. thunbergii* cultivars are propagated and sold as ornamental shrubs throughout the USA, as part of a multi-million dollar nationwide industry (Lubell *et al.*, 2008). While individual *B. thunbergii* genotypes may be deemed to pose no risk in terms of stem rust epidemiology, the results of this study raise a concern about the epidemiological risk of their progeny. Given the documented ability of *B. thunbergii* to naturalize and disperse, the prolific fruit set of many ornamental cultivars, the Herculean effort to purge the landscape of *Pg*-susceptible barberry plants in the 20th century, and the ongoing need to prevent sexual recombination of the stem rust pathogen, this study indicates a need to investigate and recon-

sider the epidemiological risk posed by *B. thunbergii*, by way of its interspecific hybrid with *B. vulgaris*. Specifically, in light of concerns around both invasiveness and *Pg* epidemiology, perhaps the minimum standard for new ornamental cultivars of *B. thunbergii* should be sterility, as pioneered by the recently patented cultivars 'UCONNBTCP4N', 'UCONNBTB039', 'UCONNBTB048', 'UCONNBTB113', and 'NCBT1' Sunjoy Mini MaroonTM.

Beyond the USA, the highly diverse *Berberis* genus is distributed nearly worldwide, with centers of diversity in southern Asia as well as Central and South America (Ahrendt, 1961; Rounsaville and Ranney, 2010), and evidence of the alternate host's role in current rust epidemics is mounting. In China, for example, ~250 *Berberis* spp. are found, accounting for nearly 50% of the species recorded globally (Ying and Chen 2001); and the sexual recombination of *Pst* observed on barberries has been implicated in the high genetic diversity of *Pst* in that country (Lu *et al.*, 2009; Mboup *et al.*, 2009; Sharma-Poudyal *et al.*, 2013; Zheng *et al.*, 2013; Wang *et al.*, 2016). Similarly, *Pg*-compatible *B. holstii* growing near wheat production areas in the highlands of eastern Africa may have played a role in the emergence of new virulence combinations in that region, including the rapidly diversifying Ug99 family of races (Singh *et al.*, 2015; Zhang *et al.*, 2017). In Iran, a country where the fruit of *B. vulgaris* is produced commercially on ~11 000 ha (Rahimi-Madiseh *et al.*, 2017), *Pg* races of highly diverse virulence profiles were recovered from aecial samples collected from infected *B. vulgaris* plants, indicating that *Pg* can complete its sexual cycle in the region (Hansen *et al.*, 2013). Indeed, barberries grow widely throughout the mountainous areas of Central West Asia and North Africa, including countries in the Ug99 pathway, yet their exact epidemiological relevance remains unclear. While the present study focuses only on the two barberry species found in New England, one imported from Europe and the other from Japan, their natural hybrid and its

complex response to *Pg* raises questions about the epidemiological role of presumably rust-resistant *Berberis* spp. worldwide. In short, the existence of interspecific hybrids may bring additional complications to the important work of understanding the contribution of *Berberis* spp. to global rust cycles.

Future work

To build on the results of this study, we are developing an F bi-parental mapping population via a controlled cross between *B. vulgaris* accession 'Wagon Hill' and *B. thunbergii* accession 'UCONN1' in order to develop linkage maps for both parental species and begin mapping the gene(s) underlying *Pg*- NHR in *B. thunbergii*. To facilitate downstream dissection of potential quantitative trait loci, a reference genome for *B. thunbergii* is also in development. While it is possible, even likely, that the mechanisms governing *Pg*-NHR at the basidiospore stage (i.e. *Pg*'s infection of its sexual host) may not be relevant to those governing *Pg*-NHR at the aeciospore or urediospore stages (i.e. the spores which infect *Pg*'s asexual hosts, including wheat and other small grains), the ongoing, centuries-long fight against this complex and historic pathogen demands that all strategies be pursued. Ultimately, the hope is that characterization of *Pg*-NHR in *B. thunbergii*, a species closely related to the pathogen's ancestral host, may provide information about the evolution of modern day heteroecious *Pg* and contribute insight into possible mechanisms of durable resistance in wheat.

Supplementary data

Supplementary data are available at *JXB* online. https://doi.org/10.1093/jxb/ery066

Fig. S1. Map of the Lime Kiln Farm Wildlife Sanctuary in Sheffield, MA.

Fig. S2. Principal components analysis showing the genetic structure of the three barberry subpopulations at the Lime Kiln Farm Wildlife Sanctuary.

Table S1. The 87 *Berberis* accessions used in the study, with passport information, genotypic data, and phenotypic features.

Table S2. Expected proportions of homozygous and heterozygous loci of hybrid accessions of various potential pedigrees.

Table S3. Disease reactions and observed proportions of homozygous and heterozygous loci within each of the 63 genotyped hybrid accessions.

Text S1. Detailed record of the GBS-SNP-CROP command lines used in this study, including all specified pipeline parameters.

Acknowledgements

We thank A. Chevoor for field support; J. Ebba and L. Hydock for greenhouse support; M. Lim and S. Gale for disease screening support; H. Gustafson for laboratory support; M. Brand for guidance on plant propagation and management; and Mass Audubon for permission to conduct research at the Lime Kiln Farm Wildlife Sanctuary. We also thank the two reviewers whose comments strengthened this manuscript. This work was supported by the Bill & Melinda Gates Foundation [OPPGD1389, OPP1133199]; and USDA NIFA Hatch Multistate Project NE009 [NH00611-R].

References

Ahrendt LWA. 1961. Berberis and Mahonia: a taxonomic revision. Botanical Journal of the Linnean Society **57,** 1–410.

An T, Cai Y, Zhao S, Zhou J, Song B, Bux H, Qi X. 2016. *Brachypodium distachyon* T-DNA insertion lines: a model pathosystem to study nonhost resistance to wheat stripe rust. Scientific Reports **6,** 25510.

Atienza SG, Jafary H, Niks RE. 2004. Accumulation of genes for susceptibility to rust fungi for which barley is nearly a nonhost results in two barley lines with extreme multiple susceptibility. Planta **220,** 71–79.

Ayliffe M, Devilla R, Mago R, White R, Talbot M, Pryor A, Leung H. 2011. Nonhost resistance of rice to rust pathogens. Molecular Plant- Microbe Interactions **24,** 1143–1155.

Ayliffe M, Singh R, Lagudah E. 2008. Durable resistance to wheat stem rust needed. Current Opinion in Plant Biology **11,** 187–192.

Borlaug NE. 2000. Ending world hunger. The promise of biotechnology and the threat of antiscience zealotry. Plant Physiology **124,** 487–490.

Cheng Y, Zhang H, Yao J, Han Q, Wang X, Huang L, Kang Z. 2013. Cytological and molecular characterization of non-host resistance in *Arabidopsis thaliana* against wheat stripe rust. Plant Physiology and Biochemistry **62,** 11–18.

Cheng Y, Zhang H, Yao J, Wang X, Xu J, Han Q, Wei G, Huang L, Kang Z. 2012. Characterization of non-host resistance in broad bean to the wheat stripe rust pathogen. BMC Plant Biology **12,** 96.

Christensen CM. 1984. E.C. Stakman, statesman of science. St. Paul, MN: American Phytopathological Society.

Connolly BA, Anderson GJ, Brand MH. 2013. Occurrence and fertility of feral hybrid barberry *Berberis* ×*ottawensis* (Berberidaceae) in Connecticut and Massachusetts. Rhodora **115,** 121–132.

Cotter RU. 1932. Factors affecting the development of the aerial stage of *Puccinia graminis*. US Department of Agriculture, No. 314.

Cummins GB. 1971. The rust fungi of cereals, grasses and bamboos. Heidelberg, Berlin: Springer-Verlag, 570.

Dawson AM, Ferguson JN, Gardiner M, Green P, Hubbard A, Moscou MJ. 2016. Isolation and fine mapping of *Rps6*: an intermediate host resistance gene in barley to wheat stripe rust. Theoretical and Applied Genetics **129,** 831–843.

Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19,** 11–15.

Dracatos PM, Ayliffe M, Khatkar M, Fetch T, Singh D, Park R. 2014. Inheritance of pre-haustorial resistance to *Puccinia graminis* f. sp. *avenae* in barley (*Hordeum vulgare* L.). Molecular Plant-Microbe Interaction **27,** 1253–1262.

Figueroa M, Alderman S, Garvin DF, Pfender WF. 2013. Infection of *Brachypodium distachyon* by formae speciales of *Puccinia graminis*: early infection events and host–pathogen incompatibility. PLoS One **8,** e56857.

Gleason HA, Cronquist A. 1963. Berberidaceae. In: Manual of vascular plants of northeastern United States and adjacent Canada, 2nd edn. Bronx, NY: The New York Botanical Garden, 64.

Grieve M. 1931. A modern herbal: the medicinal, culinary, cosmetic, and economic properties, cultivation, and folk-lore of herbs, grasses, fungi, shrubs, and trees with all their modern scientific uses. New York: Harcourt, Brace, and Company.

Haines A. 2011. *Flora Novae Angliae*: a manual for the identification of native and naturalized higher vascular plants of New England. New Haven, CT: New England Wild Flower Society and Yale University Press.

Hale IL, Connolly BA, Bartaula R. 2015. The occurrence of hybrid barberry, *Berberis* ×*ottawensis* (Berberidaceae), in New Hampshire and Rhode Island. Rhodora **117,** 384–387.

Hansen JG, Lassen P, Justesen AF, Nazari K, Hodson D, Hovmoller M. 2013. Barberry rust survey. Developing tools for data management and dissemination. Global Rust Reference Center.

Jafary H, Albertazzi G, Marcel TC, Niks RE. 2008. High diversity of genes for nonhost resistance of barley to heterologous rust fungi. Genetics **178,** 2327–2339.

Jin Y. 2011. Role of *Berberis* spp. as alternate hosts in generating new races of *Puccinia graminis* and *P. striiformis*. Euphytica **179,** 105–108.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics **27,** 3070–3071.

Lebuhn G, Anderson GJ. 1994. Anther tripping and pollen dispensing in *Berberis thunbergii*. American Midland Naturalist 257–265.

Leonard KJ, Szabo LJ. 2005. Stem rust of small grains and grasses caused by *Puccinia graminis*. Molecular Plant Pathology **6,** 99–111.

Leppik EE. 1961. Some viewpoints on the phylogeny of rust fungi. IV. Stem rust genealogy. Mycologia **53,** 378–405.

Levine MN, Cotter RU. 1932. Susceptibility and resistance of *Berberis* and related genera to *Puccinia graminis*. United States Department of Agriculture, Economic Research Service. Available at: https://ideas.repec. org/p/ags/uerstb/163218.html. [Accessed 7 March 2018].

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; **1000 Genome Project Data Processing Subgroup**. 2009. The sequence alignment/map format and SAMtools. Bioinformatics **25,** 2078–2079.

Li K, Hegarty J, Zhang C, *et al***.** 2016. Fine mapping of barley locus *Rps6* conferring resistance to wheat stripe rust. Theoretical and Applied Genetics **129,** 845–859.

Lipka U, Fuchs R, Kuhns C, Petutschnig E, Lipka V. 2010. Live and let die—*Arabidopsis* nonhost

resistance to powdery mildews. European Journal of Cell Biology **89,** 194–199.

Lu N, Zhan G, Wang J, Huang L, Kang Z. 2009. Molecular evidence of somatic genetic recombination of *Puccinia striiformis* f. sp. *tritici* in China. Acta Phytopathologica Sinica **39,** 561–568.

Lubell JD, Brand MH, Lehrer JM, Holsinger KE. 2008. Detecting the influence of ornamental *Berberis thunbergii* var. *atropurpurea* in invasive populations of *Berberis thunbergii* (Berberidaceae) using AFLP1. American Journal of Botany **95,** 700–705.

Mboup M, Leconte M, Gautier A, Wan AM, Chen W, de Vallavieille- Pope C, Enjalbert J. 2009. Evidence of genetic recombination in wheat yellow rust populations of a Chinese oversummering area. Fungal Genetics and Biology **46,** 299–307.

Mehrhoff LJ, Silander JA Jr, Leicht SA, Mosher ES, Tabak NM. 2003. IPANE: invasive plant atlas of New England. Storrs, CT: University of Connecticut Press.

Mellersh DG, Heath MC. 2003. An investigation into the involvement of defense signaling pathways in components of the nonhost resistance of *Arabidopsis thaliana* to rust fungi also reveals a model system for studying rust fungal compatibility. Molecular Plant-Microbe Interactions **16,** 398–404.

Melo AT, Bartaula R, Hale I. 2016. GBS-SNP-CROP: a reference- optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. BMC Bioinformatics **17,** 29.

Mysore KS, Ryu CM. 2004. Nonhost resistance: how much do we know? Trends in Plant Science **9,** 97–104.

Peakall ROD, Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes **6,** 288–295.

Perumalla CJ, Health MC. 1989. Effect of callose inhibition on haustorium formation by the cowpea rust fungus in the non-host, bean plant. Physiological and Molecular Plant Pathology **35,** 375–382.

Peterson PD. 2003. The common barberry: the past and present situation in Minnesota and the risk of wheat stem rust epidemics. PhD thesis, North Carolina State University.

Peterson PD. 2013. 'The Barberry or Bread': the public campaign to eradicate common barberry in the United States in the early 20th century. APS Features. doi:10.1094/APSFeature-2013-08.

Poland JA, Brown PJ, Sorrells ME, Jannink JL. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One **7,** e32253.

Prats E, Martínez F, Rojas-Molina MM, Rubiales D. 2007. Differential effects of phenylalanine ammonia lyase, cinnamyl alcohol dehydrogenase, and energetic metabolism inhibition on resistance of appropriate host and nonhost cereal–rust interactions. Phytopathology **97,** 1578–1583.

Pretorius ZA, Singh RP, Wagoire WW, Payne TS. 2000. Detection of virulence to wheat stem rust resistance gene Sr31 in *Puccinia graminis*. f. sp. *tritici* in Uganda. Plant Disease **84,** 203–203.

Rahimi-Madiseh M, Lorigoini Z, Zamani-gharaghoshi H, Rafieian- kopaei M. 2017. *Berberis vulgaris*: specifications and traditional uses. Iranian Journal of Basic Medical Sciences **20,** 569.

Rehder A. 1949. Berberidaceae. In: Manual of cultivated trees and shrubs. New York: MacMillan Co, 246.

Rounsaville TJ, Ranney TG. 2010. Ploidy levels and genome sizes of *Berberis* L. and *Mahonia* Nutt. species, hybrids, and cultivars. HortScience **45,** 1029–1033.

Shafiei R, Hang C, Kang JG, Loake GJ. 2007. Identification of loci controlling non-host disease resistance in *Arabidopsis* against the leaf rust pathogen *Puccinia triticina*. Molecular Plant Pathology **8,** 773–784.

Sharma-Poudyal D, Chen XM, Wan AM, *et al***.** 2013. Virulence characterization of international collections of the wheat stripe rust pathogen, *Puccinia striiformis* f. sp. *tritici.* Plant Disease **97,** 379–386.

Silander JA, Klepeis DM. 1999. The invasion ecology of Japanese barberry (*Berberis thunbergii*) in the New England landscape. Biological Invasions **1,** 189–201.

Singh RP, Hodson DP, Jin Y, *et al***.** 2015. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. Phytopathology **105,** 872–884.

Steffey J. 1985. Strange relatives: the barberry family. American Horticulturist **64,** 4–9.

Stokstad E. 2007. Deadly wheat fungus threatens world's breadbaskets. Science **315,** 1786–1787.

Thordal-Christensen H. 2003. Fresh insights into processes of nonhost resistance. Current Opinion in Plant Biology **6,** 351–357.

Van Splinter JL, Burgess MB, Spada DM, Werier D. 2016. *Berberis* ×*ottawensis* (Berberidaceae): a new addition to the flora of New York. Rhodora **118,** 412–414.

Wahl I, Anikster Y, Manisterski J, Segal A. 1984. Evolution at the center of origin. In: Bushnell WR, Roelfs AP, eds. The cereal rusts. Volume I. Origins, specificity, structure, and physiology. Orlando: Academic Press, Inc., 39–77.

Wang Z, Zhao J, Chen X, Peng Y, Ji J, Zhao S, Lv Y, Huang L, Kang Z.

2016. Virulence variations of *Puccinia striiformis* f. sp. *tritici* isolates collected from *Berberis* spp. in China. Plant Disease **100,** 131–138.

Ying JS, Chen DZ. 2001. Lardizabalaceae and Berberidaceae in flora of China. Beijing: Chinese Academy of Sciences, 50–214.

Zadoks JC, Bouwman JJ. 1985. Epidemiology in Europe. In: Roelfs AP, Bushnell WR, eds. The cereal rusts. Volume II. Diseases, distribution, epidemiology, and control. Orlando: Academic Press, 329–369.

Zhang H, Wang C, Cheng Y, *et al***.** 2011. Histological and molecular studies of the non-host interaction between wheat and *Uromyces fabae*. Planta **234,** 979–991.

Zhang W, Chen S, Abate Z, Nirmala J, Rouse MN, Dubcovsky J. 2017. Identification and characterization of *Sr13*, a tetraploid wheat gene that confers resistance to the Ug99 stem rust race group. Proceedings of the National Academy of Sciences, USA **114,** E9483–E9492.

Zhao BY, Ardales E, Brasset E, Claflin LE, Leach JE, Hulbert SH. 2004. The *Rxo1*/*Rba1* locus of maize controls resistance reactions to pathogenic and non-host bacteria. Theoretical and Applied Genetics **109,** 71–79.

Zhao B, Lin X, Poland J, Trick H, Leach J, Hulbert S. 2005. A maize resistance gene functions against bacterial streak disease in rice. Proceedings of the National Academy of Sciences, USA **102,** 15383–15388.

Zhao J, Wang L, Wang Z, *et al***.** 2013. Identification of eighteen *Berberis* species as alternate hosts of *Puccinia striiformis* f. sp. *tritici* and virulence variation in the pathogen isolates from natural infection of barberry plants in China. Phytopathology **103,** 927–934.

Zhao J, Yang Y, Yang D, *et al***.** 2016. Characterization and genetic analysis of rice mutant *crr1* exhibiting compromised non-host resistance to *Puccinia striiformis* f. sp. *tritici* (Pst). Frontiers in Plant Science **7,** 1822.

Zheng W, Huang L, Huang J, *et al***.** 2013. High genome heterozygosity and endemic genetic recombination in the wheat stripe rust fungus. Nature Communications **4,** 2673.

AFTERWORD

Non-host resistance (NHR) is defined as resistance exhibited by all individuals of a plant species to all genetic variants of a non-adapted pathogen [1]. Generally, a true non-host species would show complete immunity (i.e. no signs of colonization by a non-adapted pathogen), and the pathogens' inability to penetrate host tissue or form infection structures such as haustoria would be requirements for describing a plant as non-host to a pathogen [2]. In practice, however, phenotypic responses to infection by non-adapted rust pathogens range from immunity (lack of visual symptoms) to formation of visual pustules, depending on whether the NHR response arises from a basic incompatibility, in which the pathogen may be physically incapable of penetrating the host, or some active recognition and defense response, wherein the pathogen is capable of entering the host tissue but may be arrested at some stage in its development.

Some of the examples that have been collectively described as NHR response to cereal rusts in the literature include, but are not limited to, basic incompatible reactions in which the pathogen is incapable of infecting the host (e.g. flax rust on rice) [3]; pathogen penetration into host tissue followed by an inability to form haustoria (e.g. wheat stripe rust on *Arabidopsis thaliana* and bean, bean rust on wheat) [4,5]; infection in which all the fungal structures necessary for parasitism including haustoria are produced but sporulation never occurs (e.g. wheat stem rust, leaf rust, stripe rust, and barley brown rust on rice) [3]; formation of occasional tiny sporulating pustules (e.g. rusts of *Triticeae* hosts on *Brachypodium* spp.) [6]; and "near NHR," in which the majority of individuals of a potential host species exhibit immunity but a few exhibit moderate susceptibility to a non-adapted pathogen (e.g. wheat leaf rust on barley) [7]. For the cereal rusts, then, determination of whether a species is a host or non-host is complicated by the complex nature of phenotypic responses observed upon pathogen infection.

As there are no clear boundaries delimiting host and non-host species in case of cereal rusts, host resistance and NHR are most commonly differentiated based on pathogen adaptation to a particular species (host) and lack of adaptation to other species (nonhost) [8]. Despite their classification into two distinct categories of resistance, however, a growing understanding of plant immune responses suggests that host and non-host resistance may share some similarities in their underlying molecular mechanisms. There are several examples that indicate NHR consists of defense mechanisms similar to those utilized against adapted pathogens, including pathogen-associated molecular pattern (PAMP)-triggered and effector-triggered immunities [9–11]. Furthermore, both host and non-host resistance have been found to be associated with cellular responses such as hypersensitive reactions, callose deposition, production of reactive oxygen species, phytoalexin synthesis, and salicylic acid signaling [8]. Despite the fact that there is overlap between the two types of resistance, NHR is thought to be relatively durable due to the presumed complexity of its underlying mechanisms [12]; however, some studies suggest that apparent immunity may be conferred by few genes [13]. While the underlying concept of what comprises host and non-host resistance remains controversial, we speculate that with the aid of ever-improving molecular biology and genomics tools, future studies will shed light on the relationship between host and non-host resistance.

Complicating matters further, deciding whether or not two plant populations are the same species is itself non-trivial, and the subject continues to be intensely debated among taxonomists and evolutionary biologists. This debate bears directly on the work of the previous chapter because the question can be asked whether or not the two *Berberis* spp. used in this study (*B. vulgaris* and *B. thunbergii*) are actually distinct species. By extension, such a line of thinking then calls into question whether or not the resistance observed in *B. thunbergii* against *Pg* can

justifiably be referred to as NHR, as opposed to just basal resistance. According to the biological speciation concept (BSC), two taxa are defined as separate species by their inability to intercross with each other to produce fertile hybrid progeny [14]. Because *B. thunbergii* and *B. vulgaris* hybridize to produce the viable interspecific hybrid barberry *Berberis* ×*ottawensis*, one may claim on the grounds of BSC that these two taxa are not, therefore, separate species. In this afterword, rationale is provided to defend the claim that *B. thunbergii* and *B. vulgaris* are in fact two separate species despite such gene flow and that justification exists to call the resistance observed in *B. thunbergii* against *Pg* as NHR.

Although BSC is the oldest speciation concept and has been widely used to demarcate species in the animal kingdom, plant systematists have largely abandoned this concept as the sole determinant of speciation [15]. Unlike in the animal kingdom, where infertility can be easily determined by an inability to intercross, tests of infertility cannot be applied unambiguously in plants because intercrossing success in plants ranges from 0-100% and species assignment in the case of intermediate levels of infertility is ambiguous. Therefore, within the plant kingdom, unless hybridization between two taxa is so pervasive that the two effectively merge into a common genepool, the mere occurrence of some level of gene flow between taxa is insufficient grounds to declare the two as a single species [16]. Thus, in contrast to the rather binary view outlined by the BSC, species within plant kingdom are defined based on a range considerations, including morphological, ecological, geographical, reproductive difference, and other evidence of well-marked lineages [16].

Berberis vulgaris and *B. thunbergii* exhibit distinctly different morphological traits, including overall plant architecture, height, leaf shape and size, and inflorescence structure (discussed in detail in the previous chapter) [17]. The interspecific hybrid *B.* ×*ottawensis*,

typical to hybrid plants, exhibits a range of morphologies clearly intermediate between the two parental species. In addition, the two parental taxa, prior to their human-mediated introduction to North America, were geographically isolated from one another, with *B. vulgaris* being native to Asia and *B. thunbergii* to Japan. The co-occurrence of these species in North America began only recently, following the introduction of *B. vulgaris* 17th century European settlers and the introduction of *B. thunbergii* as an ornamental plant from Japan in 1875 [18]. The history of the hybrid *B. ×ottawensis* is even more recent. *Berberis ×ottawensis* was first created as an experimental hybrid in 1894 by hand-pollinating *B. vulgaris* with pollen from *B. thunbergii* [19]. Nearly 130 years later, the *B. ×ottawensis* remains an uncommon hybrid [20], with the occurrence of natural hybrids reported to be very rare until recently. The recent occurrence of the natural hybrid is further supported by the identification of 84% of the individuals in a natural stand as being true first generation hybrids in my study [21].

Besides being morphologically different and geographically isolated until very recently, *B. vulgaris* and *B. thunbergii* are also reproductively isolated, even when growing sympatrically, due to difference in flowering time. Here in the northeast, *B. thunbergii* flowers 2-4 weeks earlier than *B. vulgaris*, resulting in a very small window of flowering overlap for intercrossing [18]. In effect, this differential adaptation provides a reproductive behavior that severely restricts gene flow between the two species.

