Translating genetics of oomycete resistance from *Arabidopsis thaliana* into *Brassica* production

By

Sebastian Edward Fairhead

A thesis submitted to

The University of Warwick

For the degree of

Doctor of Philosophy

School of Life Sciences

The University of Warwick

September 2016

Appendices

List of Figures

List of Tables

Acknowledgements

I would like to first thank my supervisor Eric Holub for guiding me through this project and developing my understanding of life sciences. I would also like to thank my second supervisor Graham Teakle for his valuable input. Joana Vicente has provided me advice and training in many aspects of molecular biology and plant science. Peter Walley has been hugely supportive (and patient) training me in various aspect of computational biology. Volkan Cevik developed my understanding of the pathosystem during my master's thesis, and has continued to provide advice and support. I would like to thank Jonathan Moore, Yi-Fang Wang and Laura Baxter for their help and support with bioinformatics and data analysis. I would also like to thank our laboratory assistant William Crowther for his contribution to laboratory and plant work. Thanks to all at Warwick Crop Centre and the School of Life Sciences for their general support, making this a fun and fulfilling project.

I would finally like to thank the BBSRC and MIBTP for funding this research, and providing world class training.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The preface, images used in Figure 3.1, and the initial phenotype data in Table 4.3 have been adapted from my thesis submitted for the degree of Master of Science.

The work presented was carried out by the author except in the cases outlined below:

- Section 2.2.5, unpackaging, quality control, reference genome alignments, SNP calling and the generation of the data matrix was performed by Dr. Johnathan Moore and Dr. Yi-Fang Wang.
- Section 4.2.5 was performed by Dr. Laura Baxter.
- The photographs in Figures 2.4, 2.5 and 3.2 were taken by Professor Eric Holub.
- Genotyping by Sequencing in Section 2.2.5 was performed by the Genomic Diversity Facility at Cornell University.
- RenSeq in Section 2.2.5 and PathSeq in Section 4.2.4 were performed by Oliver Furzer and Agathe Jouet, respectively, at The Sainsbury Laboratory in Norwich.

Abstract

White blister rust caused by the obligate pathogen *Albugo candida* is infectious across the *Brassicaceae*, and is an economically important disease of cultivated *Brassica* species. The advance in genotyping technologies has made possible the understanding and deployment of host resistance to plant pathogens, previously unachievable through conventional plant breeding.

In this work, the application of Genotyping by Sequencing (GBS) and Resistant Gene Enrichment Sequencing (RenSeq) has identified a single GDSL lipase as a candidate for recessive race non-specific resistance to *A. candida* in *B. oleracea*. A second locus has been identified conferring dominant race specific resistance to an *A. candida* isolate collected from Australia.

Much work has been achieved in understanding the genetic basis of resistance to *A. candida* in the model organism *Arabidopsis thaliana*, including the identification of white rust resistance (*WRR*)*4*, a single dominant resistance (*R*) gene conferring resistance to *A. candida* races 2, 4, 7 and 9 in *A. thaliana* Columbia. In this thesis research, three Columbia-virulent isolates were characterised that are capable of breaking *WRR4-Col* mediated resistance. Two of these were used to map a new broad spectrum resistance locus, designated *WRR4-OyC1,* in the vicinity of *WRR4* in the Norwegian *A. thaliana* accession Oy-0 and two additional minor effect QTLs. All three isolates were used for association genetic analysis of genome-wide 'effectorome' sequencing to identify candidate genes for *avrWRR4* in *A. candida* for both *WRR4*, and Oy-0 recognition.

From the combined results of this research, a potential strategy for durable white rust control in oilseed and vegetable *Brassica* would be stacking of at least two *R* alleles (*WRR4-Col* and *WRR-OyC1*) in a genetic background containing the recessive, resistance allele of the GDSL lipase.

List of Abbreviations

Ac	*Albugo candida*
AcAus	Australian isolate of *Albugo candida*
AcBoWells	*Albugo candida* isolate from Wellesbourne, UK
Acr	*Albugo candida* resistance
Avr	Avirulence
CAPS	Cleaved amplified polymorphic sequence
CC	Coiled-coil
cM	centi Morgan
Col	*Arabidopsis thaliana* accession Columbia
DNA	Deoxyribonucleic acid
DH	Doubled haploid
EDS	Enhanced disease susceptibility
EHM	Extrahaustorial matrix
GBS	Genotyping by Sequencing
GLIP	GDSL lipase like protein
HpA	*Hyaloperonospora arabidopsidis*
HR	Hypersensitive response
ITS	Internal transcribed spacers
LOD	Logarithm of the odds (to the base 10)
LRR	Leucine rich repeat
MAGIC	Multiparent advanced generation Inter-Cross
MLST	Multilocus sequence typing
MAS	Marker assisted selection
NB	Nucleotide binding
NDR	Non-race-specific disease resistance
PAMP	Pathogen associated molecular pattern
PCR	Polymerase chain reaction
PDLP	Plasmodesmata Located Proteins
QTL	Quantitative trait loci
R gene	Resistance gene
RAC	Resistance to *Albugo candida*

RenSeq	Resistant gene enrichment and sequencing
RIL	Recombinant inbred lines
RNA	Ribonucleic acid
RPP	Resistance to *Peronospora parasitica*
SA	Salicylic acid
SNP	Single nucleotide polymorphism
TAIR	The Arabidopsis information resource
TIR	Toll/interleukin receptor
WRR	White rust resistance
WS	*Arabidopsis thaliana* accession Wassilewskija

Chapter 1

General introduction

1.1 Preface

The Food and Agriculture Organization predict that net food production will need to increase 60% by 2050 in response to population growth, and an increasing demand for more nutritious diets in developing nations (FAO, 2016). This rise in demand runs parallel to decreasing farmland per capita and water availability, and increasing frequency and severity of damaging weather events causing volatility in annual crop yields. The changing climate is simultaneously altering local pressures from diseases and pests, whilst tighter agrochemical legislations are reducing the number of pesticides available to growers, impairing their ability to respond to potentially devastating crop infestations.

Faced with these pressures, global agriculture needs to produce more food from the same area of land whilst reducing the inputs, water and energy required for production. Failure to achieve this will result in greater levels of environmental degradation caused by food production, undermine the ecosystem services on which food production is reliant, leading to increased levels of food insecurity. Food production now requires sustainable intensification; where output is increased whilst increasing the efficiency and sustainability of production.

Agriculture is a system comprising of biotic and abiotic interactions that determine the input requirement to achieve a viable output. Components of this system can be altered at different levels, from the regional ecology of agricultural landscapes to the interactions between organisms in given cropping systems. Sustainable intensification can therefore be achieved through multiple changes within an agricultural system.

1.2 Genetics, an applied discipline from inception

The development and cultivation of crop varieties possessing durable resistance to damaging diseases is a system change that reduces the agrochemical requirement and consequently the environmental impact and cost of production, whilst maintaining crop quality and yield. The introgression of disease resistance into crop varieties first requires an understanding of what causes some plants to be susceptible to certain pathogens, while others display varying degrees of tolerance.

Gregor Mendel was a pioneer of applying a mathematical method for experimentation in biology, with the aim to "determine the law according to which