As opposed to the concept of speciation as is commonly applied to the animal kingdom, there is no clear consensus about the definition of species in plants. Based on overall differences in morphology and geographical provenance, as well as effective reproductive isolation even between sympatric populations, I assert that multiple lines of primary evidence exist to justify the classification of *B. thunbergii* and *B. vulgaris* as two distinct species. This conclusion, in

combination with the results of nearly a century of testing at the USDA-ARS Cereal Disease Laboratory (CDL) in St. Paul, MN on diverse *B. thunbergii* accessions for their response to *Pg* and never observing infection [22], provides the rationale for referring to the resistance to *Pg* observed in *B. thunbergii* as NHR. Some caution is warranted, however, when claiming *B. thunbergii* to be a definitive non-host to *Pg*. Despite the extensive testing that has taken place, not all individuals of *B. thunbergii* have been screened; nor have all known isolates of *Pg* been used in such tests. Specifically, no *B. thunbergii* accession has yet been tested for its response to Ug99 due to the practical difficulty of conducting such a test under current international quarantine regulations. All that being said, the course of ongoing evolution may one day erase this boundary. Although co-located only a little over a century ago, the rare hybridization events observed today may signal the beginning of a slow erosion of the boundary between these two species. Indeed, as intermediate genotypes between the two parental species, *B. ×ottawensis* individuals may play a key bridging role in that fusion, increasing facilitating interspecific gene flow to the extent that their status as separate species may need to be re-evaluated. In addition, some words of caution are necessary with regard that *B. thunbergii* have not been tested for their response to Ug99

In the end, the fact is that the definitions of both species and NHR remain actively debated; thus our claim, however reasoned, that the resistance observed in *B. thunbergii* against *Pg* is NHR will be considered controversial by some researchers. Taking a more practical perspective, in terms of breeding for resistance, the salient question is not whether the resistance exhibited by *B. thunbergii* is NHR or not; the question is whether or not that resistance, whatever its mechanism or label, is *durable*. Durability of resistance is defined as the ability of widely deployed resistant gene to provide an economic level of protection over an extended period of

time [23]. To date, no accession of *B. thunbergii* has shown susceptibility to *Pg*, despite both intensive screening [22] and the proliferation of the species in the landscape (Y. Jin, personal communication); thus the resistance carried by *B. thunbergii* meets the basic definition of durability. In many ways, the question of host or non-host is immaterial. The practical value of the study as outlined in the previous chapter lies in its examination of the *Pg* resistance present in *B. thunbergii* and in what insight it may ultimately provide regarding *Pg* evolution and possible mechanisms of durable resistance to this historic pathogen.

References

1. Lipka U, Fuchs R, Kuhns C, Petutschnig E, Lipka V. Live and let die–*Arabidopsis* nonhost resistance to powdery mildews. Eur J Cell Biol. 2010;89:194–199.

2. Bettgenhaeuser J, Gilbert B, Ayliffe M, Moscou MJ. Nonhost resistance to rust pathogens–a continuation of continua. Front Plant Sci. 2014;5:664.

3. Ayliffe M, Devilla R, Mago R, White R, Talbot M, Pryor A, et al. Nonhost resistance of rice to rust pathogens. Mol Plant Microbe Interact. 2011;24:1143–1155.

4. Cheng Y, Zhang H, Yao J, Han Q, Wang X, Huang L, et al. Cytological and molecular characterization of non-host resistance in *Arabidopsis thaliana* against wheat stripe rust. Plant Physiol Biochem. 2013;62:11–18.

5. Cheng Y, Zhang H, Yao J, Wang X, Xu J, Han Q, et al. Characterization of non-host resistance in broad bean to the wheat stripe rust pathogen. BMC Plant Biol. 2012;12:96.

6. Figueroa M, Alderman S, Garvin DF, Pfender WF. Infection of *Brachypodium distachyon* by formae speciales of *Puccinia graminis*: early infection events and host-pathogen incompatibility. PLOS One. 2013;8:e56857.

7. Jafary H, Albertazzi G, Marcel TC, Niks RE. High diversity of genes for nonhost resistance of barley to heterologous rust fungi. Genetics. 2008;178:2327–2339.

8. Gill US, Lee S, Mysore KS. Host versus nonhost resistance: distinct wars with similar arsenals. Phytopathology. 2015;105:580–587.

9. Lenk A, Thordal-Christensen H. From nonhost resistance to lesion-mimic mutants: useful for studies of defense signaling. Adv Bot Res. 2009;51:91–121.

10. Niks RE, Marcel TC. Nonhost and basal resistance: how to explain specificity? New Phytol. 2009;182:817–828.

11. Schulze-Lefert P, Panstruga R. A molecular evolutionary concept connecting nonhost resistance, pathogen host range, and pathogen speciation. Trends Plant Sci. 2011;16:117–125.

12. Thordal-Christensen H. Fresh insights into processes of nonhost resistance. Curr Opin Plant Biol. 2003;6:351–357.

13. Zhao BY, Ardales E, Brasset E, Claflin LE, Leach JE, Hulbert SH. *The Rxo1/Rba1* locus of maize controls resistance reactions to pathogenic and non-host bacteria. Theor Appl Genet. 2004;109:71–79.

14. Mayr E. Populations, species, and evolution: an abridgment of animal species and evolution. Harvard University Press; 1970.

15. McDade LA. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. Evolution. 1992;46:1329–1346.

16. Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ. Plant systematics. Sinauer Sunderland; 2002.

17. Bartaula R, Melo AT, Connolly BA, Jin Y, Hale I. An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen *Puccinia graminis*. J Exp Bot. 2018;69:2483–2493.

18. Connolly BA, Anderson GJ, Brand MH. Occurrence and fertility of feral hybrid barberry *Berberis ×ottawensis* (Berberidaceae) in Connecticut and Massachusetts. Rhodora. 2013;115:121–132.

19. Saunders CE. Notes on some variations in the second generation of *Berberis* hybrids. Proc Int Conf Plant Breed Hybrid. 1902.

20. Haines A. Flora Novae Angliae: a manual for the identification of native and naturalized higher vascular plants of New England. N Engl Wild Flower Soc Yale Univ Press N Hav CT. 2011.

21. Bartaula R, Melo AT, Connolly BA, Jin Y, Hale I. An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen *Puccinia graminis*. J Exp Bot. 2018;

22. Department of Agriculture. Black Stem Rust; Identification Requirements and Addition of Rust Resistant Varieties. Federal Register, No. 36. 67.

23. Johnson R. Durable Resistance: definition of, genetic control, and attainment in plant breeding. Phytopathology. 1981;71:567–568.

CHAPTER II

MAPPING NON-HOST RESISTANCE TO THE STEM RUST PATHOGEN IN AN INTERSPECIFIC BARBERRY HYBRID

Radhika Bartaula[1*], Arthur T. O. Melo[2], Sarah Kingan[3], Yue Jin[4], Iago Hale[2*]

[1]Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA;
[2]Department of Agriculture, Nutrition, and Food Systems, University of New Hampshire, Durham, NH 03824, USA;
[3]Pacific Biosciences, Menlo Park, CA 94025, USA;
[4]USDA-ARS Cereal Disease Laboratory, St. Paul, MN 55108

[*]Correspondence:

Iago Hale: iago.hale@unh.edu,
Radhika Bartaula: rb11@wildcats.unh.edu

Submission Pending for *Genome Biology*

ABSTRACT

Non-host resistance (NHR) presents a compelling strategy for achieving long-term plant protection for global food security, yet the genetic basis of NHR remains poorly understood. For many diseases, including wheat stem rust [causal organism *Puccinia graminis* (*Pg*)], an inherent challenge in studying the mechanism of NHR lies in the difficulty of developing a genetically tractable system within which NHR segregates. The present study turns to the pathogen's alternate host, barberry (*Berberis* spp.), to overcome this challenge. An interspecific F_1 bi-parental mapping population was developed via a controlled cross between *Pg*-susceptible European barberry (*B. vulgaris*) and *Pg*-resistant Japanese barberry (*B. thunbergii*) to dissect the genetic mechanism of the apparent *Pg*-NHR exhibited by *B. thunbergii*. Using genotyping-by-sequencing (GBS) data, the first genetic linkage maps were constructed for the two parental species; and QTL analysis of *Pg*-NHR resulted in a single QTL on the short arm of chromosome 3 of *B. thunbergii*. To gain further insight into this QTL region, dubbed *QPgr-3S*, a chromosome-level 1.2 Gb draft genome for *B. thunbergii* was assembled using long PacBio reads and chromosome conformation capture data. By anchoring the *B. thunbergii* linkage map to the draft genome and using a 189.3 Mbp transcriptome, the 13 cM *QPgr-3S* region was found to contain 10 contigs, ~3.4 Mbp of sequence, and 99 high-confidence genes. To narrow this list to candidate genes of high priority for subsequent investigation, differential gene expression analysis was combined with functional annotation, resulting in the identification of 12 candidate genes, of which two were particularly noteworthy. Both *GG9708* and *GG9868* appear to be differentially expressed in *B. thunbergii* during *Pg* inoculation and belong to gene families implicated in durable disease resistance in other plant-pathogen systems, namely leucine-rich

repeat receptor-like kinases (*GG9708*) and zinc ion binding SSM4 proteins (*GG9868*). Although subsequent validation and fine mapping studies are required, this study demonstrates the feasibility and lays the groundwork for dissecting the mechanism of *Pg*-NHR in the alternate host, with the hope that such work may contribute insight into possible novel mechanisms of durable rust resistance in wheat.

Keywords: Wheat | Stem rust | Barberry | Non-host resistance | Durable resistance | Linkage mapping | Genome assembly

Introduction

Stem rust, caused by the fungal pathogen *Puccinia graminis* (*Pg*), is one of the most destructive diseases of wheat and related small grains [1], causing severe epidemics and major recurring yield losses globally for millennia [2,3]. Since the middle of the 20th century, stem rust has been effectively controlled through a combination of the ongoing development of resistant wheat varieties and the removal of the pathogen's alternate host *Berberis vulgaris* near major wheat growing areas [3,4]. In the last 20 years, however, the emergence of new virulent stem rust races has rendered some long-used resistance genes ineffective [5,6]. For example, when the *Ug99* stem rust race was detected in East Africa in 1998, more than 80% of the world's wheat germplasm was estimated to be vulnerable to its novel virulence on the widely deployed resistance gene *Sr31* [7]. The rapid distribution and continued evolution of the *Ug99* family of races, combined with recent outbreaks of stem rust in Europe [8], underscore the importance to long-term wheat security of developing new strategies of durable resistance [9]. Traditionally, researchers have turned to the diverse *Triticum* genepool for new sources of resistance. Although translatability to wheat improvement may be less straightforward, a complementary approach is to look outside the *Triticum* genepool for insight into mechanisms of intrinsically durable non-host resistance (NHR) to the complex *Pg* pathogen.

The defining characteristics of NHR is that it is a form of resistance in which all individuals of a potential host species exhibit immunity to all individuals (e.g. races) of a potential pathogen [10]. As the most common form of resistance, NHR presents a compelling vision to provide broad-spectrum, durable protection from many plant pathogens, including the causal organism of wheat stem rust [11,12]. The genetic mechanisms underlying *Pg*-NHR remain largely unknown, however, especially in comparison to the relatively well-studied

mechanisms of race specific and quantitative, race non-specific host resistance. While current knowledge of true *Pg*-NHR is limited to one study conducted in rice [13], intermediate *Pg*-NHR resistance, in which the majority of individuals of a potential host species exhibit immunity but a few exhibit moderate susceptibility to a non-adapted pathogen, has been investigated in barley and *Brachypodium distachyon* [14,15]. The study of *Pg*-NHR in rice demonstrated that, despite the macroscopic appearance of immunity, the pathogen has the capacity to develop haustoria and colonize mesophyll. Such a result suggests that rice mounts a post-haustorial defense response, one found largely to be comprised of callose and reactive oxygen species production [13]. It would be valuable to identify the gene(s) underlying such a response in a non-host species, but traditional inheritance and mapping studies cannot be used because the non-host, by definition, fails to segregate for resistance. To facilitate the study of the genetics of this type of resistance, a genetically tractable system segregating for *Pg*-NHR is desired.

The *Berberis*-*Pg* system recently has been proposed as a unique pathosystem to investigate the genetic basis of *Pg*-NHR [16]. Numerous species within the highly diverse *Berberis*, or barberry, genus are susceptible to *Pg* infection (e.g. *B. vulgaris*) [17,18], but others are considered non-hosts (e.g. *B. thunbergii*) [19]. Interspecific hybridization between such host and non-host barberry species is known to occur in nature [20]; thus mapping populations derived from such hybridization provides a potential means of mapping and dissecting the genetic basis of *Pg*-NHR.

Unlike rice, which has no co-evolutionary relationship with *Pg*, barberries are thought to be one of the first eudicots parasitized by the rusts; thus they may have played an important role in the evolution of several groups of rust fungi (Fig 1). Multiple lines of evidence support this line of thinking: 1) *Berberis* spp. host a wide diversity of rust species, including numerous

macrocyclic, heteroecious species of *Puccinia* (e.g. *P. graminis, P. striiformis, P. montanensis, P. brackypodii, P. pigmea, P. koeleriae*, and *P. errhenatheri*), a number of autoecious rusts (e.g. *Cumminsiella* spp. belonging to Pucciniaceae, *Edythea* spp. belonging to Uropyxidaceae, and *Pucciniosira* spp. belonging to Pucciniosiraceae), as well as anamorphic rusts (e.g. *Acedidium* and *Uredo* spp.); 2) Only slight morphological differences exist among the teliospores of the various macrocyclic rusts, suggesting a single evolutionary origin of these pathogens; and 3) A recent palaeobotanical finding of *B. wuganensis* from a sediment layer between 55 to 65 Mya in northeastern China suggests that the barberries are one of the earliest groups of angiosperms. Finally, there are eight known species of the macrocyclic, heteroecious *Puccinia* spp. that undergo their sexual (aecial) stage on barberry and their asexual (uredinial and telial) stages on graminaceous plants from the Poaceae family, further suggesting that *Puccinia* spp. likely parasitized *Berberis* spp. prior to their host expansion to the grasses.

Today, the genus *Puccinia* comprises more than 2,000 species; and within that diverse genus, host jump rather than co-speciation is believed to be the more frequent form of speciation [21]. In other words, host jumps from the Poaceae into other, newer plant families likely explains the origins of the majority, if not all, of the diverse *Puccinia* spp. that parasitize the grasses. For example, a host jump from Poaceae to Ranunculaceae likely produced the *P. recondita* complex and aligned species, a jump to Liliaceae likely produced *P. hordei* and aligned species, and a jump to Oxalidaceae likely produced *P. sorghi* and aligned species. Because the relationship between barberries and the rusts likely predates such speciation (Fig 1), it is of interest to understand the mechanism of NHR exhibited by some barberry species today, the descendants of the ancestral hosts of the cereal rust pathogens.

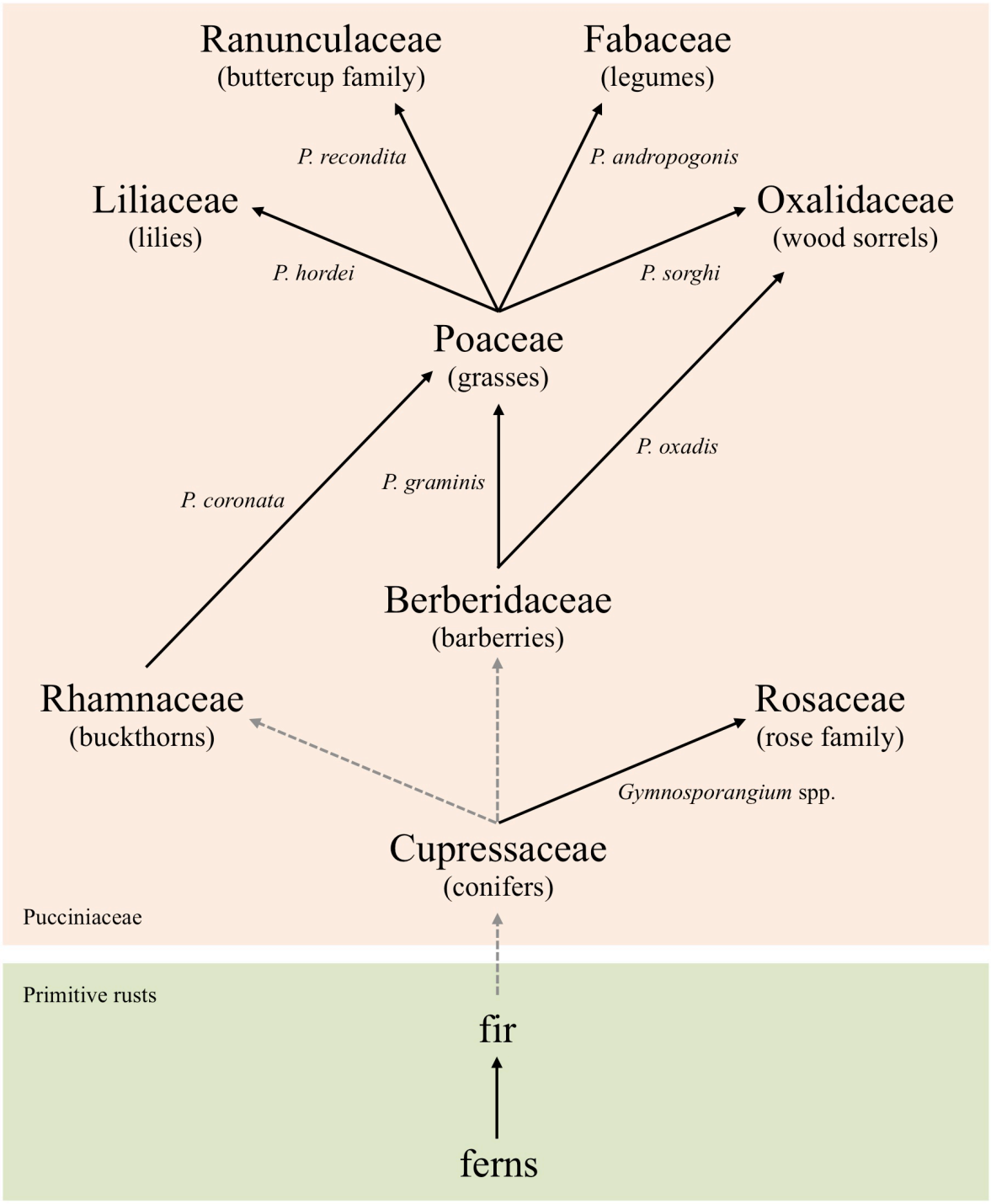


Fig 1 Hypothetical evolutionary chart showing evolution of modern day macrocyclic, heteroecious *Puccinia* spp. Solid arrow represents host jump between respective host group supported by the presence of existing rust species connecting the groups. Dotted line represents host jump between respective host group indirectly supported by similarity of teliospores but a lack of existing rust species connecting the respective host groups.

Quantitative trait loci (QTL) mapping is a proven method for dissecting the genetic underpinnings of both simple and complex traits in plants, including disease resistance [22]; but successful QTL analysis in a species is possible only once basic genomic tools and resources become available [23]. For example, it is the availability of extensive genomic resources, such as a high quality reference genome [24] and dense genetic linkage maps [25–28], that has facilitated significant progress in wheat research in recent years. For the barberries, in contrast, available resources prior to this study were limited to 7 microsatellite markers [29], 238 AFLPs [30–32], and a set of GBS data from natural populations recently deposited on Sequence Read Archive database [16]. A high-quality genetic linkage map is a necessary prerequisite to begin mapping genes governing *Pg*-NHR in barberry, and a reliable reference genome is required to facilitate QTL dissection and motivate candidate gene postulation. Once developed, such genomic resources also have the potential to support researchers tasked with identifying *Berberis* spp. in the field as part of a global rust surveillance network [33] as well as provide molecular tools to ornamental breeders.

The primary objective of this research is to identify gene(s) associated with *Pg*-NHR in *B. thunbergii*. To achieve this, an F_1 bi-parental mapping population was developed via an inter-specific cross between *Pg* susceptible *B. vulgaris* and *Pg* non-host *B. thunbergii*; and that population was used to construct the first genetic linkage maps for both parental species and perform QTL analysis. To gain insight into the detected *Pg*-NHR QTL region, second and third generation sequencing platforms were used to assemble a high quality reference genome and transcriptome for *B. thunbergii*. Combining the results of QTL mapping with differential expression analysis and *B. thunbergii* - *B. vulgaris* sequence variation analysis, two noteworthy candidate genes for *Pg*-NHR were identified. This study not only establishes foundational

resources for the novel *Berberis-Pg* pathosystem but also demonstrates its use in an initial dissection of *Pg*-NHR, with the long-term hope of contributing insight into possible novel mechanisms of durable resistance to the wheat stem rust pathogen.

Materials and Methods

Mapping population development

An F_1 mapping population consisting of 182 individuals was derived from an inter-specific cross between *B. vulgaris* accession 'Wagon Hill' (female parent) and *B. thunbergii* accession 'BtUCONN1' (pollen parent). Wagon Hill is susceptible to stem rust and is a relatively taller shrub (~ 3 m height) that displays 2-5 cm long obovate to obovate-oblong leaves with highly serrated margins (>50 serrations) and has 5-8 cm long pendant racemes of bright yellow flowers. In contrast, BtUCONN1 is a non-host to the stem rust pathogen and is a small shrub (0.5–2.5 m height) that displays 1.3–3.8 cm long entire leaves and 1-2 cm long inflorescences with few umbellate but mostly solitary flowers. The female parent Wagon Hill is a feral plant growing along the shoreline of the Great Bay Estuary in Durham, New Hampshire (N43°07'30.64", W70°52'17.95") and the pollen parent BtUCONN1 is a feral plant maintained in the barberry collection at the research farm of the University of Connecticut (N41°47'40.63", W072°13'39.61").

To make the inter-specific cross, pollen was harvested from mature flowers of BtUCONN1 using the previously described N-pentane method [34] and stored at 4ºC until flowers of Wagon Hill reached reproductive maturity. Emasculation and hand pollination of female flowers was performed at the so-called balloon stage, when the petals begin to part slightly at the top, giving the appearance of an inflated balloon prior to opening. Seeds from

successful crosses were individually handpicked and stratified in wet sand in a petri dish at 4 ºC for three months to break dormancy before sowing. Propagated cuttings of the two parents were maintained along with the F_1 mapping population in plastic pots (11.5 cm diameter; 6.5 cm tall) filled with PRO-MIX HP growth media in the Macfarlane Greenhouse facility at the University of New Hampshire, Durham.

To verify the putative F_1 status of the individuals in the mapping population, a PCR-based species-specific marker was designed based on available GBS data [16]. A universal primer pair was designed to amplify a short genomic sequence exhibiting a length polymorphism between the two parents. Specifically, the primers (F: 5'-CCTGATTGGGGCTCATTATC-3'; R: 5'-AGTGAGGAATTCCGAGCTGA-3') amplified a 208 bp fragment in Wagon Hill but only a 195 bp fragment in BtUCONN1, due to the presence of a 13 bp indel (see Text S1). PCR was conducted with a total reaction volume of 20 µl (0.25 mM of each primer, 100 µM of each dNTP, 0.75 U Taq DNA Polymerase, 10x standard Taq buffer, and 100 ng of template DNA). Cycling conditions consisted of 5 minutes at 94°C, 32 cycles of 30 s at 94°C, 30 s at the primer annealing temperature of 52°C, and 15 s at 68°C, followed by a final 5 minute elongation step. Amplified products were separated on a 3% TBE/EtBr agarose gel for 60 min at 75 V and imaged with UV transillumination. The F_1 status of a putative hybrid individual was considered validated if both bands from the two parental species were detected (Fig S1).

Genotyping and variant detection

Genomic DNA of the 182 verified F_1 individuals and both parents was extracted from ~100 mg of lyophilized leaf tissue using a modified CTAB method [35]. Prior to GBS library preparation, isolated DNA was purified using Zymo Research's Genomic DNA Clean &

Concentrator^{TM-10} column (Catalog # D4011), following manufacturer's protocol. Reduced representation libraries were constructed using the two-enzyme (*PstI-MspI)* GBS protocol described by Poland et al. [36] and sequenced via 150 bp paired-end (PE) reads on an Illumina HiSeq 2500 at the Hubbard Center for Genome Studies, UNH.

Raw FASTQ files were generated by CASAVA 1.8.3 and analyzed using the reference-free bioinformatics pipeline GBS-SNP-CROP v.3.0 [37]. A Mock Reference (MR) was constructed using the high quality PE reads from the two parents; and putative variants, both SNPs and indels, were identified via alignment of high quality PE reads from the parents and all F_1 progeny to the MR, following the pipeline's recommended parameters for diploid species. Complete details of the GBS-SNP-CROP command lines used in this analysis, including all specified pipeline parameters, are provided in Text S2.

Genetic linkage map construction

A detailed account of the series of filters applied to obtain the final set of markers for linkage map construction is provided in result. In short, a marker was culled if it met any of the following criteria: 1) It was unscored for more than 30% of the individuals in the population; 2) It was heterozygous for both parents; 3) It did not segregate in the population (i.e. all progeny were heterozygous for the marker); 4) Its mean ratio of primary to alternate allele depth deviated significantly from the expected ratio of 1:1; and 5) Its segregation ratio deviated significantly from the expected ratio of 1:1, according to its marker class (Table 1). As a final filter, genotypes with >30% missing data were removed.

Linkage analysis was performed using ONEMAP v2.0-4 R package [38], and separate linkage maps were constructed for the two parents according to a two-way pseudo-test-cross

mapping strategy [39]. The BtUCONN1 linkage map was constructed using marker sets 1 and 2, while the Wagon Hill map was constructed using marker sets 3 and 4 (Table 1). For each map, a two-point test was first performed for all marker pairs, using a minimum LOD score of 4 and a maximum recombination fraction of 0.25 to group markers into linkage groups (LGs). Next, markers within each LG were ordered using the 'try' algorithm within ONEMAP.

Table 1 Four classes (sets) of markers were selected for linkage map construction.

Marker Set	BtUCONN1	Wagon Hill	Expected segregation of F_1 genotypes
For generating a linkage map for *B. thunbergii*			
1	ab	aa	aa:ab (1:1)
2	cd	--	c:d (1:1)
For generating a linkage map for *B. vulgaris*			
3	ee	ef	ee:ef (1:1)
4	--	gh	g:h (1:1)

To identify potential genotyping errors, common in GBS data [40], maps were manually inspected for the presence of singletons (apparent double crossovers) [41], which were replaced with missing values. If multiple markers were found to map to the same genetic bin, a consensus of the set of markers was chosen to represent the linkage bin for final mapping iterations, which were made until no alternative orders were generated by the 'ripple.seq' function. Final map distances were calculated with the Kosambi mapping function [42], and ideograms were generated using Mapchart 2.0 [43]. All mapped markers were tested for locus specific segregation distortion. Specifically, the chi-square goodness fit was used to test for significant deviation (p-value < 0.01) from expected Mendelian inheritance, which may arise as a result of gametic selection or post post-zygotic selection.

Stem rust disease phenotyping

To determine disease responses, the parents and all F_1 individuals in the mapping population were inoculated with *Pg* basidiospores, using with *Pg* telia found on naturally infected *Elymus repens*, as previously described [16]. The pollen parent BtUCONN1 exhibits the clear non-host reaction typical of *B. thunbergii*. In contrast, the female parent Wagon Hill exhibits the clear susceptible reaction of *B. vulgaris*, with well-developed mature aecia visible on the abaxial surfaces of leaves. Images of typical reactions of the parents and of individuals in the F_1 mapping population are presented in Fig 2. As detailed in Table 2, a 4-point scale was developed in response to the particular segregating characteristics observed in this population. The levels of this scale are based on consideration of the following symptoms: 1) Degree of flecking; 2) Presence and intensity of necrotic lesions; and 3) Presence and density of pycnia and aecia. All plants were scored for reaction to stem rust 14 days after inoculation.

Table 2 Descriptions of the levels of the 4-point scale developed for phenotyping the disease reactions of the individuals comprising the F_1 mapping population in this study.

Scale[1]	Description
1	Sparse flecking and necrotic lesions, sometimes < 3 tiny pycnia without aecia in a leaf
2	Evident necrotic lesions; 5 to 15 obvious pycnia with or without aecia in a leaf
3	5 to 15 well developed pycnia with aecia in a leaf; sparse necrotic lesions
4	>15 well developed pycnia and aecia; not preceded by necrosis

[1] In general, a resistant (R) reaction is indicated by a score of 1, a moderate resistant (MR) reaction by 2, a moderate susceptible (MS) reaction by 3, and a susceptible (S) reaction by 4. Representative images for each disease class are presented in Fig 2.

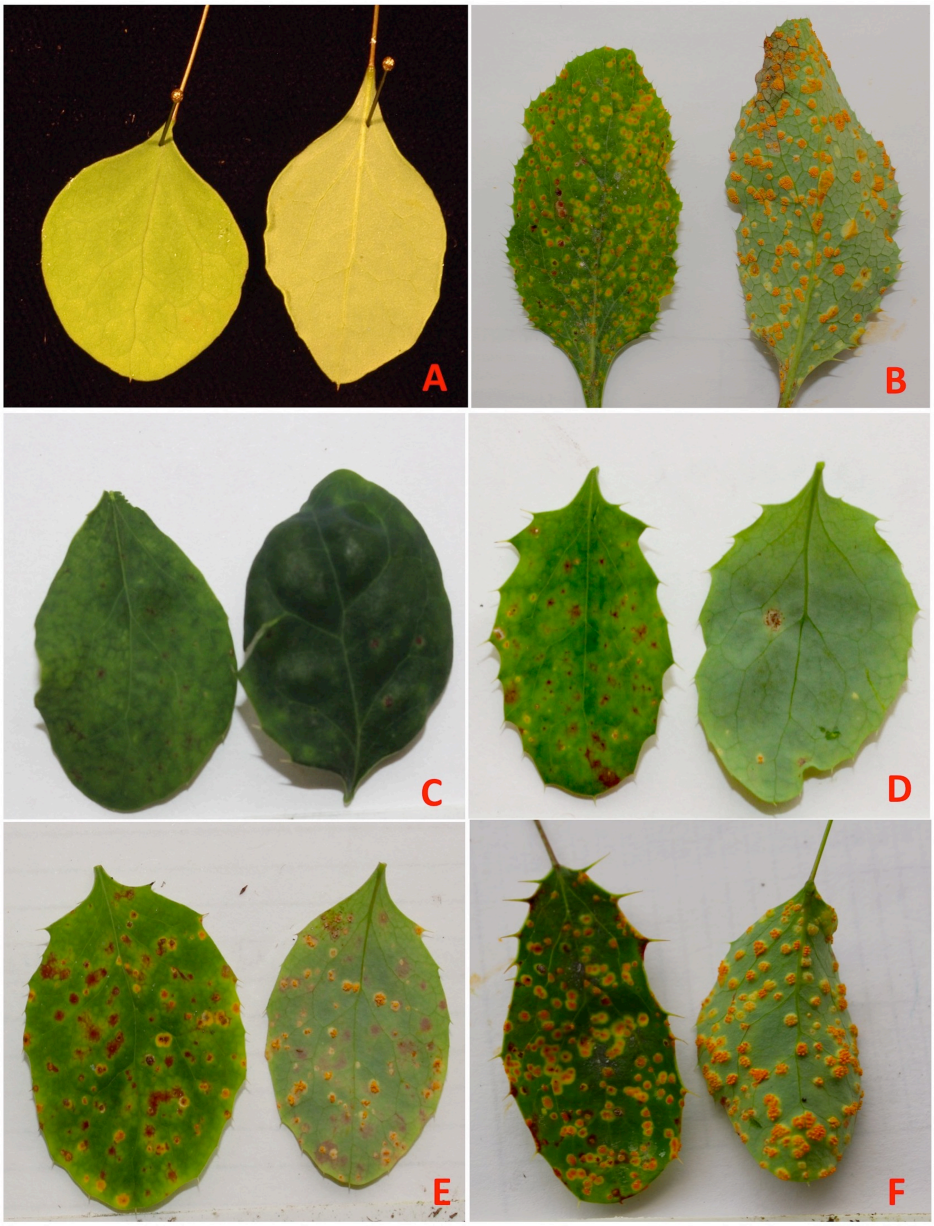


Fig 2 Representative disease responses of the two mapping population parents, *B. thunbergii* accession 'BtUCONN1' and *B. vulgaris* accession 'Wagon Hill', motivate the four-point disease reaction scale used to phenotype the F_1 mapping population. **(A)** Resistant reaction of *B. thunbergii* accession 'BtUCONN1', showing no visual symptoms; **(B)** Susceptible reaction of *B. vulgaris* accession 'Wagon Hill', showing dense pycnia on the upper leaf surface and prolific, well-developed aecia on the lower surface; **(C)** Resistant reaction (score of 1 on the four-point scale) of *B.* ×*ottawensis* progeny 'WH15-039', showing sparse flecking; **(D)** Moderate resistant reaction (score 2) of *B.* ×*ottawensis* progeny 'WH15-063', showing evident necrotic lesions and some pycnia formation; **(E)** Moderate susceptible reaction (score 3) of *B.* ×*ottawensis* progeny 'WH15-128', showing well-developed pycnia and aecia, alongside sparse necrotic lesions; and **(F)** Susceptible reaction (score 4) of *B.* ×*ottawensis* progeny 'WH15-149', showing well-developed pycnia and aecia and no evident necrosis. All photos were taken 14 days post-inoculation.

QTL analysis

QTL analysis for *Pg* resistance was performed using both the parental and maternal genetic linkage maps in R using R/qtl v1.39-5 [44,45]. Haley-Knott regression [46] was used, based on the composite interval mapping method (CIM); and a QTL was deemed significant if its peak LOD score exceeded the threshold determined via permutation analysis (1,000 permutations, 5% significance level).

Reference genome assembly

B. thunbergii cv 'Kobold', a commercial green-leafed cultivar common in the ornamental industry, was selected for whole genome sequencing. Kobold is a heterozygous diploid (2n = 2x= 28) and is a non-host to stem rust [47]. Cuttings of Kobold were obtained from the barberry collection at the University of Connecticut, rooted, and maintained in the MacFarlane Greenhouses at UNH under standard conditions for barberry [16]. For sequencing, ~2 g of fresh young leaves were collected from 4-6 clonally propagated plants and flash frozen in liquid nitrogen. Genomic DNA was extracted using modified a CTAB procedure [48] and quantified via both fluorometry (Qubit, Thermo Fisher Scientific, Waltham, U.S.A.) and agarose gel electrophoresis with a lambda DNA standard. A 20-kb BluePippin kit (PacBio) was used for Single Molecule Real Time (SMRT) library preparation; and 115 SMRT cells were sequenced on the PacBio RS II system at the UC Davis Genome Center, using P6-C4 chemistry. All data were collected as 6-h sequencing videos.

The FALCON v0.3.0 and FALCON-Unzip toolkits [49] were used for whole genome assembly and phasing. FALCON is described as a Hierarchical Genome Assembly Process (HGAP) pipeline that generates a genome assembly from long PacBio reads through the

following basic steps: 1) Raw read error correction via alignment of sub-reads; 2) Pre-assembly of long, error-corrected reads; 3) Overlap detection pre-assembled reads; 4) Overlap filtering; 5) Overlap graph construction; and 6) Graph-based contig construction. After this initial assembly, FALCON-Unzip is used in highly heterozygous species to resolve the distinct haplomes (i.e. unzip the genome) based on patterns of structural variants and associated SNPs (i.e. haplotype blocks). This unzip process gives rise to a set of so-called primary contigs (the primary assembly) and a set of associated haplotigs (phased variants of the primary contigs, in regions of high heterozygosity). Complete details of the FALCON assembly parameters used in this study are provided in Text S3. Finally, the Arrow algorithm from the 'GenomicConsensus' PacBio package (https://github.com/PacificBiosciences/GenomicConsensus) was used to polish the phased primary contigs and their associated haplotigs. The genome size was estimated using k-mer analysis of the error-corrected PacBio reads [50] and propidium iodide flow cytometric analysis using *Pisum sativum* L. Citrad (2C= 9.09 pg) as an internal standard BD Accuri™ C6 Cytometer [51].

Further polishing and curation of the assembly was accomplished using the Purge Haplotigs pipeline [52]. High levels of heterozygosity in some genomic regions can lead to the incorrect assignment of haplotigs as distinct primary contigs [52]. To identify such errors and correctly assign homologous contigs into the haplotig pool, the Purge Haplotigs pipeline first performs a read-depth analysis [53] to flag abnormally low or high coverage contigs as potential chimeras and then performs a BLAST [54] against the entire assembly to identify putative primary contigs exhibiting high homology to one another. During this process, alignment dotplots are produced, and these are manually screened to break likely chimeras, to define the

final set of primary contigs as the reference sequence, and to assign residual syntenic contigs as haplotigs. Complete details of the script used for purging halpotigs are provided in Text S4.

Assessment of genome assembly quality and Hi-C scaffolding

Quality of the final curated assembly was assessed using QUAST [55], and assembly completeness was evaluated using the set of 1,440 core plant genes in BUSCO v3 [56]. To identify and purge contaminant contigs, the final assembly was aligned using BLAST to the following databases of possible contaminants: plasmid DNA (cpDNA and mtDNA) from angiosperms, the human genome (GRCh38.p7), the *Escherichia coli* genome (CP017100.1), and 16S and 18S rRNAs. The rRNA database was created using the SILVA project [57], and the others were created via sampling from Genbank. To further evaluate completeness, the PacBio error corrected reads (preads), the RNA-seq data generated for transcriptome assembly (see below), and the GBS data from the BtUCONN1 parent generated for linkage mapping were also aligned to the final assembly.

To linearly orient and order the primary contigs into chromosome-scale pseudomolecules, Proximity-Guided Assembly was performed using Phase Genomics' Proximo^{TM} chromosome conformation capture (Hi-C) analysis [58]. Tissue processing, chromatin isolation, library preparation, sequencing, and Hi-C analysis were performed by Phase Genomics (Seattle, WA, USA). Finally, the BtUCONN1 genetic linkage map was used to manually curate the Hi-C assembly using JuiceBox [59], bringing independent information to guide the ordering of a set of anchor contigs in instances of ambiguity.

Anchoring of the genetic linkage maps to the physical assembly

Orthogonal sets of markers were used to build the genetic linkage maps of the two parents; thus the two maps share no markers in common, preventing a direct assessment of synteny between the two species. The physical assembly, however, presents a potential "common language" by which the two maps can be compared, provided the markers in the linkage maps can be uniquely located in (i.e. anchored to) the physical assembly. To accomplish this, BLASTn [60] was performed between the MR centroids (queries) and the curated assembly (subject). Using only those centroids exhibiting unique positions within the reference genome, synteny plots were generated using the Pacth function of the Matplotlib plotting library (https://matplotlib.org/index.html). The above anchoring method was also used to project the detected *Pg*-NHR QTL region onto the physical map, thus permitting insight into its underlying physical sequence.

Transcriptome assembly

For transcriptome assembly, ten different tissues, including immature leaf tissue at various time points after *Pg* inoculation, were collected from a clonally propagated plant of *B. thunbergii* cv 'Kobold' (Table S1). Fresh tissues were flash frozen in liquid nitrogen and ground to fine powder using mortar and pestle. Total RNA was isolated using the Zymo Research RNA Clean & Concentrator™ kit (Catalog # R1015), according to the manufacturers' protocol. RNA-seq libraries were prepared with Illumina TruSeq® RNA Library Prep Kits and sequenced via 150 bp paired-end (PE) reads on an Illumina HiSeq 2500 at the Hubbard Center for Genome Studies, UNH.

CASAVA-processed raw sequences were error-corrected using the software BFC v1.0

[61], following recommendations from the Oyster River Protocol For Transcriptome Assembly [62]. Error-corrected reads were then processed to remove Illumina adapters and gently trimmed to remove low quality reads (Phred ≤ 5) using Trimmomatic v.0.33 [63]. All post-processed reads from the 10 tissues were pooled, and the transcriptome was assembled using Trinity (reference-guided *de novo* assembly) [64,65]. Assembly quality was evaluated using TransRate [66], and its completeness was assessed using the set of 1,440 core plant genes in BUSCO v3 [56]. In addition to providing basic summary statistics and quality metrics, TransRate provides an overall score of transcriptome contiguity based on various mapping metrics; and BUSCO evaluates assembly content based on the representation of expected single copy orthologs.

Identification of candidate genes

To facilitate the identification of candidate genes that may explain the association of the detected QTL region to *Pg* response, the physical contigs spanning the QTL region were first annotated using the RepeatMasker software [67]. Next, a functional annotation was performed with the Maker pipeline [68], using both *ab-initio* and transcriptome-based analyses. The set of well-supported genes within the QTL region, hereafter referred to as high-confidence (HC) genes, were defined based on Maker's Annotation Edit Distance quality metric (AED < 0.5) and non-overlapping genes greater than 500 bp.

Combinations of approaches were taken to pare down the full set of HC genes to those more likely to contribute to *Pg*-NHR. A differential gene expression (DGE) analysis experiment was conducted to identify genes whose levels of expression change under challenge by *Pg*. Three biological replicates of immature leaves were sampled from clonally propagated *B. thunbergii* cv. 'Kobold' plants at four different time points: pre-inoculation (T0) and 48, 72, and 144 hrs post-inoculation (T48, T72, and T144). Total RNA was extracted, sequenced, and

processed as described above. Transcript abundance was quantified using Kallisto [69], and time course analysis was performed using Sleuth [70]. Complete details of the parameters used for transcript abundance and time course analysis are provided in Text S5.

The final list of high-priority candidate genes is composed of those HC genes in the QTL region that are differentially expressed under *Pg* inoculation and show homology to gene families implicated in disease resistance in other plant pathosystems. Putative protein functions and Gene Ontology (GO) terms were assigned to the candidate genes using both the Phytozome v.12.1 [71] and UniProtKB [72] databases.

Results

Variant detection and linkage map construction

GBS libraries were constructed for the two parental lines and 182 F_1 progeny, generating a total of 60 Gb of data (~401 million 150-bp PE reads). After quality parsing and demultiplexing, the average of 3 million high quality reads per genotype were retained by the GBS-SNP-CROP pipeline (Table S2). Using the high quality reads from the two parents, a mock reference (MR) comprised of 87,089 centroids (i.e. consensus GBS fragments) was generated, with a total length of approximately 15.35 Mbp.

A total of 15,411 polymorphic markers, including 14,043 SNPs (average depth D_{SNPs} = 41.5) and 1,368 indels (D_{indels} = 36.4), were identified by mapping all high quality reads from the population to the MR. A detailed account of the progression of filters applied for obtaining the final sets of markers for linkage map construction is provided in Table 3. Separate genetic linkage maps were constructed for each species, according to a two-way pseudo-test-cross mapping strategy. After filtering individual F_1 lines with >30% missing data, 161 and 162

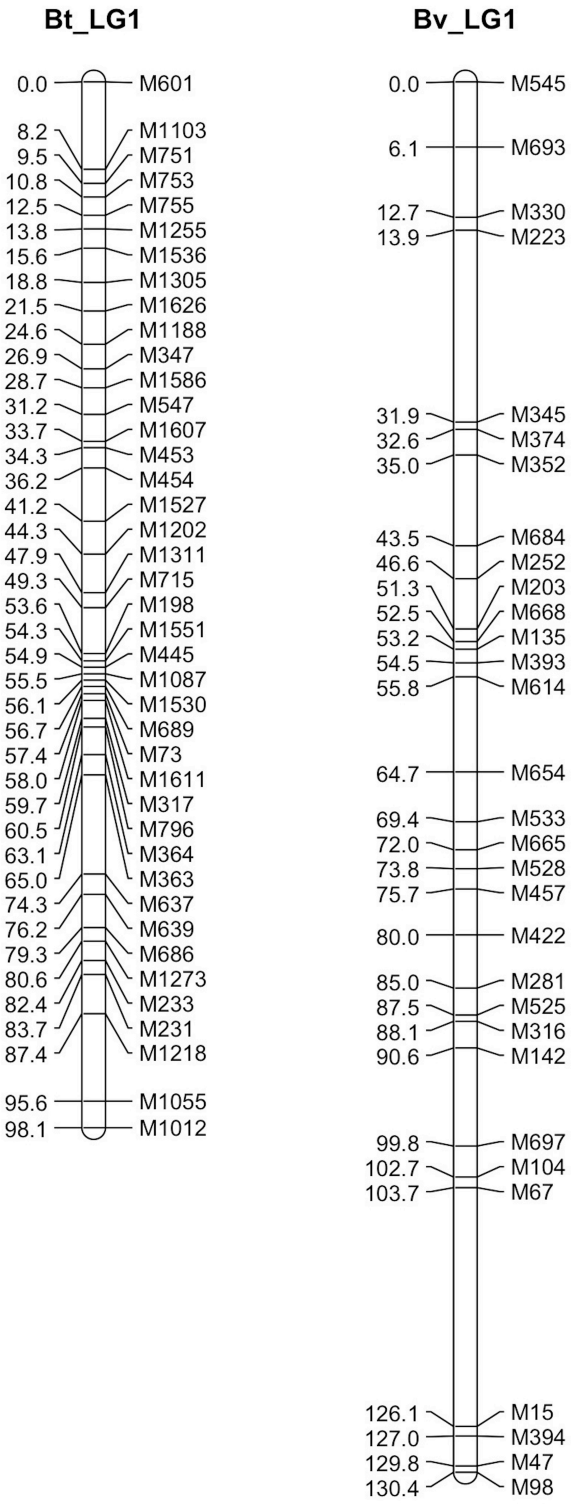
individuals were retained for BtUCONN1 and Wagon Hill linkage map construction, respectively. The BtUCONN1 map was constructed using a total of 1,757 makers (1,497 and 260 from Marker Sets 1 and 2, respectively; see Table 3), and the Wagon Hill map was constructed using a total of 706 markers (600 and 106 from Marker Sets 3 and 4, respectively). Two additional markers were removed from the BtUCONN1 analysis due to lack of linkage with any other markers. From the Wagon Hill analysis, seven ambiguous markers were similarly removed. For both parental species, the remaining markers coalesced into 14 distinct linkage groups, in agreement with the 14 chromosomes present in these *Berberis* spp (Fig 3).

Table 3 Description of the progression of filters applied to the initial set of 15,411 markers (SNP's and indels) identified by the GBS-SNP-CROP pipeline to obtain the final sets of markers for construction of linkage maps

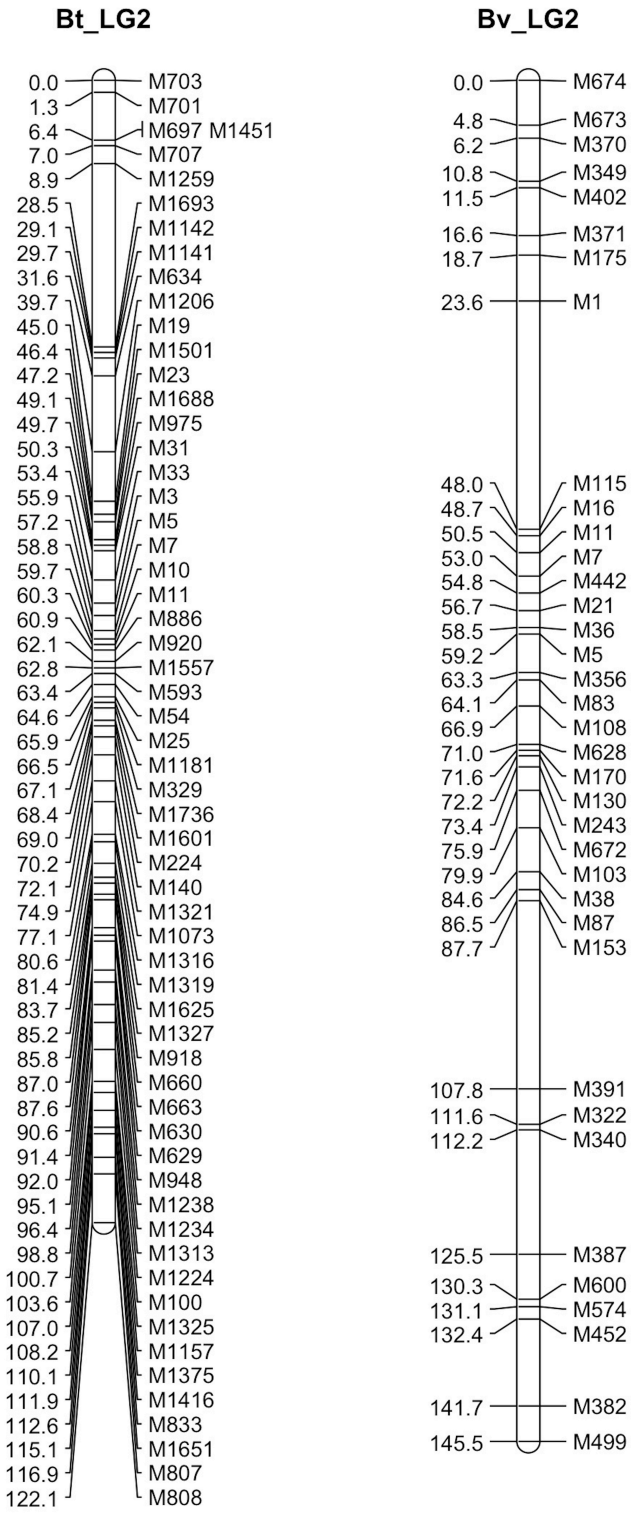
Filter descriptions	Markers removed at each step	Markers retained
1. More than 30% missing genotype calls across the population	6,106	9,305
2. Heterozygous in both parents	272	9,033
3. Homozygous for alternate alleles in the two parents [a]	3,982	5,051
4. Deviating significantly from expected allele depth ratio in heterozygotes [b]	1,801	3,250
5. Segregating genotypes unsupported by parental genotypes [c]	697	2,553
6. Deviating significantly from expected Mendelian segregation [d]	90	2,463
Final markers for the *B. thunbergii* linkage map (*Bt* × *Bv*)		**1,757**
Set 1: ab × aa		1,497
Set 2: cd × --		260
Final markers for the *B. vulgaris* linkage map (*Bt* × *Bv*)		**706**
Set 3: ee × ef		600
Set 4: -- × gh		106

[a] Both parents are homozygous for the marker and no variation is observable among the F_1 progeny (i.e. all are heterozygous for the marker)
[b] Mean allele depth ratio across heterozygote F_1 progeny deviates > 25% from the expected bi-allelic depth ratio of 1:1
[c] Lack of parental genotypes (missing data) and/or parental genotyping errors can prevent the unique assignment of gametic origin. For example, while ab × aa is expected to segregate only as aa and ab among the progeny, the alternate homozygote (bb) may be observed due to parental genotyping error. All such markers were removed from the analysis.
[d] Segregation ratio of genotypes deviates > 2 SD from the expectation for each marker set; such markers were removed due to their high segregation distortion

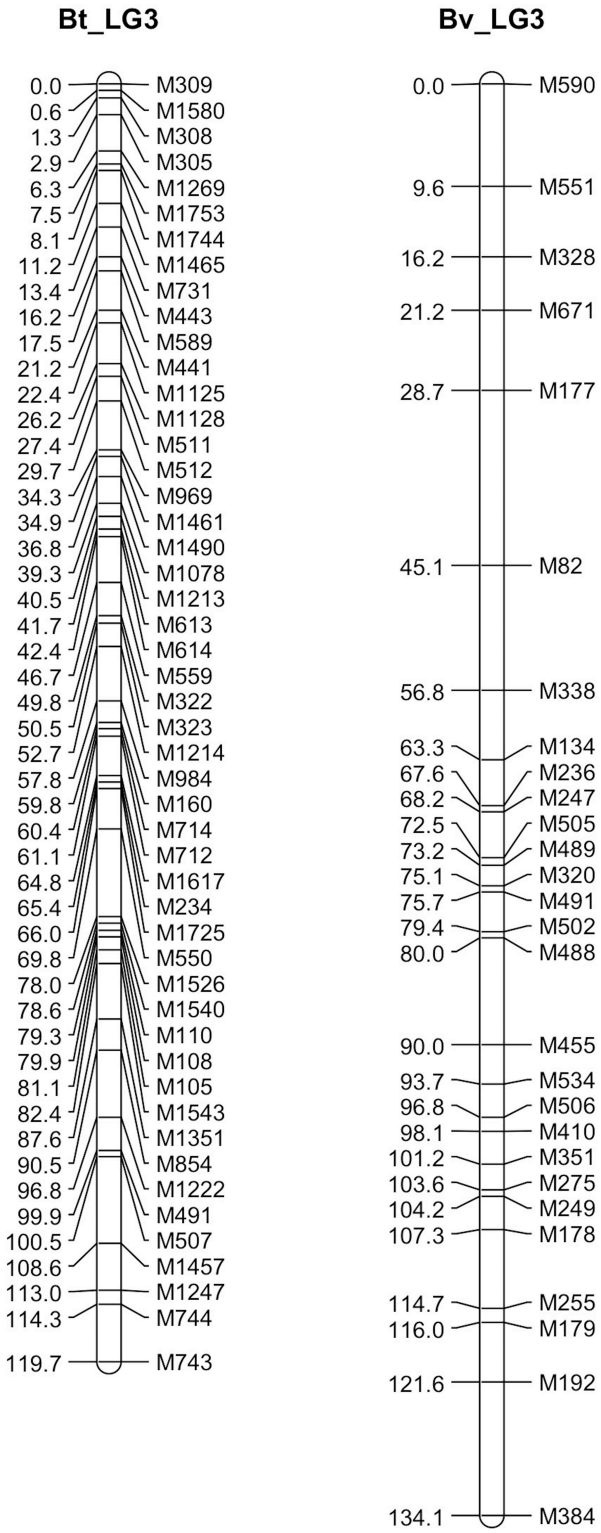
Chromosome 1



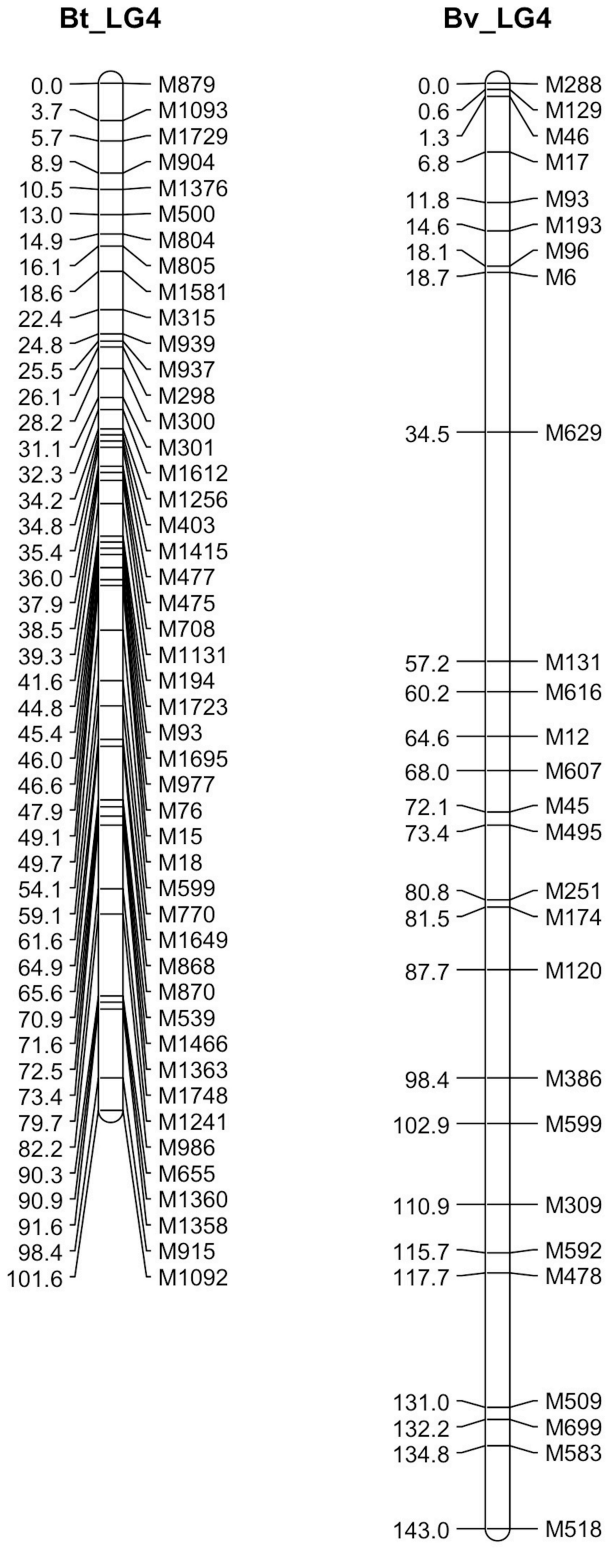
Chromosome 2



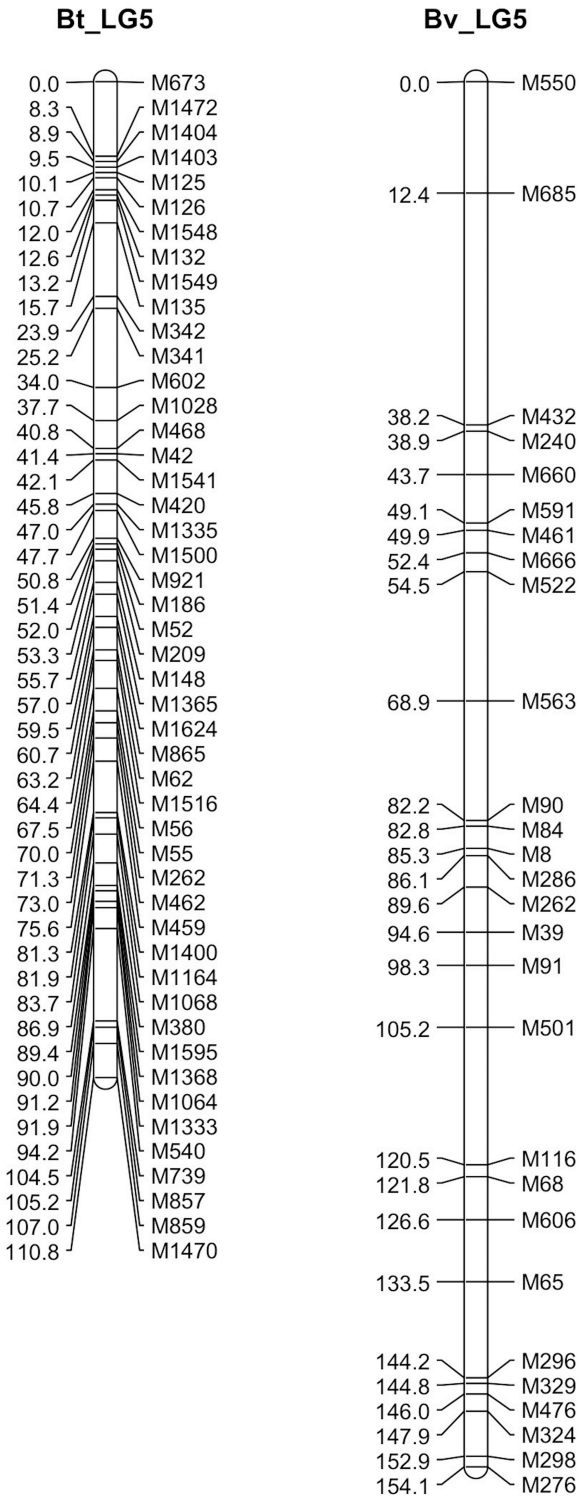
Chromosome 3



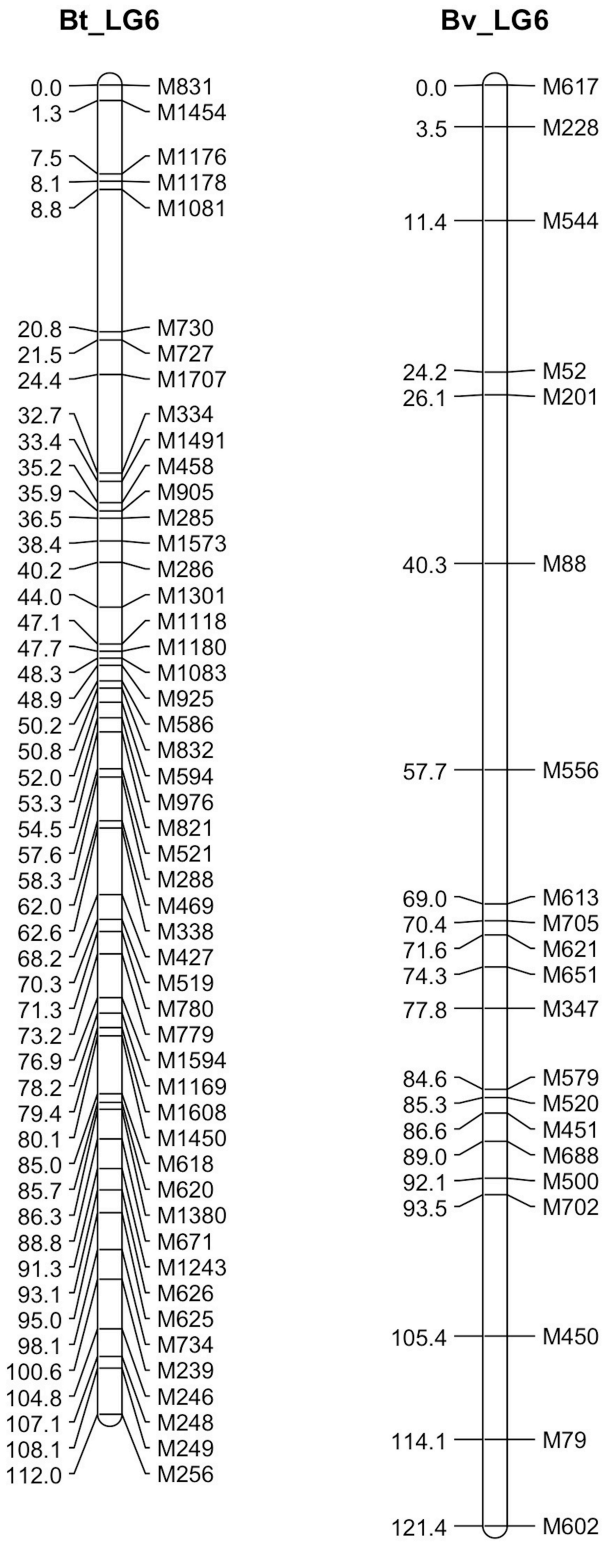
Chromosome 4



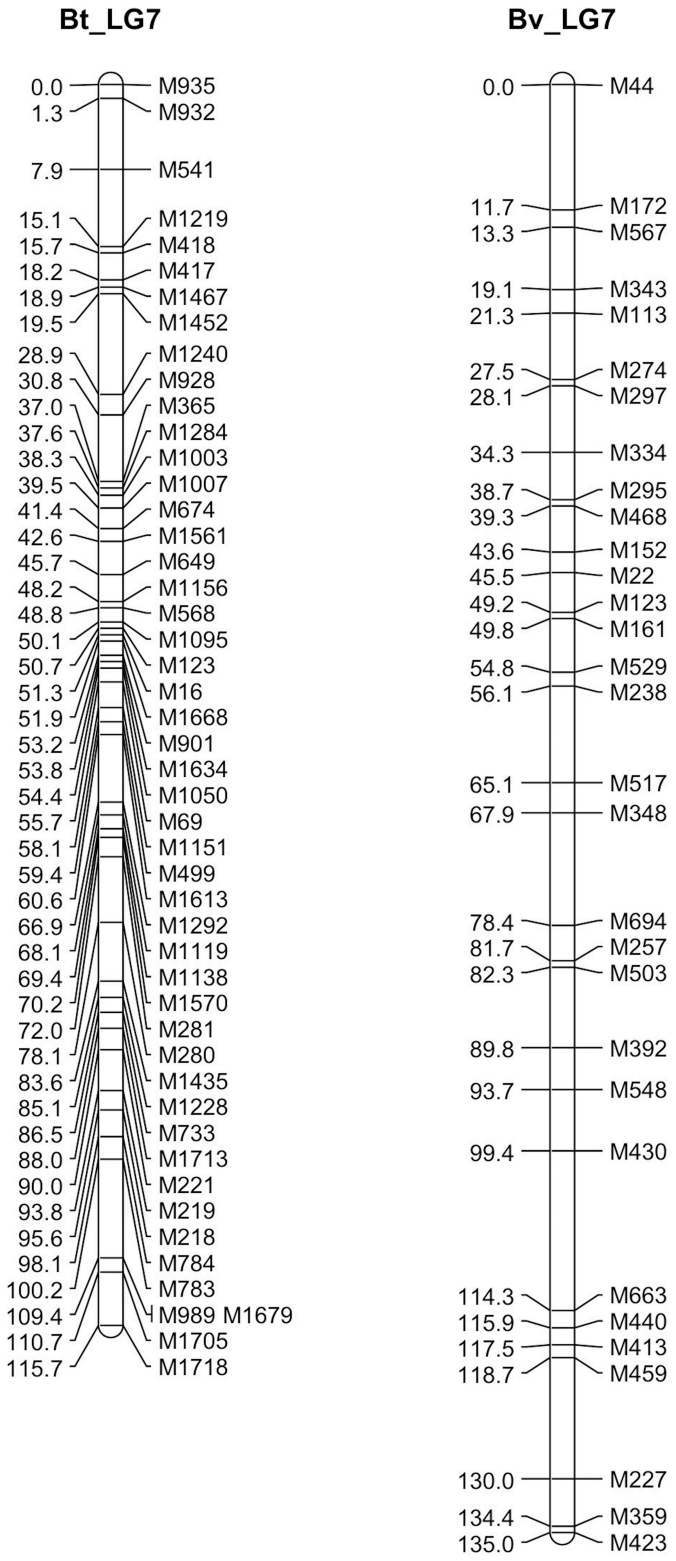
Chromosome 5



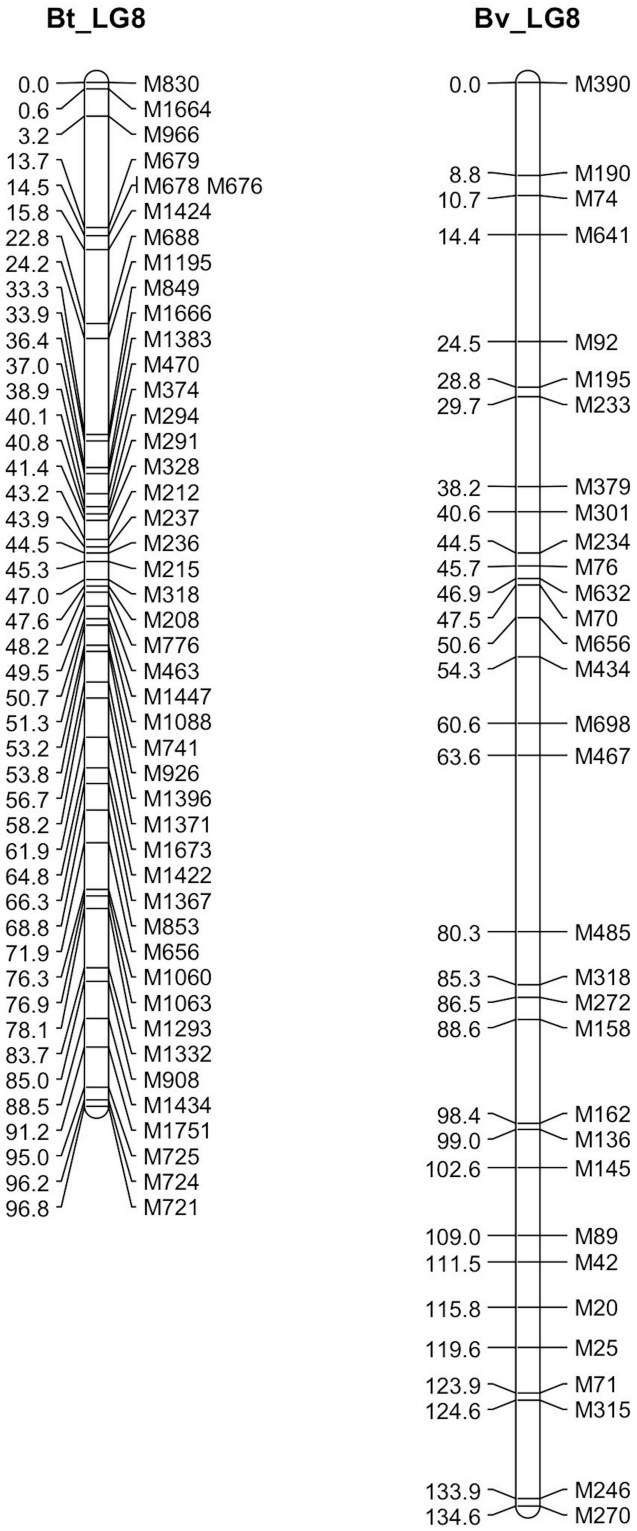
Chromosome 6



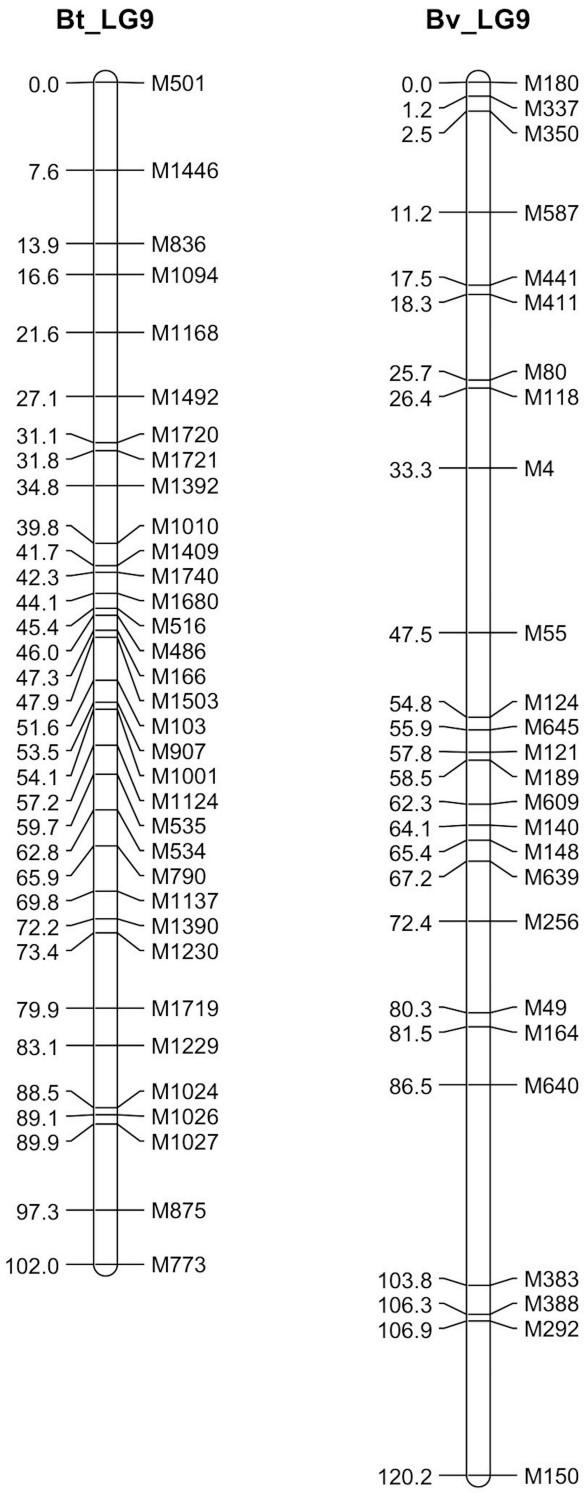
Chromosome 7



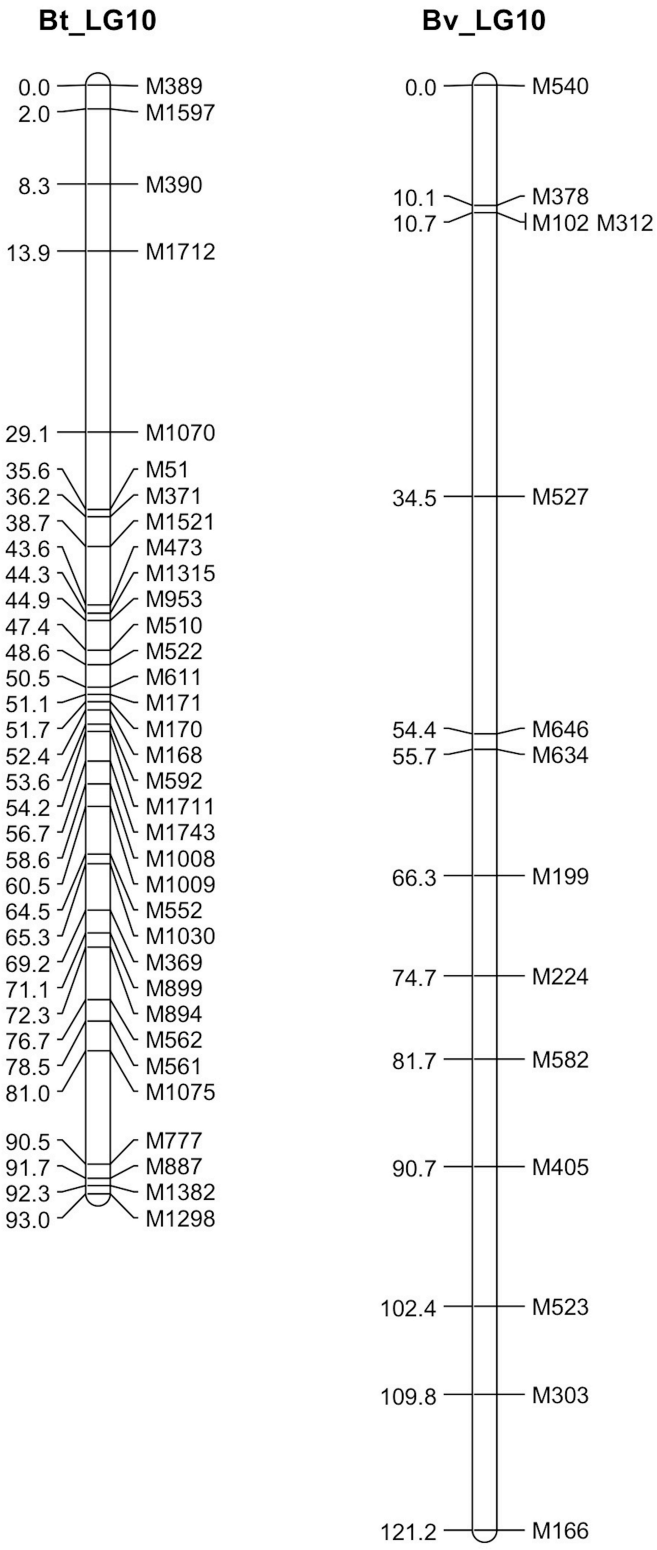
Chromosome 8



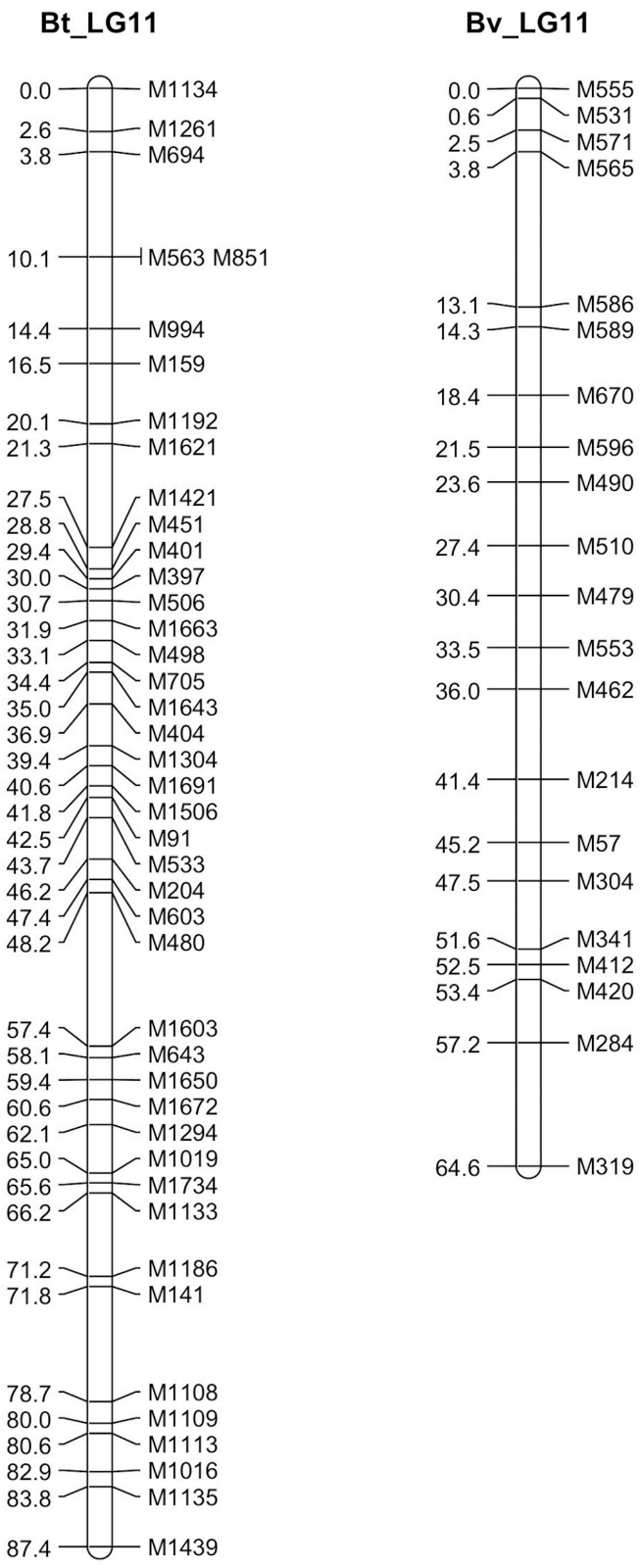
Chromosome 9



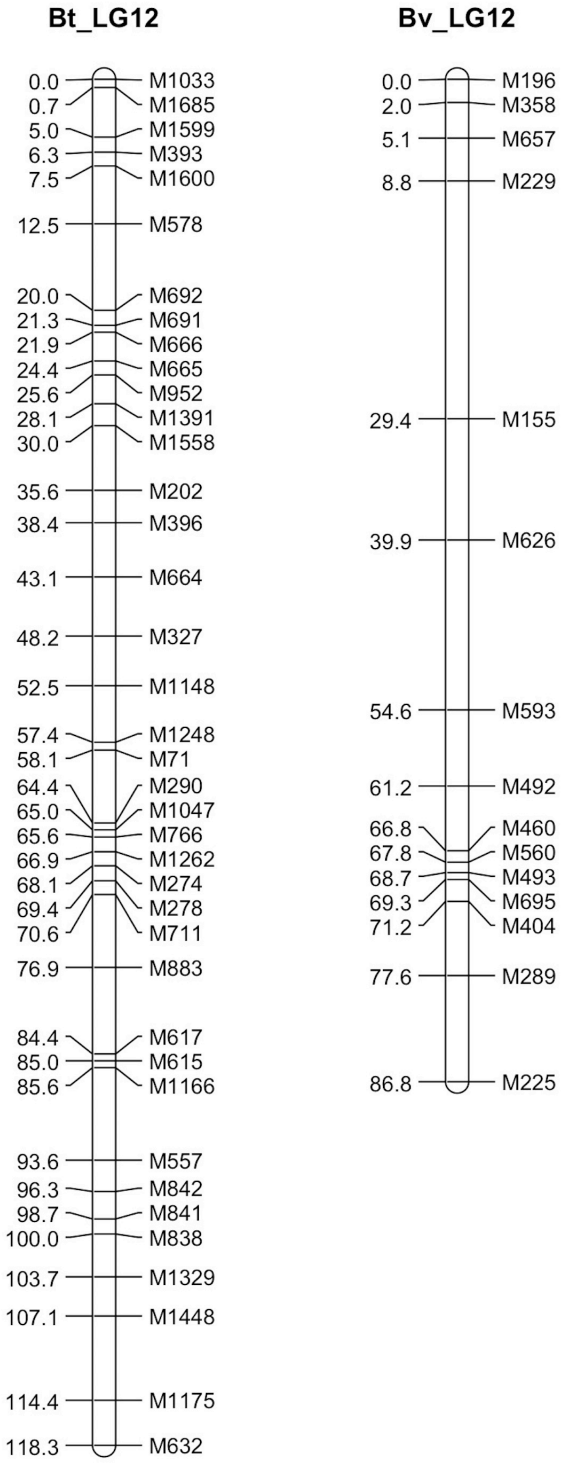
Chromosome 10



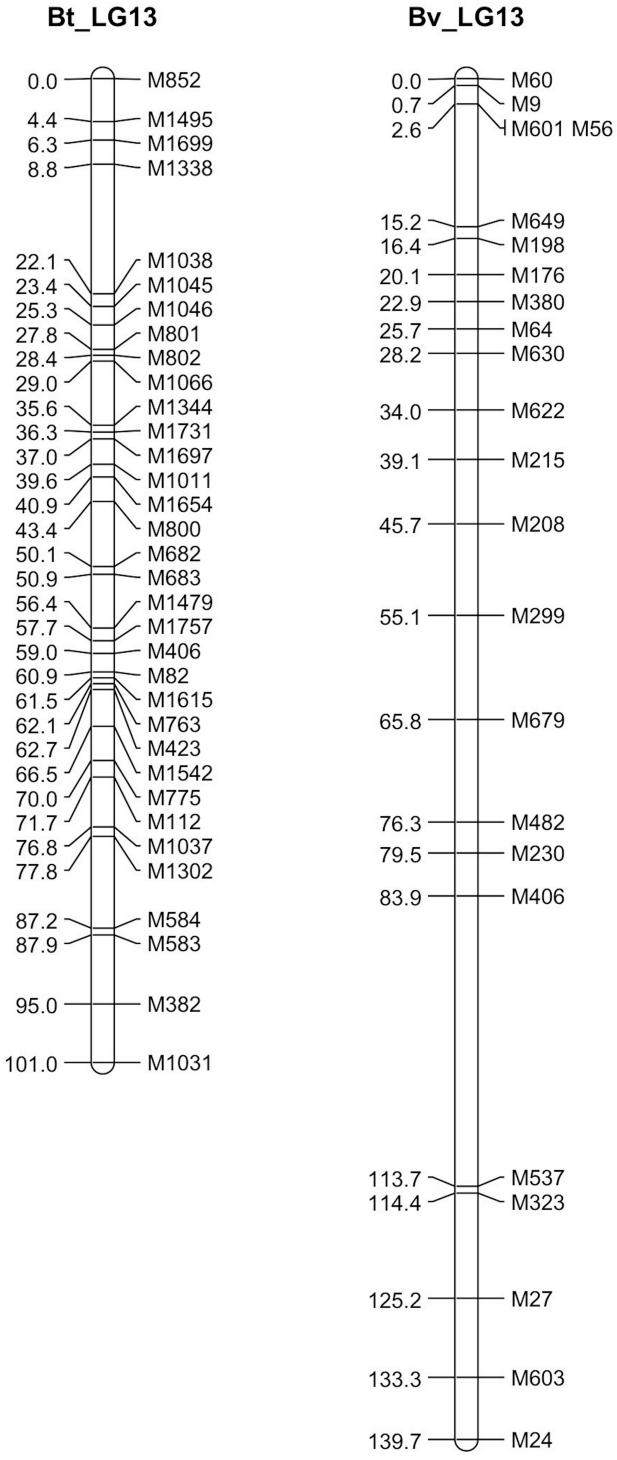
Chromosome 11



Chromosome 12



Chromosome 13



Chromosome 14

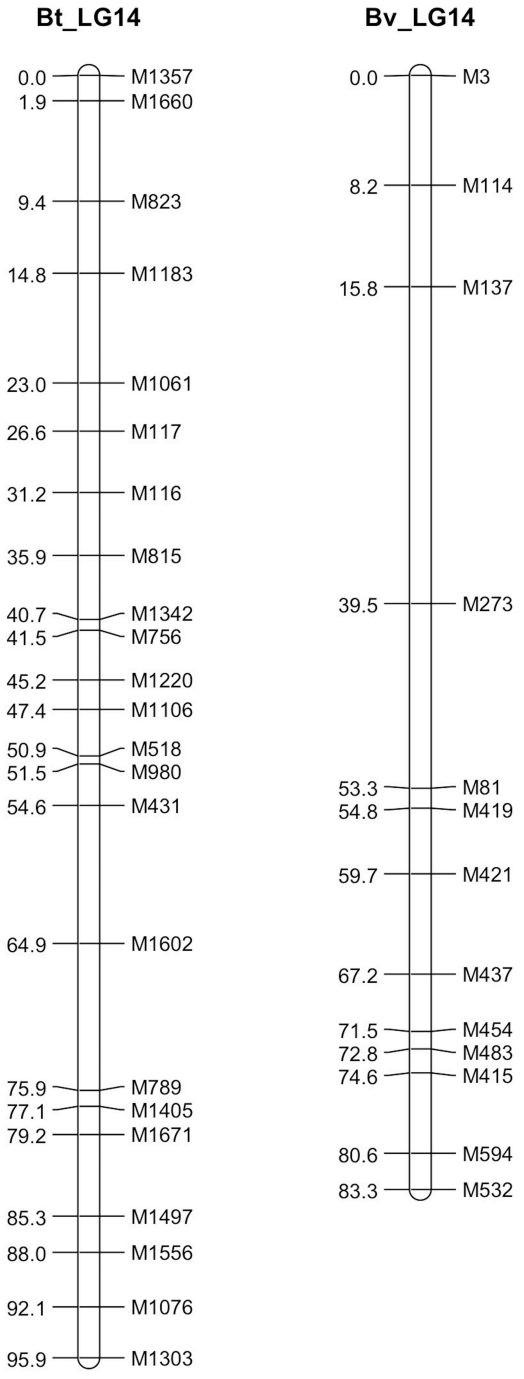


Fig 3 Detailed genetic linkage maps of *B. thunbergii* accession BtUCONN1 and *B. vulgaris* accession Wagon Hill. For each of the 14 chromosomes (paired based on synteny assessed via the Kobold reference assembly), side-by-side ideograms are shown for BtUCONN1 and Wagon Hill. All chromosomes share the same scale (Kosambi cM) and are oriented such that 0 cM corresponds to the telomere of the short-arm.

Summary statistics of the two genetic linkage maps are detailed in Table 4. The BtUCONN1 map consists of 598 recombination bins (mapped loci) and a total length of 1,474 cM. The number of loci in each linkage group varies from 23 (LG14) to 60 (LG7), with an average distance between adjacent loci of 2.6 cM. The Wagon Hill map consists of 347 loci and total length of 1,713.96 cM. The number of loci in each linkage group varies from 25 (LG7) to 76 (LG7), with an average distance between adjacent loci of 5.5 cM. Marker names, alleles, and genetic positions (cM), as well as a color-coded visualization of the recombination events within all members of the mapping population genotypes are provided in Table S3 for BtUCONN1 and Table S4 for Wagon Hill.

All loci were evaluated for segregation distortion and the distribution of distorted loci determined, relative to the maps. For BtUCONN1, a total of 71 mapped loci exhibited significant distortion ($p < 0.01$), and a high proportion of these loci (~86%) were localized to three linkage groups (37% within 63 cM of LG2, 24% within 44 cM of LG3, and 25% within 44 cM of LG11), indicating the possibility of a true biological basis of distortion rather than it being a random artifact. On the other hand, the Wagon Hill map contains comparatively fewer distorted loci (~4%), with no clear pattern of distribution.

Table 4 Comparative summary statistics of the *Berberis thunbergii* 'BtUCONN1' and *B. vulgaris* 'Wagon Hill' linkage maps

Linkage Group	Length (cM)		Number of markers		Number of loci		Mean distance between loci (cM)	
	BtUCONN1	Wagon Hill	BtUCONN1	Wagon Hill	BtUCONN1	Wagon Hill	BtUCONN1	Wagon Hill
1	98.1	130.4	122	63	41	31	2.5	4.3
2	122.1	145.5	178	74	60	37	2.1	4.0
3	119.7	134.1	140	58	50	28	2.4	5.0
4	101.6	143.0	109	54	47	27	2.2	5.5
5	110.8	154.1	139	44	48	28	2.4	5.7
6	112.0	121.4	151	33	50	21	2.3	6.1
7	115.7	135.0	195	76	49	31	2.4	4.5
8	96.8	134.6	148	63	46	32	2.2	4.3
9	102.0	120.2	145	44	34	26	3.1	4.8
10	93.0	121.2	88	25	34	14	2.8	9.3
11	87.4	64.6	110	39	43	21	2.1	3.2
12	118.3	86.8	75	35	39	15	3.1	6.2
13	101.0	139.7	100	52	34	23	3.1	6.3
14	95.9	83.3	55	32	23	13	4.4	6.9
Average	105.3	122.4	125.4	49.4	42.7	24.8	2.6	5.5

Disease phenotyping

To determine disease responses to *Pg*, the parents and all F_1 progeny were inoculated using overwintered telia of *Pg* found on naturally infected *E. repens*. The F_1 progeny segregated into four clear phenotypic classes, ranging from resistant to susceptible (Fig 2, Table 2). Of the 162 F_1 individuals used to build linkage map, phenotype is available for 153 individuals. 25 exhibited a clear resistant reaction similar to that of the *B. thunbergii* parent (Fig 2C) and 61 exhibited a clear susceptible reaction similar to that of the *B. vulgaris* parent (Fig 2F). Of the remaining 67 lines, 38 exhibited moderate resistance (Fig 2D) and 29 exhibited moderate susceptibility (Fig 2E).

QTL analysis

To map regions associated with *Pg*-NHR in *B. thunbergii*, CIM analysis was conducted using the linkage maps of both parents and the 4-point stem rust reaction type described above. According to the LOD threshold score of 3.9 declared via permutation analysis, CIM analysis resulted in the identification of a single significant QTL of peak LOD value 28.2, centered 25 cM from the telomere of *B. thunbergii* chromosome 3 (Fig 4). The flanking markers for this QTL, hereafter referred to as *QPgr-3S* (QTL for *Pg* resistance on the short arm of chromosome 3), were determined via a detailed characterizations of the F_1 individuals with recombination events on either side of peak QTL marker *M1128*. The distal flanking marker *M411* is set by phenotype of the WH15-191, with a recombination event with *M1125*. Similarly, the proximal flanking marker *M969* is set by phenotype of the WH15-101, with a recombination event with *M512*. Both WH15-191 and WH15-101 are clear resistant lines with score of 1. Hence, the *QPgr-3S* region spans approximately 13 cM bounded by markers *M411* and *M969*. No significant QTL

was detected in the *B. vulgaris* map.

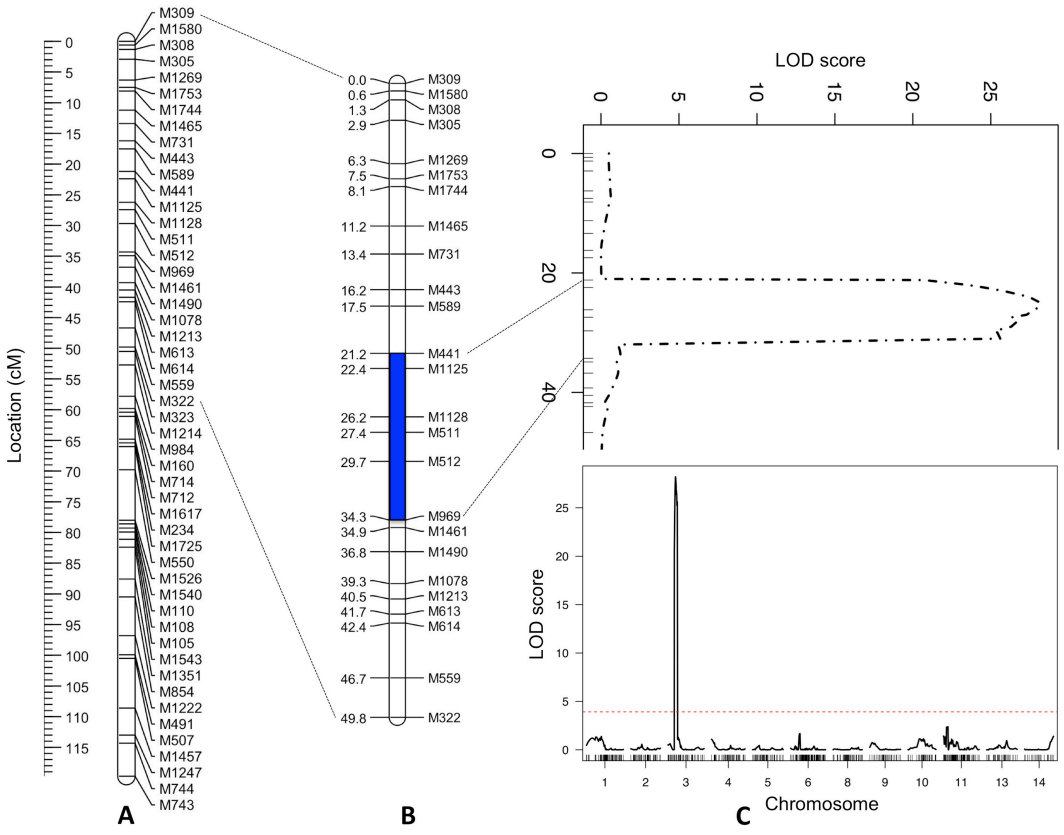


Fig 4 Genetic and physical positions of the *QPgr-3S* region on the short arm of chromosome 3 of *B. thunbergii*. **(A)** Full linkage map of the chromosome 3; **(B)** Enlarged linkage map of the chromosome 3, with the QTL region indicated in blue between its two flanking markers, *M411* and *M969*; **(3)** LOD plot of the QTL region (top) and the context of the single QTL peak across the 14 chromosomes of *B. thunbergii* (bottom). The dotted red line indicates the threshold for QTL significance (LOD = 3.9), determined via permutation analysis.

Draft genome of B. thunbergii cv. 'Kobold'

Approximately 129 Gbp of sequence data was generated from 115 PacBio SMRT cells (P6-C4 chemistry), with an average read length of 10,409 bp and a read length N50 of 15,021 bp (Table S5). The haploid genome size of Kobold was estimated to be 1.37 Gbp based on k-mer analysis and 1.72 Gbp based on flow cytometry, both of which are close to the previously

published *B. thunbergii* haploid genome size (1C) of 1.51 Gb [73]. The FALCON-Unzip pipeline resulted in an assembly of 1.36 Gb, consisting of 4,671 primary contigs and a read length N50 of 669,177 bp (Table 4). The corresponding phased haplotigs numbered 7,144 contigs, with a total length of 0.87 Gb, approximately 64% of the primary contig space. Further curation, in the form of chimera breaking and cryptic haplotig identification, resulted in a final 1.23 Gbp assembly consisting of 2,698 primary contigs with a read length N50 of 0.76 Mbp (Table 5). The number of haplotigs in the final assembly increased to 8,790, with a combined length of 0.99 Gb (>80% of the primary contig space).

Table 5 Summary statistics of the main stages of the *B. thunbergii* cv 'Kobold' genome assembly

Variables	FALCON-Unzip		Final assembly [1]		Hi-C scaffolding [2]
	Primary contigs	Haplotigs	Primary contigs	Haplotigs	
Number of contigs	4,671	7,144	2,698	8,790	15
Total length (Gbp)	1.36	0.88	1.23	0.99	1.23
Longest (Mbp)	8.60	1.49	8.60	1.49	100.80
Shortest (bp)	8,581	561	20,469	561	1,319
> 100 kbp (%)	2,551 (54.6)	2,836 (39.7)	2,229 (82.6)	3126 (35.6)	15 (100)
> 1 Mbp (%)	289 (6.2)	9 (0.1)	289 (10.7)	9 (0.1)	15 (100)
Mean length (Mbp)	0.29	0.12	0.46	0.11	5.99
N50 length (Mbp)	0.67	0.21	0.76	0.19	86.95
GC content (%)	37.6	37.7	37.7	37.7	37.7

[1] After application of the Purge Haplotigs pipeline [52] and manual curation (chimera breaking and haplotig re-assignment)
[2] All statistics for the Hi-C assembly refer to scaffolds rather than contigs

Genome completeness and contamination analyses revealed a final genome assembly of acceptable quality, featuring complete representation of 80.9% of the BUSCO core plant gene set and only 15.1% missing BUSCO genes. While only 83.0% of the BtUCONN1 GBS fragments aligned to the final assembly, 100% of the PacBio preads and 100% of the RNA-seq data (92.2% in proper pair) aligned to the final assembly. After the initial FALCON-Unzip assembly, 119 primary contigs showed significant sequence similarity with plant cpDNA and mtDNA

sequence; but this number dropped to only one primary contig in the final assembly, following haplotig purging and curation.

Although BUSCO analysis represented relatively lower percentage of complete single copy plant genes (80%) in final assembly, 100% alignments of RNA-seq data indicate relative completeness of the final assembly than represented by BUSCO scores. Some BUSCOs classified as missing in the Kobold final genome assembly could be due to *Berberis* spp. being too divergent from model plants used to build BUSCO core genes for plant datasets. Additionally, *Berberis* spp. may have complex gene structures than in model plants that render the gene prediction tool difficult to locate and predict correctly or even partially. Indeed, BUSCO plant gene sets are defined based on 30 model plants that are phylogenetically distantly related from any plant species within order Ranunculaes.

The primary contigs from the final assembly were ordered into chromosome-level scaffolds (pseudo-molecules) on the basis of three-dimensional proximity information obtained via chromosome conformation capture analysis (Hi-C). As shown in the Hi-C heatmap (Fig 5), 97.8% of the primary contigs (2,824 contigs, 1.2 Gbp) successfully assembled into 14 pseudo-molecules representing the 14 chromosomes of *B. thunbergii*. The remaining 2.2% (157 primary contigs, 49.9 Mbp) were assigned as unscaffolded contigs.

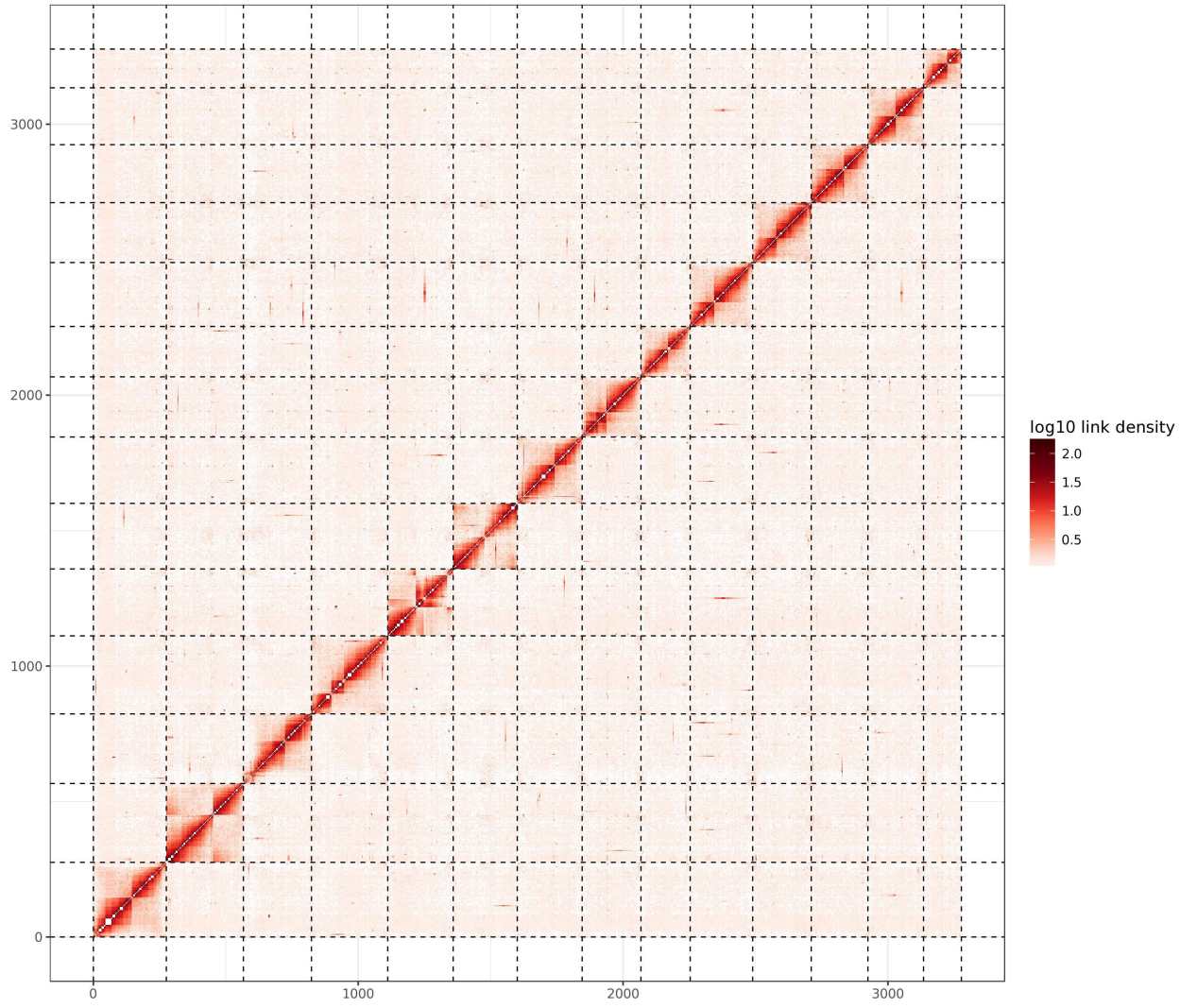


Fig 5 Hi-C post scaffolding Heat Map of *B. thunbergii*. Primary contigs assembled into 14 pseudo-molecules representing the 14 chromosomes of *B. thunbergii*

Anchoring of the genetic linkage maps to the physical assembly

Using BLASTn with MR centroids as queries, the positions of the mapped GBS markers in the final Hi-C assembly was used to anchor the genetic linkage maps of both the parental species to the Kobold physical map. As illustrated in Fig 6 a very high degree of synteny is observed between the two species, with co-linearity to the *B. thunbergii* (Kobold) physical map being 92.9% and 95.1% for the *B. vulgaris* (Wagon Hill) and *B. thunbergii* (BtUCONN1) genetic

linkage maps, respectively. The physical positions of a small percentage of loci in both linkage maps, 3.9% in *B. thunbergii* and 5.1% in *B. vulgaris*, were ambiguous, in that they could not be associated with unique positions in the physical assembly. Another small percentage of loci, 0.93% in *B. thunbergii* and 1.12% in *B. vulgaris*, exhibited unambiguous BLAST hits to different chromosomes than in the linkage map, as indicated by dots in Fig 4. Approximate centromere positions were visually inferred from the Hi-C heatmap plot (Fig 5).

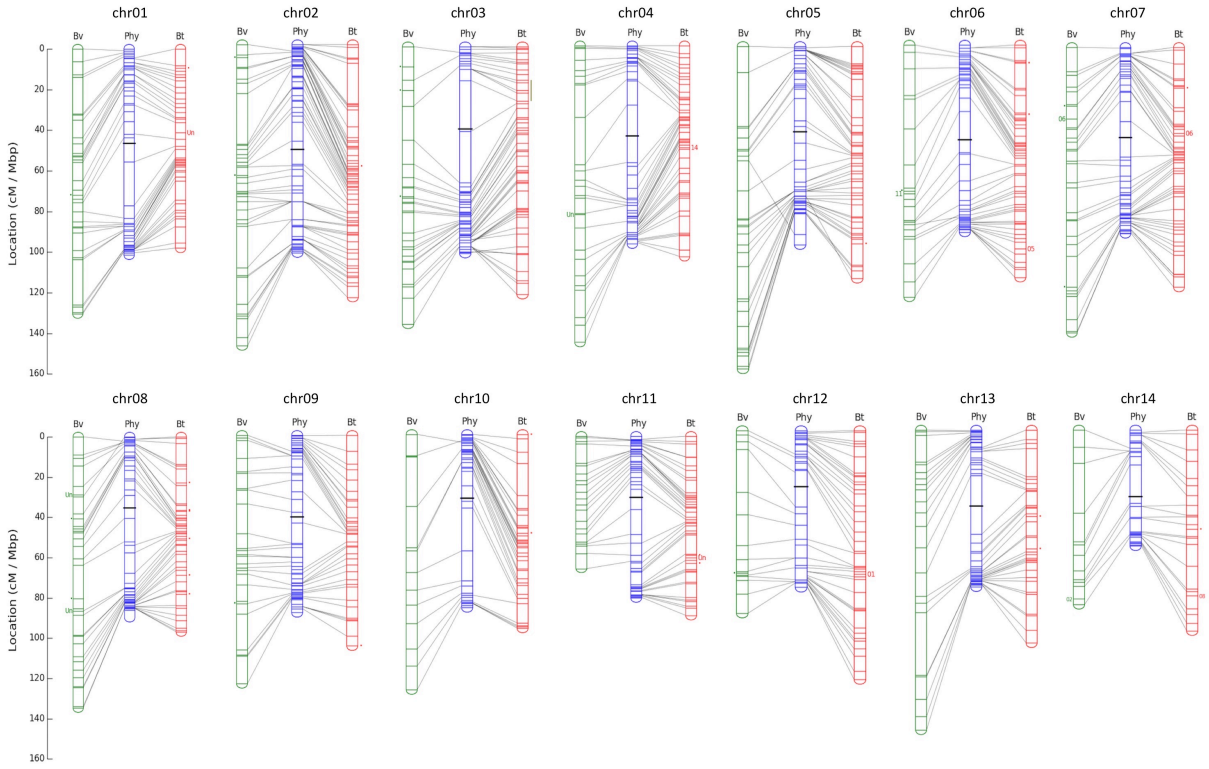


Fig 4 Visualization of synteny of the BtUCONN1 (red) and Wagon Hill (green) genetic maps (cM), via anchoring of GBS centroids to the *B. thunbergii* cv Kobold physical reference (blue; Mbp). The seven GBS markers that BLAST outside their home linkage group are indicated by small numbers (01 – 14) that signify the linkage groups with which they associate. The four GBS markers that BLAST to unscaffolded contigs are indicated by "Un". Small dots beside linkage maps indicate loci with multiple, ambiguous alignments throughout the genome. Bold horizontal black bars on the Kobold physical map indicate approximate centromere positions, based on the Hi-C heatmap. Finally, the grey side bar along the chromosome 3 linkage map for BtUCONN1 indicates the position of *QPgr*-3S.

Transcriptome assembly

A total of 62 Gbp of data, comprised of ~206 million 150-bp PE reads, was obtained by sequencing a library of 10 different tissues from the reference accession Kobold, including immature leaf tissues sampled as various time points following inoculation with *Pg*. Using the Trinity pipeline and the final Kobold assembly as a guide, a 189.3 Mbp transcriptome was assembled, containing 131,407 putative transcripts and 55,186 cDNA sequences (complete ORFs) (see Table 5 for summary statistics). Quality and completeness of the transcriptome assembly were assessed by TransRate and BUSCO. A recent study reports that a TransRate score of 0.22 exceeds 50% of the published *de novo* assembled transcriptomes deposited in the NCBI TSA to date [66]. In comparison, the TransRate score of the Kobold transcriptome is 0.40, indicating its relative quality. Completeness statistics are also acceptable, as indicated by the fact that, of the BUSCO set of 1,440 core plant genes, 1,286 (89.3%) were represented in the transcriptome, of which 501(34.8%) were single copy and 754 (52.4%) were duplicated.

Table 6 Descriptive statistics of the *B. thunbergii* cv. 'Kobold' reference-guided transcriptome assembly

Trinity reference-based assembly results	
Number of transcripts	122,872
Total length (bp)	189,291,041
Mean length (bp)	1,541
Number of ORFs (%)	55,186 (44.28%)
Transcript length N50 (bp)	1,991
GC Content	40.0%

TransRate results	
TransRate score	0.403
TransRate optimal score	0.427
TransRate cutoff	0.037
Number of good contigs (%)	120,972 (98.5%)

BUSCO results	
Complete (%)	1,286 (89.3%)
Complete and single-copy (%)	651 (45.2%)
Complete and duplicated (%)	635 (44.1%)
Fragmented (%)	47 (3.3%)
Missing (%)	107 (7.4%)

Identification of candidate genes

Functional annotation of the *QPgr-3S* region resulted in the identification of 99 high confidence (HC) genes. Of these 99 genes, 61 were annotated based on the reference transcriptome (evidence-based approach) and 38 were annotated based on gene prediction models (*Ab-initio* approach). To identify a short list of candidate genes potentially associated with *Pg*-NHR, the HC genes were further inspected using two different analyses (see Materials and Methods). Time course DGE analysis identified 12 genes that exhibit differential expression 72 hours after *Pg* inoculation (Table S6).

Combined with the underlying linkage evidence from the QTL analysis, the results of the time course DGE analyses elevated the 12 genes identified above to a status

of potential candidates associated with *Pg*-NHR. All of these candidate genes are found on the same primary contig (000400F, 8.6 Kbp). Of these 12 genes, only two simultaneously met all three criteria of linkage, differential expression, and homology to protein families implicated in disease resistance in other systems. As high-priority candidates, these two genes were selected for detailed functional annotation (Table S7). The first gene (*GG9708*) was found to be a leucine-rich repeat belonging to receptor-like kinase (LRR-RLKs), and the second (*GG9868*) a zinc ion binding SSM4 protein, two gene families that have been implicated in durable disease resistance in other systems (Fig S4).

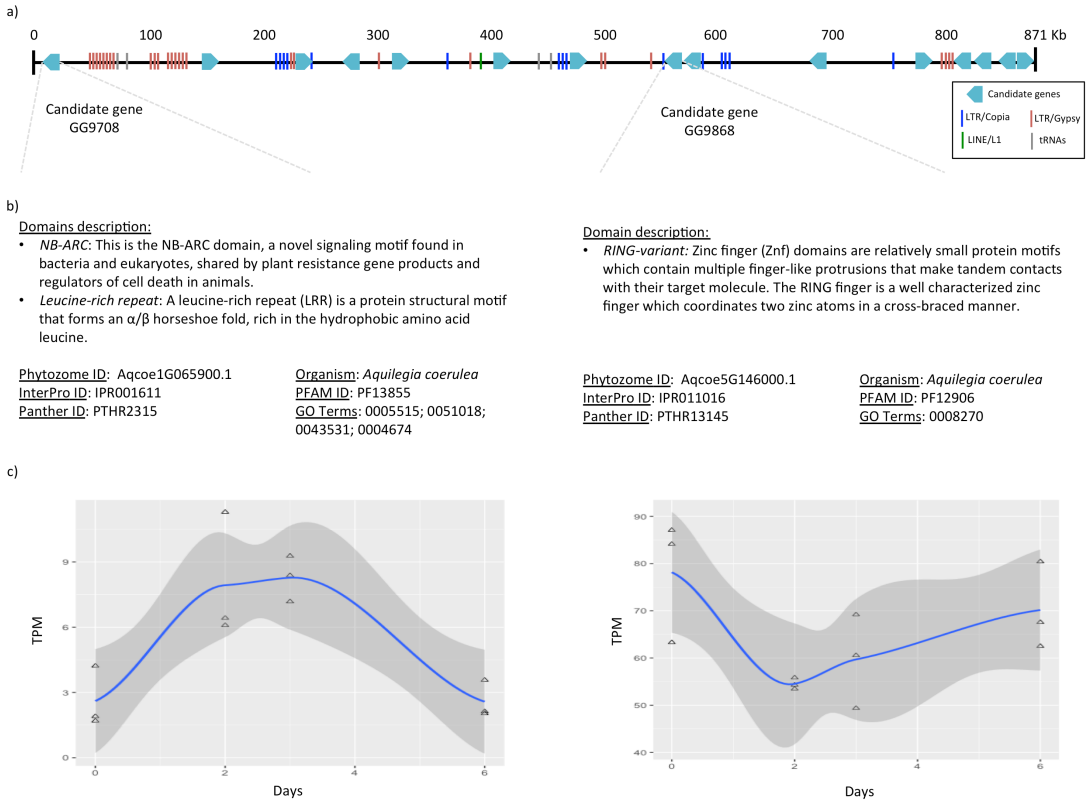


Fig 7 Candidate genes present in contig 000400F. **(A)** All 12 candidate genes were identified in contig 000400F. **(B)** Functional description of 2 high priority candidate genes (*GG9708* and *GG9868*) **(C)** The time course differential gene expression analysis showing the transcript abundance (TPM) across the evaluated time.

Discussion

As the most common form of disease resistance, and one that possesses intrinsic durability, non-host resistance (NHR) presents itself as a compelling strategy for protecting crops from plant diseases. Despite well-articulated visions of such a strategy and the great need for durable disease resistance in many crops, however, the genetic basis of NHR remains poorly understood for many plant-microbe pathosystems, including wheat stem rust. Over the past decade, in light of growing global concern about the ever-evolving cereal rusts, efforts have been mounting to understand NHR to rust pathogens using various model and non-model plants. Many plant species, including *A. thaliana*, *Brachypodium distachyon*, rice, barley, and cowpea [74–79], have been used to study NHR to *P. striiformis* f. sp. *tritici* (*Pst*), the causal organism of wheat stripe rust. In contrast, NHR to the wheat stem rust pathogen *P. graminis* has thus far been studied only in rice [74], as distinct from the studies of intermediate *Pg* resistance conducted in barley and *B. distachyon* [14,15].

As the only globally important small grain that is immune (i.e. a non-host) to all known rust diseases, rice (*Oryza* spp.) is an attractive source for genes conferring durable rust resistance. Genetic studies for *Pg*-NHR in rice are difficult, however, precisely because all individuals exhibit resistance to the pathogen. Although studies have shown some limited progression of Pg infection in rice, thus raising the possibility of dissecting *Pg*-NHR in that system [74], variation in the infection process is small, requires tedious microscopic studies, and ultimately fails to complete.

Inheritance and mapping studies which make use of interspecific crosses (host × non-host) present an alternative means of probing the genetics of *Pg*-NHR. Developing such mapping populations can be challenging, however, as interspecific crosses

frequently suffer from sterility, abnormal growth, and poor seed viability [80]. Fortunately, a host and non-host species within *Berberis*, the ancestral host genus of *Pg*, successfully hybridize, making the development of an F_1 interspecific hybrid mapping population possible. In this study, a controlled cross between *Pg* susceptible *B. vulgaris* and *Pg* non-host *B. thunbergii* was done to study the inheritance of the gene(s) underlying the putative *Pg*-NHR of *B. thunbergii*. To support this work, foundational genetic and genomic resources were developed to help establish the *Berberis*-*Pg* pathosystem as a viable research model for studying *Pg*-NHR.

Genetic and genomic resource development

A high-quality genetic linkage map is required for quantitative trait loci (QTL) mapping, but the relatively long generation time of some perennial plant species, like barberries, precludes the efficient generation of inbred lines, thus impeding map construction. Despite its many non-model attributes, however, barberry displays a high level of heterozygosity due to its outcrossing nature. And interspecific hybrids obtained by crossing highly heterozygous species often display sufficient segregation to allow genetic map construction in an F_1 population. In this study, using a pseudo test cross strategy, high-quality linkage maps were developed for both *B. vulgaris* and *B. thunbergii* from a single F_1 population. As a result of the stringent quality filters applied to the set of *de novo* GBS markers, nearly 100% of the markers were placed successfully in the linkage maps of the two species. Although flow-cytometry analysis indicates comparable genome sizes between the two parents (1.72 Gb vs.1.69 Gb), the total length of the BtUCONN1 (*B. thunbergii*) linkage map is roughly 15% smaller than that of the Wagon

Hill (*B. vulgaris*) map (1,474 cM vs. 1,714 cM). This incongruity with the expected physical genome sizes is likely due to the significantly fewer markers available for the *B. vulgaris* map as compared to those available for the *B. thunbergii* map (706 vs 1,757). Low marker density often results in inflation of linkage maps [81], so it is expected that additional markers may act to reduce the overall recombination length of *B. vulgaris* map. The significantly lower number of markers available for the *B. vulgaris* map is likely a result of the relatively lower level of heterozygosity in this species as a result of the severe genetic bottleneck presumed during its colonial introduction from Europe to North America [16].

Despite this deviation from expected overall similarity in length, the two linkage maps developed in this study, the first of their kind for any species within the plant order Ranunculales, represent a solid contribution to the field. The relatively even distribution of markers across the 14 chromosomes of both species permits initial QTL analysis of acceptable resolution, with approximately 30% and 20% of the inter-marker distances being less than 1 cM for *B. thunbergi*i and *B. vulgaris*, respectively. In addition, the strong synteny observed between the two maps gives further evidence of their reliability.

As a complement to genetic resources like mapping populations and linkage maps, a high-quality reference genome can serve as an invaluable resource in dissecting QTLs, identifying underlying candidate genes, and facilitating their detailed characterization. In this study, cutting-edge sequencing and scaffolding technologies were used to develop a highly contiguous *de novo* reference genome of *B. thunbergii*. Using PacBio SMRT sequencing and chromosome conformation capture data, a 1.2 Gb haploid assembly of *B. thunbergii* cv. 'Kobold' was successfully assembled into 14

chromosome-scale pseudo-molecules. As with the linkage maps, this reference is the first of its kind for a member of both the *Berberidaceae* family as well as the order *Ranunculales* more broadly. Given the previous lack of molecular resources for barberries, the reference assembled in this study exemplifies the power of recent technologies to make rapid progress even in non-model systems and establishes a benchmark for the *de novo* assembly of a highly heterozygous plant species with a moderately sized genome.

In conclusion, the development of foundational genetic and genomic resources, including a genotyped interspecific mapping population, linkage maps for its two parental species, a chromosome-scale reference genome, and a multiple-tissue transcriptome establishes *Berberis* spp. as a viable research model for studying *Pg*-NHR. Furthermore, such resources promise to facilitate related endeavors, including global rust surveillance work and ornamental horticulture breeding.

QPgr-3S *and the identification of candidate genes for* Pg-*NHR*

The long-term goal of this research is to identify candidate gene(s) governing *Pg*-NHR in *B. thunbergii*. As an initial step in that direction, the genetic and genomic resources developed here enabled the identification of a single QTL of large effect (LOD > 28) on the short arm of chromosome 3 of *B. thunbergii* (Fig 3). This 13 cM QTL region, dubbed *Qpgr-3S*, was found to span 10 physical contigs and contain a total of 99 high-confidence genes. Of these, 12 were identified as candidate genes and short-listed as relatively high priority for follow up studies, including one (GG9708) exhibiting homology to a leucine-rich repeat containing receptor-like kinase and the other (GG9868) to a zinc ion binding SSM4 protein.

Receptor-like kinases (RLKs) in plants are a large superfamily of proteins involved in a diverse array of plant responses, including development, growth, hormone perception, and recognition of and response to pathogens [82]. Although new data are highlighting different classes of RLKs, many known defense-related RLKs belong to the leucine-rich repeat (LRR) subclass; and RLKs have been identified to regulate both plant innate immune response and R-gene mediated pathogen specific responses. The gene *GG9708* identified in this study exhibits homology to the LRR subclass of RLKs.

The current model of plant NHR suggests that plant immune responses can be broadly grouped in two major classes, namely those triggered by pathogen-associated molecular patterns (PAMP-triggered immunity-PTI) and those triggered by pathogens effectors (effector-triggered immunity- ETI) [83,84]. PTI in immune plants is often recognized via a receptor kinase located in the plant plasma membrane, whereas ETI invokes intracellular NB-LRR proteins after detecting actions or structures of pathogen effectors [85]. PAMPs, upon recognition by so-called pattern-recognition receptors (PRRs), are known to activate RLKs [86–88]. As a specific example, Rajaraman et al. (2016) identified a LRR-malectin domain-containing transmembrane RLK that appears to mediate intermediate NHR of barley to the non-adapted wheat powdery mildew fungus *Blumeria graminis* f.sp. *tritici* [89]. Similarly, the barley *Rpg1* gene, known to confer durable resistance to the stem rust pathogen, also encodes a receptor-like kinase protein [90,91]. Thus, in addition to other evidence (linkage and DGE), the functional annotation of GG9708 underscores its status as a noteworthy candidate gene, perhaps triggering the *Pg*-NHR response in *B. thunbergii* via PTI.

The functional annotation of *GG9868* also supports its status as a candidate gene of interest. Mechanical barriers to foliar pathogen penetration, mostly consisting of the waxy cuticle present on a leaf's surface, are considered an important factor in many forms of NHR [92]. Such barriers are often the first line of defense against pathogen attack, and potential pathogens must overcome them for the successful colonization of the host. Zinc finger transcription factors have been implicated in the regulation of wax biosynthesis [92]; and a gene affecting rust germ tube differentiation has been identified to encode a zinc finger transcription factor in alfalfa [92]. The identification of zinc ion binding protein as one of the candidate genes suggests that pre-invasive penetration barrier upon pathogen attack may also trigger NHR response in *B. thunbergii*.

The identification of both the *QPgr-3S* region and a pair of high-priority candidate genes demonstrates the utility of the genetic and genomic resources developed in the study to probe the genes underlying *Pg*-NHR exhibited by *B. thunbergii*. Such results, however, are but the first step toward identifying the genes governing *Pg*-NHR; and further work is required to validate and dissect the QTL region, in addition to testing candidate gene hypotheses.

Possible modes of inheritance of* Pg-*NHR

From the practical standpoint of breeding for improved resistance to wheat stem rust, the central questions regarding *Pg*-NHR concern the nature and modes of inheritance of the underlying genes. As previously observed in a natural interspecific barberry hybrid population [16], F_1 interspecific hybrids exhibit a range of reactions to *Pg*, from fully resistant to fully susceptible, with many intermediate forms. This range of

reactions was similarly observed in the F_1 mapping population used in this study (Figure 2C-2F and Table 2), suggesting that *Pg*-NHR in *B. thunbergii* is most likely polygenic. Polygenic NHR has been suggested in other studies as well, including rice NHR to wheat stem rust, barley NHR to powdery mildews, barley NHR to oat stem rust, and barley NHR to other non-adapted rust species [14,93,94].

If indeed *QPgr-3S* is implicated in *Pg*-NHR, the data suggest that its underlying gene(s) are necessary but not sufficient for resistance. In other words, this study at most provides a first insight into a larger gene network regulating *Pg*-NHR in *B. thunbergii*. Indeed, in light of the lack of segregation in the non-host parental species *B. thunbergii*, the segregation of resistance among F_1 hybrids suggests the possible existence of some critical gene(s), by definition fixed within the *B. thunbergii* genepool, upstream of *QPgr-3S*. Because of their fixed state within *B. thunbergii*, such gene(s) cannot be mapped in an F_1 population; but if recessive, their single dosage in an F_1 would permit susceptibility to *Pg*, thus allowing the detection of background resistance genes, (e.g. *QPgr-3S*). In all likelihood, then, *QPgr-3S* is not a critical region conferring *Pg*-NHR but is rather a region contributing to *Pg* resistance. Strategic crosses among the F_1 progeny and/or backcrosses to *B. thunbergii* will be necessary to identify those critical gene(s) regulating *Pg*-NHR in *B. thunbergii*; and the current work demonstrates the feasibility of such studies.

Data availability

All raw sequence data and final assemblies (genome and transcriptome) generated in this study are available through the NCBI database. The parsed, high-quality GBS data

generated for the two parental lines and the 182 F_1 genotypes are available through the NCBI Short Read Archive, with SRA ID's provided in Supplementary Data 1. RNA-seq data from the ten *B. thunbergii* cv. 'Kobold' tissues used for DGE and transcriptome assembly are linked to NCBI BioProject PRJNA478022. The assembled transcriptome itself is available under TSA ID GGRA00000000, and the final Kobold genome assembly is deposited in NCBI under accession number QNQO00000000.

Supplementary Materials

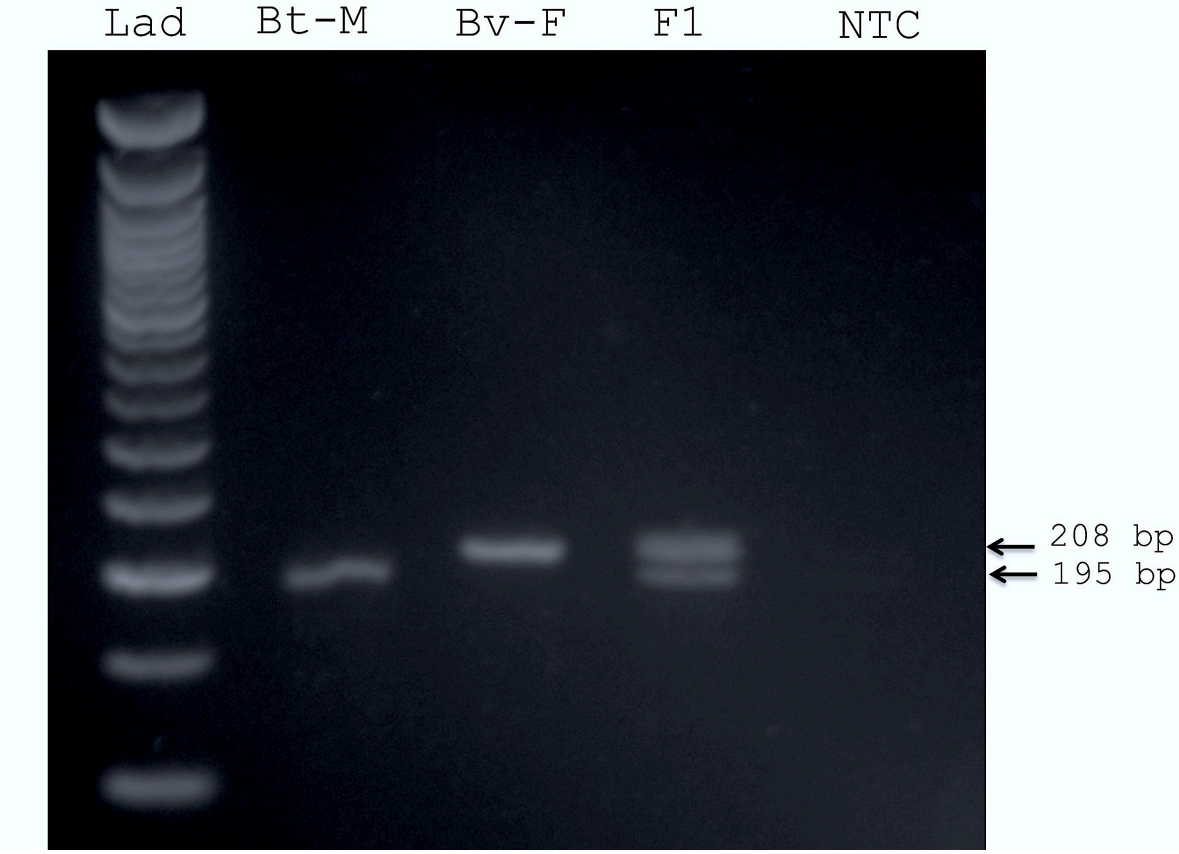


Fig S1 Gel image for F_1 mapping population validation. The F_1 status of a putative hybrid individual was considered validated if both bands from the two parental species were detected.

Text S1 Cluster sequences and primers information for PCR based markers used for validation of F_1 status of mapping population. **A)** 222 bp sequence of *Berberis thunbergii*; Deleted sequence is highlighted in green. **B)** 235 bp sequence of *B. vulgaris*. **C)** Alignment of *B. vulgaris* and *B. thunbergii* cluster. D) Primer and final PCR product information

A.

>B.thunbergii_Cluster45383_222bp

AATAATAGCTCCCTGATTGGGGCTCATTATCAGTCGCCATGTTTAGAATCCCGAGCAAGAAGCTCGGGAAAAATCAGAA
GCACATGCAATAGATAGAAGAAGCAAAGAGGTATTCTTCAAAGCTCCAAATCCCCTAAAAAAAACGTGGTCGGTTCTGC
TCGGCTAGGTTGCTTCAAATATAGCATTTCAGCTCGGAATTCCTCACTTTTCTTCTCAGCTCCG

B.

>B.vulgaris_Cluster45383_235bp

AATAATAGCTCCCTGATTGGGGCTCATTATCAGTCGCCATGTTTAGAATCCCGAGCAAGAAGCTCGGGAAAAATCAGAA
GCACATGCAATAGATAGAAGAAG**CAAAGAGATCGAT**CAAAGAGGTATTCTTCAAAGCTCCAAATCCCCTAAAAAAAACG
TGGTCGGTTCTGCTCGGCTAGGTTGCTTCAAATATAGCATTTCAGCTCGGAATTCCTCACTTTTCTTCTCAGCTCCG

C.

```
Bv      AATAATAGCTCCCTGATTGGGGCTCATTATCAGTCGCCATGTTTAGAATCCCGAGCAAGA   1
Bt      ..................................................
Bv      AGCTCGGGAAAAATCAGAAGCACATGCAATAGATAGAAGAAGCAAAGAG-----------   61
Bt      ..............................................ATCGATCAAAG
Bv      --GTATTCTTCAAAGCTCCAAATCCCCTAAAAAAAACGTGGTCGGTTCTGCTCGGCTAGG   121
Bt      AG..................................................
Bv      TTGCTTCAAATATAGCATTTCAGCTCGGAATTCCTCACTTTTCTTCTCAGCTCCG        181
Bt      ..................................................
```

D.

Species	Primer ID	Primer sequence	Annealing Tm	Band size (bp)
B. thunbergii	Cluster45383_UF	CCTGATTGGGGCTCATTATC	52 ºC	195
B. vulgaris	Cluster45383_UR	AGTGAGGAATTCCGAGCTGA		208

Text S2 Detailed record of the GBS-SNP-CROP v.3.0 command lines used in this study, including all specified pipeline parameters.

GBS-SNP-CROP-1.pl
perl /path-to-workdir/GBS-SNP-CROP-1.pl -d PE -b barcodesIDs.txt -fq L001 -s 1 -e 48 -enz1 TGCA -enz2 CGG

GBS-SNP-CROP-2.pl
perl /path-to-workdir/GBS-SNP-CROP-2.pl -d PE -fq L001 -t 10 -ph 33 -ad TruSeq3-PE.fa:2:30:10 -l 30 -sl 4:30 -tr 30 -m 32

GBS-SNP-CROP-3.pl
perl /path-to-workdir/GBS-SNP-CROP-3.pl -d PE -b barcodesIDs.txt -fq L001

GBS-SNP-CROP-4.pl
perl /path-to-workdir/GBS-SNP-CROP-4.pl -d PE -b barcodeID.txt -rl 150 -pl 32 -p 0.01 -id 0.93 -t 10 -MR MockRefName

GBS-SNP-CROP-5.pl
perl /path-to-workdir/GBS-SNP-CROP-5.pl -d PE -b barcodeID.txt -ref MockRefName.MockRef.Genome.fasta -Q 30 -q 0 -f 2 -F 2308 -t 10 -Opt 0

GBS-SNP-CROP-6.pl
perl /path-to-workdir/GBS-SNP-CROP-6.pl -b barcodeID.txt -out SNPs.summary.txt

GBS-SNP-CROP-7.pl
perl /path-to-workdir/GBS-SNP-CROP-7.pl -in SNPs.summary.txt -out SNPs.call.txt -mnHoDepth0 5 -mnHoDepth1 20 -mnHetDepth 3 -altStrength 0.962 -mnAlleleRatio 0.25 -mnCall 0.75 -mnAvgDepth 7 -mxAvgDepth 200

```
# GBS-SNP-CROP-8.pl
perl /path-to-workdir/GBS-SNP-CROP-8.pl -in SNPs.call.txt -out SNP.Rmatrix -b
barcodesIDs.txt -formats R
```

Text S3 Complete details of the FALCON assembly parameters used in this study.

```
[General]
# list of files of the initial fasta
input_fofn = input.fofn

input_type = raw
#input_type = preads

#openending = True

stop_all_jobs_on_failure = False

# The length cutoff used for seed reads used for initial mapping
length_cutoff = 5000

genome_size = 1400000000
seed_coverage = 30

# The length cutoff used for seed reads usef for pre-assembly
length_cutoff_pr = 9000

sge_option_da = -pe smp 5 -q bigmem
sge_option_la = -pe smp 20 -q bigmem
sge_option_pda = -pe smp 6 -q bigmem
sge_option_pla = -pe smp 16 -q bigmem
sge_option_fc = -pe smp 24 -q bigmem
sge_option_cns = -pe smp 12 -q bigmem

pa_concurrent_jobs = 96
cns_concurrent_jobs = 96
ovlp_concurrent_jobs = 96


pa_HPCdaligner_option =  -v -B128 -M32 -e.70 -l4800 -s100 -k18 -h480 -w8
ovlp_HPCdaligner_option = -v -B128 -M32 -h1024 -e.96 -l2400 -s100 -k18


pa_DBsplit_option = -a -x500 -s400

ovlp_DBsplit_option = -s400

falcon_sense_option = --output_multi --min_idt 0.70 --min_cov 2 --max_n_read 200 --n_core 8
falcon_sense_skip_contained = True

overlap_filtering_setting = --max_diff 85 --max_cov 87 --min_cov 2 --n_core 12
```

Text S4 Complete details of script used for purging halpotigs. This code is available at https://bitbucket.org/mroachawri/purge_haplotigs

Text S5 Complete details of parameters used for quantifying transcript and the sleuth R code for the time course differential expression analysis.

Kallisto command lines for quantify transcript abundance. The quant function was performed for all seven RNA-seq libraries:

```
$ kallisto index -i Ber_ALL7.idx ALL7.TrinityGG.Assembly.fa
$ kallisto quant -t 20 -i Ber_ALL7.idx -o IM0_R1_KOUT \
/path/to/dir/IM0-1_PE_R1.fq
/path/to/dir/IM0-1_PE_R1.fq
$ kallisto quant -t 20 -i Ber_ALL7.idx -o IM0_R2_KOUT \
/path/to/dir/IM0-2_PE_R1.fq
/path/to/dir/IM0-2_PE_R1.fq
$ kallisto quant -t 20 -i Ber_ALL7.idx -o IM0_R3_KOUT \
/path/to/dir/IM0-3_PE_R1.fq
/path/to/dir/IM0-3_PE_R1.fq
```

Sleuth R command lines for a time course analysis:

```
library("sleuth")
sample_id = dir(file.path(".", "kallistoOUT"))
kal_dirs = file.path(".", "KallistoOUT", sample_id, "kallisto")
s2c = read.table(file.path(".", "hiseq_info.txt"), header = T, stringsAsFactors=F)
s2c = dplyr::select(s2c, sample = sample, condition)
s2c = dplyr::mutate(s2c, path = kal_dirs)
so = sleuth_prep(s2c, extra_bootstrap_summary=T,read_bootstrap_tpm=T)
so = sleuth_fit(so, ~condition, 'full')
so = sleuth_fit(so, ~1, 'reduced')
so = sleuth_lrt(so, 'reduced', 'full')
sleuth_table = sleuth_results(so, 'reduced:full', 'lrt', show_all = F)
sleuth_significant = dplyr::filter(sleuth_table, qval <= 0.01)
wtest = sleuth_wt(so, 'conditionTime0', 'full')
swbeta = sleuth_lrt(wtest, 'reduced',  'full')
sleuth_live(swbeta)
```

Table S1 *B. thunbergii* cv 'Kobold' tissues used for transcriptome assembly. Immature leaf tissue were collected at four time points (before and after *Pg* inoculation, were collected from a clonally propagated plant of *B. thunbergii* cv 'Kobold'

Tissue ID	Tissue	Time after *P. graminis* inoculation (hrs)
IM0	Immature leaf	0
IM48	Immature leaf	48
IM72	Immature leaf	72
IM144	Immature leaf	144
ML	Mature leaf	--
AM	Apical meristem	--
YS	Young stem	--
RO	Root	--
FR	Fruit	--
FL	Flowers	--

Table S2 Table of the *Berberis thunbergii* accession 'BtUCONN1', *B. vulgaris* accession 'Wagon Hill', and the interspecific F_1 mapping population used in the study. For each accession, the following information is provided: 1) Accession ID, 2) GBS library ID, 3) GBS barcode, 4) Assigned NCBI Sequence Read Archive (SRA) number, 5) Number of high-quality paired-end (PE) reads used for variant calling, 6)% of missing data

Accession ID	GBS Library ID	GBS Barcode	SRA Number	PE reads	% Missing Data
BtUCONN1	Lib4_01	TGACGCCA	SRR5712434	2,790,916	28.40
Wagon Hill	Lib2_86	TAGCAG	SRR5712049	4,007,578	12.81
WH15-001	Lib3_01	TGACGCCA	SRR7450012	2,112,038	41.68
WH15-002	Lib3_02	CAGATA	SRR7450011	5,751,546	5.03
WH15-003	Lib3_03	GAAGTG	SRR7450018	3,388,764	26.48
WH15-004	Lib4_02	CAGATA	SRR7450107	2,842,666	6.89
WH15-005	Lib3_04	TAGCGGAT	SRR7450017	4,446,406	8.45
WH15-006	Lib3_05	TATTCGCAT	SRR7450033	3,349,136	8.33
WH15-007	Lib3_06	ATAGAT	SRR7450032	5,045,010	6.40
WH15-008	Lib3_07	CCGAACA	SRR7450031	4,525,094	6.57
WH15-009	Lib3_08	GGAAGACAT	SRR7450110	4,920,230	6.49
WH15-010	Lib4_03	GAAGTG	SRR7450071	3,295,594	7.53
WH15-011	Lib3_09	GGCTTA	SRR7450008	5,759,290	5.89
WH15-012	Lib3_10	AACGCACATT	SRR7450007	4,065,098	6.85
WH15-013	Lib3_11	GAGCGACAT	SRR7450053	1,554,146	47.19
WH15-014	Lib4_04	TAGCGGAT	SRR7450070	2,003,012	17.82

WH15-015	Lib3_12	CCTTGCCATT	SRR7450052	5,659,240	5.96
WH15-016	Lib3_13	GGTATA	SRR7450055	2,517,512	37.56
WH15-018	Lib3_14	TCTTGG	SRR7450054	5,408,654	5.85
WH15-019	Lib3R_41	ACCAGGA	SRR7450037	248,778	5.61
WH15-020	Lib3_15	GGTGT	SRR7450057	6,678,362	5.68
WH15-021	Lib3_16	GGATA	SRR7450056	6,628,450	5.66
WH15-022	Lib3_17	CTAAGCA	SRR7450059	1,369,444	56.34
WH15-023	Lib3_18	ATTAT	SRR7450058	2,281,894	14.20
WH15-024	Lib3R_19	GCGCTCA	SRR7450088	365,442	52.83
WH15-025	Lib3_20	ACTGCGAT	SRR7450050	1,823,152	35.27
WH15-026	Lib3_21	TTCGTT	SRR7450049	1,227,128	33.92
WH15-027	Lib3_22	ATATAA	SRR7450027	2,341,268	39.82
WH15-028	Lib4_05	TATTCGCAT	SRR7450036	1,079,506	69.82
WH15-030	Lib3_23	TGGCAACAGA	SRR7450028	2,713,452	8.24
WH15-031	Lib3_24	CTCGTCG	SRR7450029	3,051,676	8.55
WH15-032	Lib3_25	GCCTACCT	SRR7450030	3,861,554	6.77
WH15-033	Lib3R_42	CCACTCA	SRR7450040	47,664	11.19
WH15-034	Lib3_26	CACCA	SRR7450023	3,777,170	8.00
WH15-035	Lib3_27	AATTAG	SRR7450024	3,381,012	8.29
WH15-036	Lib4_06	ATAGAT	SRR7449972	3,622,674	7.53
WH15-037	Lib3_28	GGAACGA	SRR7450025	1,126,232	68.77
WH15-038	Lib4_07	CCGAACA	SRR7450085	3,294,350	8.19
WH15-039	Lib3_29	ACAACT	SRR7450026	2,237,874	36.38
WH15-040	Lib4_08	GGAAGACAT	SRR7450082	2,550,346	14.03
WH15-041	Lib3_30	ACTGCT	SRR7450020	5,120,752	5.44
WH15-042	Lib3_31	CGTGGACAGT	SRR7450021	3,213,812	7.03
WH15-043	Lib4_09	GGCTTA	SRR7450087	2,997,460	7.60
WH15-044	Lib3_32	TGGCACAGA	SRR7449998	1,226,120	60.28
WH15-045	Lib3_33	TGCTT	SRR7449997	3,105,110	33.54
WH15-046	Lib3_34	GCAAGCCAT	SRR7449996	2,400,666	37.51
WH15-047	Lib4_10	AACGCACATT	SRR7450086	3,262,280	9.88
WH15-048	Lib3_35	CGCACCAATT	SRR7449995	1,832,180	24.66
WH15-049	Lib4_11	GAGCGACAT	SRR7450081	3,129,446	9.29
WH15-050	Lib3_37	AACTGG	SRR7450002	1,214,340	59.30
WH15-051	Lib3_38	ATGAGCAA	SRR7450001	1,170,156	65.38
WH15-052	Lib4_12	CCTTGCCATT	SRR7450080	2,755,322	12.50
WH15-053	Lib3_39	CTTGA	SRR7450000	5,082,016	5.54
WH15-054	Lib3_40	GCGTCCT	SRR7449999	2,500,852	28.39
WH15-055	Lib4_13	GGTATA	SRR7450123	2,826,948	8.04
WH15-056	Lib4_14	TCTTGG	SRR7450124	2,880,728	10.71
WH15-057	Lib3R_43	TCACGGAAG	SRR7450034	504,342	73.23
WH15-058	Lib4_15	GGTGT	SRR7450125	3,871,918	8.28
WH15-059	Lib4_16	GGATA	SRR7450126	3,668,444	7.09

WH15-060	Lib4_17	CTAAGCA	SRR7450127	610,598	56.78
WH15-061	Lib4_18	ATTAT	SRR7450128	776,324	65.32
WH15-062	Lib4_19	GCGCTCA	SRR7450129	1,176,104	37.64
WH15-063	Lib3R_47	CTCTA	SRR7450141	230,072	15.20
WH15-064	Lib4_20	ACTGCGAT	SRR7450130	1,322,092	31.40
WH15-065	Lib3_49	CTCTCGCAT	SRR7450004	4,724,580	7.27
WH15-066	Lib3_50	CAGAGGT	SRR7450003	2,593,728	26.14
WH15-067	Lib4_21	TTCGTT	SRR7450131	1,658,562	21.84
WH15-068	Lib3R_44	TATCA	SRR7450035	322,004	2.44
WH15-069	Lib3_51	GCGTACAAT	SRR7449979	2,488,068	36.07
WH15-070	Lib4_22	ATATAA	SRR7450132	959,976	56.59
WH15-071	Lib3_52	ACGCGCG	SRR7449980	9,830,004	5.45
WH15-072	Lib4_23	TGGCAACAGA	SRR7450153	968,652	49.78
WH15-073	Lib3_53	GTCGCCT	SRR7449977	5,544,474	5.94
WH15-074	Lib4_24	CTCGTCG	SRR7450152	374,958	83.23
WH15-075	Lib3_54	AATAACCAA	SRR7449978	2,536,380	29.21
WH15-076	Lib3R_45	TAGCCAA	SRR7449981	611,760	61.70
WH15-077	Lib3_55	AATGAACGA	SRR7449975	7,813,208	6.26
WH15-078	Lib3_56	CGTCGCCACT	SRR7449976	3,174,022	11.11
WH15-079	Lib4_96	CGTTCA	SRR7450115	2,839,052	7.38
WH15-080	Lib3_58	GAAGCA	SRR7449973	8,406,004	5.24
WH15-081	Lib3_59	AACGTGCCT	SRR7449974	3,493,722	11.58
WH15-082	Lib3R_46	ATATCGCCA	SRR7449994	244,814	5.15
WH15-083	Lib4_25	GCCTACCT	SRR7450151	2,769,448	21.89
WH15-084	Lib4_26	CACCA	SRR7450150	1,594,186	35.02
WH15-085	Lib3R_48	GGTGCACATT	SRR7450051	229,258	47.35
WH15-086	Lib3_60	CCTCG	SRR7449982	4,394,588	25.11
WH15-087	Lib3_61	CTCAT	SRR7449983	3,045,094	27.45
WH15-088	Lib3_62	ACGGTACT	SRR7450138	3,509,716	36.59
WH15-089	Lib3_63	GCGCCG	SRR7450137	2,323,742	36.73
WH15-090	Lib4_27	AATTAG	SRR7450149	1,621,560	49.77
WH15-091	Lib4_28	GGAACGA	SRR7450148	2,313,742	13.12
WH15-092	Lib4_29	ACAACT	SRR7450147	3,225,278	9.40
WH15-093	Lib4_30	ACTGCT	SRR7450146	3,453,890	10.47
WH15-094	Lib4_31	CGTGGACAGT	SRR7450145	2,483,696	15.15
WH15-096	Lib4_32	TGGCACAGA	SRR7450144	2,819,988	14.31
WH15-098	Lib4_34	GCAAGCCAT	SRR7449988	3,158,644	9.46
WH15-099	Lib3_65	TCCGAG	SRR7450140	1,672,242	48.55
WH15-100	Lib4_35	CGCACCAATT	SRR7449989	1,974,742	23.71
WH15-101	Lib3_66	TAGATGA	SRR7450139	4,757,720	6.42
WH15-102	Lib3_67	TGGCCAG	SRR7450134	5,428,268	6.41
WH15-103	Lib3R_66	TAGATGA	SRR7450106	571,520	59.52
WH15-104	Lib3_68	GCACGAT	SRR7450133	3,309,024	36.90

WH15-105	Lib3_69	TTGCTG	SRR7450136	2,870,162	43.83
WH15-106	Lib3_70	CGCAACCAGT	SRR7450135	1,005,144	62.00
WH15-107	Lib3_71	TCACTG	SRR7450143	1,606,654	58.45
WH15-108	Lib3_72	ACAGT	SRR7450142	1,889,622	52.26
WH15-109	Lib3_73	GGAGTCAAG	SRR7450113	4,858,370	7.57
WH15-110	Lib3_74	TGAAT	SRR7450116	6,207,404	6.39
WH15-111	Lib4_37	AACTGG	SRR7449986	2,913,062	8.10
WH15-112	Lib3_75	CATAT	SRR7450122	2,286,546	47.14
WH15-113	Lib3_76	GTGACACAT	SRR7450117	7,132,616	16.98
WH15-114	Lib3_77	TATGT	SRR7450118	3,406,668	43.75
WH15-115	Lib3_78	CAGTGCCATT	SRR7450119	4,697,626	7.51
WH15-116	Lib3_79	ACAACCAACT	SRR7450120	5,118,394	27.09
WH15-117	Lib3_80	TGCAGA	SRR7450121	5,167,950	19.39
WH15-118	Lib3_81	CATCTGCCG	SRR7450111	5,493,618	17.98
WH15-119	Lib4_38	ATGAGCAA	SRR7449987	3,065,592	11.71
WH15-120	Lib3_82	GGACAG	SRR7450112	2,840,562	45.71
WH15-121	Lib4_39	CTTGA	SRR7449992	3,991,212	7.55
WH15-122	Lib3_83	ATCTGT	SRR7450097	3,243,778	22.30
WH15-123	Lib4_40	GCGTCCT	SRR7449993	2,628,172	14.48
WH15-125	Lib3_84	AAGACGCT	SRR7450096	2,966,272	30.45
WH15-126	Lib4_41	ACCAGGA	SRR7449990	2,941,652	14.60
WH15-127	Lib4_42	CCACTCA	SRR7449991	2,080,476	33.29
WH15-128	Lib4_43	TCACGGAAG	SRR7449984	3,454,326	9.95
WH15-129	Lib4_44	TATCA	SRR7449985	1,826,916	18.32
WH15-130	Lib4_45	TAGCCAA	SRR7450010	3,460,470	9.62
WH15-131	Lib4_46	ATATCGCCA	SRR7450009	2,477,832	13.60
WH15-132	Lib4_47	CTCTA	SRR7450022	1,337,624	46.23
WH15-133	Lib3_85	GAATGCAATA	SRR7450095	4,046,612	22.06
WH15-135	Lib4_48	GGTGCACATT	SRR7450019	2,468,640	26.36
WH15-136	Lib4_49	CTCTCGCAT	SRR7450014	2,477,022	17.00
WH15-137	Lib4_50	CAGAGGT	SRR7450013	4,132,720	6.70
WH15-138	Lib4_51	GCGTACAAT	SRR7450016	3,029,696	9.46
WH15-140	Lib3_87	ATCCG	SRR7450094	5,166,780	10.17
WH15-141	Lib3_88	CTTAG	SRR7450093	5,671,298	5.47
WH15-142	Lib3_89	TTATTACAT	SRR7450092	2,341,538	31.74
WH15-143	Lib4_52	ACGCGCG	SRR7450015	1,598,226	26.81
WH15-144	Lib3_90	GCCAACAAG A	SRR7450091	5,124,162	9.49
WH15-147	Lib3_93	CAACCACACA	SRR7450090	3,537,756	10.17
WH15-148	Lib3_94	GCTCCGA	SRR7450089	2,533,302	34.33
WH15-149	Lib4_53	GTCGCCT	SRR7450006	2,366,658	14.42
WH15-150	Lib4_54	AATAACCAA	SRR7450005	1,954,346	18.63
WH15-151	Lib4_55	AATGAACGA	SRR7450045	2,467,150	15.62

WH15-152	Lib4_56	CGTCGCCACT	SRR7450046	2,709,674	10.36
WH15-153	Lib4_57	ATGGCAA	SRR7450047	2,960,330	9.90
WH15-155	Lib4_58	GAAGCA	SRR7450048	3,374,758	8.53
WH15-156	Lib4_59	AACGTGCCT	SRR7450041	2,127,672	36.16
WH15-157	Lib4_60	CCTCG	SRR7450042	1,060,136	37.04
WH15-158	Lib4_61	CTCAT	SRR7450043	3,079,852	12.54
WH15-159	Lib4_62	ACGGTACT	SRR7450044	2,593,134	31.17
WH15-160	Lib4_63	GCGCCG	SRR7450038	1,787,808	26.72
WH15-161	Lib4_64	CAAGT	SRR7450039	3,486,468	17.54
WH15-162	Lib4_65	TCCGAG	SRR7450065	2,139,246	17.86
WH15-163	Lib4_66	TAGATGA	SRR7450064	3,736,462	8.35
WH15-164	Lib4_67	TGGCCAG	SRR7450063	3,155,444	7.99
WH15-165	Lib4_68	GCACGAT	SRR7450062	3,940,062	6.63
WH15-166	Lib4_69	TTGCTG	SRR7450069	3,836,980	6.78
WH15-167	Lib4_70	CGCAACCAGT	SRR7450068	2,423,538	10.80
WH15-168	Lib4_71	TCACTG	SRR7450067	2,856,424	8.41
WH15-169	Lib4_72	ACAGT	SRR7450066	2,303,656	12.69
WH15-170	Lib4_73	GGAGTCAAG	SRR7450061	3,335,196	8.56
WH15-171	Lib4_74	TGAAT	SRR7450060	3,356,992	7.14
WH15-172	Lib4_75	CATAT	SRR7450078	4,391,264	6.90
WH15-173	Lib4_76	GTGACACAT	SRR7450079	4,404,624	6.89
WH15-174	Lib4_77	TATGT	SRR7450076	3,884,856	6.32
WH15-175	Lib4_78	CAGTGCCATT	SRR7450077	3,621,816	8.16
WH15-176	Lib4_79	ACAACCAACT	SRR7450074	3,059,892	12.99
WH15-177	Lib4_80	TGCAGA	SRR7450075	3,906,346	7.86
WH15-178	Lib4_81	CATCTGCCG	SRR7450072	2,538,896	11.43
WH15-179	Lib4_82	GGACAG	SRR7450073	3,243,944	7.97
WH15-180	Lib4_83	ATCTGT	SRR7450083	2,384,508	12.19
WH15-181	Lib4_84	AAGACGCT	SRR7450084	2,246,822	19.55
WH15-182	Lib4_85	GAATGCAATA	SRR7450103	3,182,898	11.10
WH15-183	Lib4_86	TAGCAG	SRR7450102	3,073,906	8.04
WH15-184	Lib4_87	ATCCG	SRR7450105	2,084,460	12.45
WH15-185	Lib4_88	CTTAG	SRR7450104	3,944,864	6.22
WH15-186	Lib4_89	TTATTACAT	SRR7450099	2,138,918	16.72
WH15-187	Lib4_90	GCCAACAAG A	SRR7450098	3,007,726	10.88
WH15-188	Lib4_91	TGCCGCAT	SRR7450101	2,444,658	12.48
WH15-189	Lib4_92	CGTGTCA	SRR7450100	2,372,566	11.19
WH15-190	Lib4_93	CAACCACACA	SRR7450109	2,313,580	14.07
WH15-192	Lib4_94	GCTCCGA	SRR7450108	1,154,644	44.34
WH15-193	Lib4_95	TCAGAGAT	SRR7450114	3,018,696	13.08

Table S3 Linkage map of *B. thunbergii* accession 'BtUCONN1' and associated information. Marker names, alleles, genetic positions (cM), and color-coded visualization of the recombination events within all members of the mapping population genotypes are provided. This file is available at https://unh.box.com/s/ogzs6zvrxlczsmsg6h8t7y6gzuvqckga

Table S4 Linkage map of *B. vulgaris* accession 'Wagon Hill' and associated information. Marker names, alleles, genetic positions (cM) and color-coded visualization of the recombination events within all members of the mapping population genotypes are provided. This file is available at https://unh.box.com/s/ogzs6zvrxlczsmsg6h8t7y6gzuvqckga

Table S5 Summary table of PacBio raw data obtained by sequencing 116 SMRT cells

	Basepair (bp_
Total sequence amount	128,875,283,508
Number of reads	12,069,440
Median read length	10,630
Mean read length	10,409
N75 read length	18,777
N50 read length	15,021
N25 read length	11,266
Maximum read length	46,433

Table S6 Maker functional annotation features for the 12 candidate genes for time course differential gene expression. The two high priority candidate genes are highlighted in dark green. This file is available at https://unh.box.com/s/ogzs6zvrxlczsmsg6h8t7y6gzuvqckga

Table S7 Phytozome-based detailed annotation of the two high priority candidate genes associated with the *Pg* resistance in *B. thunbergii*, describing a detailed functional annotation as well the associated reference ID's for different databases. This file is available at https://unh.box.com/s/ogzs6zvrxlczsmsg6h8t7y6gzuvqckga

References

1. Leonard KJ, Szabo LJ. Stem rust of small grains and grasses caused by *Puccinia graminis*. Mol Plant Pathol. 2005;6:99–111.

2. Singh RP, Hodson DP, Jin Y, Lagudah ES, Ayliffe MA, Bhavani S, et al. Emergence and spread of new races of wheat stem rust fungus: continued threat to food security and prospects of genetic control. Phytopathology. 2015;105:872–884.

3. Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Njau P, Wanyera R, et al. Will stem rust destroy the world's wheat crop? Adv Agron. 2008;98:271–309.

4. Peterson PD, Leonard KJ, Roelfs AP, Sutton TB. Effect of barberry eradication on changes in populations of *Puccinia graminis* in Minnesota. Plant Dis. 2005;89:935–940.

5. Jin Y, Szabo LJ, Pretorius ZA, Singh RP, Ward R, Fetch Jr T. Detection of virulence to resistance gene *Sr24* within race TTKS of *Puccinia graminis* f. sp. *tritici.* Plant Dis. 2008;92:923–926.

6. Pretorius ZA, Singh RP, Wagoire WW, Payne TS. Detection of virulence to wheat stem rust resistance gene *Sr31* in *Puccinia graminis.* f. sp. *tritici* in Uganda. Plant Dis. 2000;84:203–203.

7. Singh RP, Hodson DP, Huerta-Espino J, Jin Y, Bhavani S, Njau P, et al. The emergence of Ug99 races of the stem rust fungus is a threat to world wheat production. Annu Rev Phytopathol. 2011;49:465–481.

8. Lewis CM, Persoons A, Bebber DP, Kigathi RN, Maintz J, Findlay K, et al. Potential for re-emergence of wheat stem rust in the United Kingdom. Commun Biol. 2018;1:13.

9. Ayliffe M, Singh R, Lagudah E. Durable resistance to wheat stem rust needed. Curr Opin Plant Biol. 2008;11:187–192.

10. Lipka U, Fuchs R, Kuhns C, Petutschnig E, Lipka V. Live and let die–*Arabidopsis* nonhost resistance to powdery mildews. Eur J Cell Biol. 2010;89:194–199.

11. Thordal-Christensen H. Fresh insights into processes of nonhost resistance. Curr Opin Plant Biol. 2003;6:351–357.

12. Mysore KS, Ryu C-M. Nonhost resistance: how much do we know? Trends Plant Sci. 2004;9:97–104.

13. Ayliffe M, Jin Y, Kang Z, Persson M, Steffenson B, Wang S, et al. Determining the basis of nonhost resistance in rice to cereal rusts. Euphytica. 2011;179:33–40.

14. Dracatos PM, Ayliffe M, Khatkar MS, Fetch Jr T, Singh D, Park RF. Inheritance of prehaustorial resistance to *Puccinia graminis* f. sp. *avenae* in barley (*Hordeum vulgare* L.). Mol Plant Microbe Interact. 2014;27:1253–1262.

15. Figueroa M, Alderman S, Garvin DF, Pfender WF. Infection of *Brachypodium distachyon* by formae speciales of *Puccinia graminis*: early infection events and host-pathogen incompatibility. PLOS One. 2013;8:e56857.

16. Bartaula R, Melo AT, Connolly BA, Jin Y, Hale I. An interspecific barberry hybrid enables genetic dissection of non-host resistance to the stem rust pathogen *Puccinia graminis*. J Exp Bot. 2018;69:2483–2493.

17. Jin Y. Role of Berberis spp. as alternate hosts in generating new races of *Puccinia graminis* and *P. striiformis.* Euphytica. 2011;179:105–108.

18. Zhao J, Wang L, Wang Z, Chen X, Zhang H, Yao J, et al. Identification of eighteen *Berberis* species as alternate hosts of *Puccinia striiformis* f. sp. *tritici* and virulence variation in the pathogen isolates from natural infection of barberry plants in China. Phytopathology. 2013;103:927–934.

19. Peterson PD. "The Barberry or Bread": The Public Campaign to Eradicate Common Barberry in the United States in the Early 20th Century. Available from: https://www.apsnet.org/publications/apsnetfeatures/Pages/Barberry.aspx

20. Hale IL, Connolly BA, Bartaula R. The Occurrence of Hybrid Barberry, *Berberis* ×*ottawensis* (Berberidaceae), in New Hampshire and Rhode Island. Rhodora. 2015;117:384–7.

21. Bettgenhaeuser J, Gilbert B, Ayliffe M, Moscou MJ. Nonhost resistance to rust pathogens–a continuation of continua. Front Plant Sci. 2014;5:664.

22. Huihui L, Hearne S, Banziger M, Li Z, Wang J. Statistical properties of QTL linkage mapping in biparental genetic populations. 2010.

23. Snape JW, Moore G. Reflections and opportunities: gene discovery in the complex wheat genome. Wheat Prod Stress Environ. Springer. 2007; p. 677–684.

24. Consortium IWGS. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science. 2014;345:1251788.

25. Boyko EV, Gill KS, Mickelson-Young L, Nasuda S, Raupp WJ, Ziegle JN, et al. A high-density genetic linkage map of *Aegilops tauschii*, the D-genome progenitor of bread wheat. Theor Appl Genet. 1999;99:16–26.

26. Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, et al. A high-density genetic map of hexaploid wheat (*Triticum aestivum L.*) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet. 2005;110:865–880.

27. Xue S, Zhang Z, Lin F, Kong Z, Cao Y, Li C, et al. A high-density intervarietal map of the wheat genome enriched with markers derived from expressed sequence tags. Theor Appl Genet. 2008;117:181–189.

28. Somers DJ, Isaac P, Edwards K. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). Theor Appl Genet. 2004;109:1105–1114.

29. Lambert KA, Obae SG. Molecular characterisation of *Berberis thunbergii* cultivars using microsatellite markers. J Hortic Sci Biotechnol. 2016;91:156–160.

30. Côté M-J, Leduc L. Molecular identification of Japanese barberry (*Berberis thunbergii*) cultivars using amplified fragment length polymorphism. HortScience. 2007;42:478–482.

31. Lubel JD, Brand MH, Lehrer JM. AFLP identification of *Berberis thunbergii* cultivars, inter-specific hybrids, and their parental species. J Hortic Sci Biotechnol. 2008;83:55–63.

32. Lubell JD, Brand MH, Lehrer JM, Holsinger KE. Amplified Fragment Length Polymorphism and parentage analysis of a feral barberry (*Berberis thunberg*ii DC.) population to determine the contribution of an ornamental landscape genotype. Hortscience. 2009;44:392–395.

33. Ali S, Hodson D. Wheat Rust Surveillance: Field Disease Scoring and Sample Collection for Phenotyping and Molecular Genotyping. Wheat Rust Dis. Springer; 2017. p. 3–11.

34. CADIC A. Breeding for ever-red barberries (Berberis spp.). ISHS Int Symp Sel Breed Woody Ornam 320. 1992; p. 85–90.

35. Doyle JJ. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–15.

36. Poland JA, Brown PJ, Sorrells ME, Jannink J-L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLOS One. 2012;7:e32253.

37. Melo AT, Bartaula R, Hale I. GBS-SNP-CROP: a reference-optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. BMC Bioinformatics. 2016;17:29.

38. Margarido GRA, Souza AP, Garcia AAF. OneMap: software for genetic mapping in outcrossing species. Hereditas. 2007;144:78–79.

39. Grattapaglia D, Sederoff R. Genetic linkage maps of Eucalyptus grandis and Eucalyptus urophylla using a pseudo-testcross: mapping strategy and RAPD markers. Genetics. 1994;137:1121–1137.

40. Melo AT, Guthrie RS, Hale I. GBS-Based Deconvolution of the Surviving North American Collection of Cold-Hardy Kiwifruit (*Actinidia* spp.) Germplasm. PLOS One. 2017;12:e0170580.

41. Van Os H, Stam P, Visser RG, Van Eck HJ. SMOOTH: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. Theor Appl Genet. 2005;112:187–194.

42. Kosambi DD. The estimation of map distances from recombination values. DD Kosambi. Springer; 2016. p. 125–130.

43. Voorrips RE. MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002;93:77–78.

44. Broman KW, Wu H, Sen Ś, Churchill GA. R/qtl: QTL mapping in experimental crosses. Bioinformatics. 2003;19:889–890.

45. Broman KW, Wu H, Churchill G, Sen S, Yandell B. qtl: Tools for Analayzing QTL Experiments. R Package Version. 2010;1–15.

46. Haley CS, Knott SA. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity. 1992;69:315–324.

47. Department of Agriculture. Black Stem Rust; Identification Requirements and Addition of Rust Resistant Varieties. Federal Register , No. 36. 67.

48. Stoffel K, van Leeuwen H, Kozik A, Caldwell D, Ashrafi H, Cui X, et al. Development and application of a 6.5 million feature Affymetrix Genechip® for massively parallel discovery of single position polymorphisms in lettuce (*Lactuca* spp.). BMC Genomics. 2012;13:185.

49. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, et al. Phased diploid genome assembly with single-molecule real-time sequencing. Nat Methods. 2016;13:1050.

50. Chikhi R, Medvedev P. Informed and automated k-mer size selection for genome assembly. Bioinformatics. 2013;30:31–37.

51. Doležel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. Nat Protoc. 2007;2:2233.

52. Roach MJ, Schmidt SA, Borneman AR. Purge Haplotigs: Synteny Reduction for Third-gen Diploid Genome Assemblies. bioRxiv. 2018;286252.

53. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–842.

54. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.

55. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–1075.

56. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31:3210–3212.

57. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2012;41:D590–D596.

58. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol. 2013;31:1119.

59. Dudchenko O, Shamim MS, Batra S, Durand NC, Musial NT, Mostofa R, et al. The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. bioRxiv. 2018;254797.

60. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–3402.

61. Li H. BFC: correcting Illumina sequencing errors. Bioinformatics. 2015;31:2885–2887.

62. MacManes MD. Establishing evidenced-based best practice for the de novo assembly and evaluation of transcriptomes from non-model organisms. bioRxiv. 2016;035642.

63. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–2120.

64. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644.

65. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8:1494.

66. Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. TransRate: reference-free quality assessment of de novo transcriptome assemblies. Genome Res. 2016;26:1134–1144.

67. Chen N. Using Repeat Masker to identify repetitive elements in genomic sequences. Curr Protoc Bioinforma. 2004;5:4–10.

68. Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 2008;18:188–196.

69. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34:525.

70. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods. 2017;14:687.

71. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2011;40:D1178–D1186.

72. Consortium U. UniProt: the universal protein knowledgebase. Nucleic Acids Res. 2016;45:D158–D169.

73. Rounsaville TJ, Ranney TG. Ploidy levels and genome sizes of *Berberis* L. and Mahonia Nutt. species, hybrids, and cultivars. HortScience. 2010;45:1029–1033.

74. Ayliffe M, Devilla R, Mago R, White R, Talbot M, Pryor A, et al. Nonhost resistance of rice to rust pathogens. Mol Plant Microbe Interact. 2011;24:1143–1155.

75. Cheng Y, Zhang H, Yao J, Wang X, Xu J, Han Q, et al. Characterization of non-host resistance in broad bean to the wheat stripe rust pathogen. BMC Plant Biol. 2012;12:96.

76. Cheng Y, Zhang H, Yao J, Han Q, Wang X, Huang L, et al. Cytological and molecular characterization of non-host resistance in *Arabidopsis thaliana* against wheat stripe rust. Plant Physiol Biochem. 2013;62:11–18.

77. An T, Cai Y, Zhao S, Zhou J, Song B, Bux H, et al. *Brachypodium distachyon* T-DNA insertion lines: a model pathosystem to study nonhost resistance to wheat stripe rust. Sci Rep. 2016;6:25510.

78. Dawson AM, Ferguson JN, Gardiner M, Green P, Hubbard A, Moscou MJ. Isolation and fine mapping of *Rps6:* an intermediate host resistance gene in barley to wheat stripe rust. Theor Appl Genet. 2016;129:831–843.

79. Li K, Hegarty J, Zhang C, Wan A, Wu J, Guedira GB, et al. Fine mapping of barley locus *Rps6* conferring resistance to wheat stripe rust. Theor Appl Genet. 2016;129:845–859.

80. Atienza SG, Jafary H, Niks RE. Accumulation of genes for susceptibility to rust fungi for which barley is nearly a nonhost results in two barley lines with extreme multiple susceptibility. Planta. 2004;220:71–79.

81. Di Pierro EA, Gianfranceschi L, Di Guardo M, Koehorst-van Putten HJ, Kruisselbrink JW, Longhi S, et al. A high-density, multi-parental SNP genetic map on apple validates a new mapping approach for outcrossing species. Hortic Res. 2016;3:16057.

82. Goff KE, Ramonell KM. The role and regulation of receptor-like kinases in plant defense. Gene Regul Syst Biol. 2007;1:117762500700100020.

83. Chisholm ST, Coaker G, Day B, Staskawicz BJ. Host-microbe interactions: shaping the evolution of the plant immune response. Cell. 2006;124:803–814.

84. Jones JD, Dangl JL. The plant immune system. Nature. 2006;444:323.

85. Ronald PC, Beutler B. Plant and animal sensors of conserved microbial signatures. science. 2010;330:1061–1064.

86. Gómez-Gómez L, Boller T. FLS2: an LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol Cell. 2000;5:1003–1011.

87. Krol E, Mentzel T, Chinchilla D, Boller T, Felix G, Kemmerling B, et al. Perception of the *Arabidopsis* danger signal peptide 1 involves the pattern recognition receptor *AtPEPR1* and its close homologue *AtPEPR2*. J Biol Chem. 2010;285:13471–13479.

88. Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones JD, Boller T, et al. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. Cell. 2006;125:749–760.

89. Rajaraman J, Douchkov D, Hensel G, Stefanato FL, Gordon A, Ereful N, et al. An LRR/malectin receptor-like kinase mediates resistance to non-adapted and adapted powdery mildew fungi in barley and wheat. Front Plant Sci. 2016;7:1836.

90. Brueggeman R, Drader T, Kleinhofs A. The barley serine/threonine kinase gene *Rpg1* providing resistance to stem rust belongs to a gene family with five other members encoding kinase domains. Theor Appl Genet. 2006;113:1147–1158.

91. Nirmala J, Drader T, Lawrence PK, Yin C, Hulbert S, Steber CM, et al. Concerted action of two avirulent spore effectors activates Reaction to *Puccinia graminis 1* (*Rpg1*)-mediated cereal stem rust resistance. Proc Natl Acad Sci. 2011;108:14676–14681.

92. Lee H-A, Lee H-Y, Seo E, Lee J, Kim S-B, Oh S, et al. Current understandings of plant nonhost resistance. Mol Plant Microbe Interact. 2017;30:5–15.

93. Romero CC, Vermeulen JP, Vels A, Himmelbach A, Mascher M, Niks RE. Mapping resistance to powdery mildew in barley reveals a large-effect nonhost resistance QTL. Theor Appl Genet. 2018;1–15.

94. Jafary H, Szabo LJ, Niks RE. Innate nonhost immunity in barley to different heterologous rust fungi is controlled by sets of resistance genes with different and overlapping specificities. Mol Plant Microbe Interact. 2006;19:1270–1279.

CHAPTER III

SUMMARY AND FUTURE DIRECTIONS

Summary of work to date

Wheat stem rust, caused by the fungal pathogen *Puccinia graminis* (*Pg*), is one of the most agriculturally important diseases in the world; and many of the rust resistance genes deployed in improved wheat varieties have been overcome since strategic breeding began in the early 20th century. The overarching goal of the research presented in this dissertation was to search for novel sources of durable resistance to the stem rust pathogen. Toward that end, the mechanism of the non-host resistance (NHR) in Pg's alternate (sexual) host, barberry, was investigated, with the hope of identifying a novel source of durable resistance beyond the *Triticum* genepool.

The research began with an investigation of whether or not the interspecific hybrid *Berberis ×ottawensis*, obtained via an intercross between *Pg* non-host *B. thunbergii* and *Pg*-susceptible *B. vulgaris*, can be utilized to dissect the genetic mechanism(s) of Pg-NHR exhibited by *B. thunbergii*. For this Berberis-Pg pathosystem to be useful for genetic studies of NHR, segregation for Pg resistance must be observable with a hybrid population. To determine if the hybrid satisfied that minimum criterion, a natural population of *B. ×ottawensis* in western Massachusetts was surveyed, propagated, and screened for reaction to *Pg*. The results of this study demonstrated that the Pg-NHR observed in *B. thunbergii* indeed segregates in a population of first-generation (F_1) interspecific hybrids; therefore, the gene(s) underlying *Pg*-NHR in *B. thunbergii* are in theory mappable in an F_1 population derived from the controlled hybridization of *B. vulgaris* and *B. thunbergii*.

Given the confirmed segregation of Pg-NHR in an F1 mapping population, the scope of the research broadened. To map the gene(s) underlying Pg-NHR in *B. thunbergii*, an F_1 bi-parental population was developed via a controlled cross between *B. vulgaris* accession Wagon Hill and *B. thunbergii* accession BtUCONN1. Using markers identified between the parents via a

novel genotyping-by-sequencing pipeline which I helped develop, the first-ever genetic linkage maps were constructed for the two parental species. Subsequent QTL analysis led to the identification of a single QTL of large effect, dubbed *QPgr-3S*, on the short arm of chromosome 3 of *B. thunbergii*

To facilitate both dissection of the 13 cM *QPgr*-3S region as well as eventual candidate gene postulation, a 1.2 Gbp chromosome-scale reference genome of *B. thunbergii cv.* 'Kobold' was assembled using long-read PacBio SMRT sequencing and three-dimensional proximity information obtained from Hi-C. To support the functional annotation of this reference and investigate differential gene expression under *Pg* challenge, a transcriptome of the reference accession was also assembled from a multiple tissue RNA library. The reference genome generated in this work represents the first genome sequence available not only within the *Berberidaceae* family but for the entire plant order *Ranunculales*. More broadly, the genetic and genomic resources developed in this study lay the groundwork for the *Berberis*-*Pg* pathosystem as a viable research system for dissecting Pg-NHR while simultaneously providing valuable resources for global rust surveillance work and ornamental horticulture breeding.

Combining the results of the QTL analysis with those of functional annotation, differential gene expression analysis, 12 candidate genes were identified for ongoing investigation of which two genes, *GG9708* and *GG9868*, were noteworthy due to homology with disease resistance protein family. *GG9708* exhibits homology to leucine-rich repeat receptor-like kinases (LRR-RLKs), and *GG9868* exhibits homology to zinc ion binding SSM4 proteins. Both are noteworthy *Pg*-NHR gene candidates due to their differential expression under *Pg* inoculation, their absence of homologs in the *B. vulgaris* parent, and the fact that their associated gene families have been implicated in durable resistance in other studies. While not ignoring

other lower priority candidate genes in the region, these two genes should be prioritized for future validation studies.

Although the *QPgr*-3S region is relatively large (13 cM, 3.4 Mb, 10 contigs, 99 HC genes), the identification of a few promising candidate genes demonstrates the utility of the aggregated genetic and genomic resources developed in this work to dissect the Pg-NHR expressed in *B. thunbergii*. These resources should prove valuable in the future fine mapping and functional studies necessary to dissect not only the *QPgr*-3S region but other genomic regions containing genes contributing to *Pg*-NHR in *B. thunbergii.*

Future directions

To build on the results of this study, future work should focus not only validation and dissection of the *QPgr-3S* region but also a serious exploration of the potential relevance of genes regulating *Pg*-NHR in *B. thunbergii* to durable stem rust resistance strategies in wheat. Progeny testing in an independent population derived either from the same cross (Wagon Hill × BtUCONN1) or a cross between hybrids in the original mapping population could be used for validation of the *QPgr*-3S region. In either case, the larger set of segregating lines (and critical recombinants) will be useful for saturating the linkage map, thereby delimiting the QTL to a smaller region, as per a traditional positional cloning approach.

As discussed in Chapter 2, the data suggest that the Pg-NHR in *B. thunbergii* is likely polygenic; and because no segregation for disease response has yet been observed in *B. thunbergii*, we hypothesize the existence of at least one critical NHR gene, perhaps fixed and in a recessive state, which simply cannot be mapped using the current F_1 population. According to this thinking, the heterozygous *B. thunbergii* locus mapped using the current population harbors

gene(s) necessary but not sufficient for *Pg* resistance in the hybrids. In order to detect and identify the critical gene(s) governing Pg-NHR in *B. thunbergii*, a mapping population derived from a strategic crosses between F_1 progeny that controls for the *QPgr*-3S region is required. For example, a cross between two F_1 lines carrying the resistant haplotype at *QPgr*-3S but exhibiting extreme disease phenotypes (e.g. WH15-004 and WH15-129 should allow segregation of this hypothesized critical Pg-NHR gene, thereby further elucidating the network of gene(s) governing *Pg*-NHR in *B. thunbergii.*

The research reported in this dissertation lays the foundation for a novel approach to studying NHR and demonstrates the viability of the *Berberis-Pg* pathosystem toward that end; but some words of caution are necessary with regard to the potential relevance of genes regulating Pg-NHR in *B. thunbergii* to durable stem rust resistance strategies in wheat. As described in the introduction, while wheat and barberry are both hosts of the stem rust pathogen, each are hosts (and non-hosts) to that pathogen at its various life stages (spore types). While wheat is susceptible to infection by *Pg* aeciopores and urediniospores, it is a non-host to *Pg* basidiospores. Conversely, barberry is susceptible to infection by *Pg* basidiospores but is a non-host to aeciospores and urediniospores.

It is possible, even likely, that the mechanisms governing Pg-NHR at the basidiospore stage in barberry may have no relevance at all to those governing Pg-NHR at the aeciospore or urediniospores stages in wheat. However, no study has yet been undertaken to ask the question if there is any overlap in these resistance mechanisms. It is possible that a histo-chemical study comparing the reaction of resistant barberry to infection by both urediniospores and basidiospores could shed light on whether any overlap exists in the mechanisms of recognition of and resistance to the two spore types. Complementary histo-chemical studies could be performed

comparing the reactions of a stem rust resistant wheat variety to both urediniospore and basidiospore infection. Because the various life stages in question belong to the same pathogenic organism and because *Berberis* spp. are the ancestral hosts of *Pg* prior to its host expansion to the grasses, the possibility exists that the mechanism of Pg-NHR in *B. thunbergii* may provide information about the evolution of modern day heteroecious Pg and contribute insight into possible mechanisms of durable resistance in wheat. Now that the *Berberis*-*Pg* pathosystem has been shown to be a viable research system for dissecting *Pg*-NHR, investment should be made in investigating its potential relevance to wheat improvement.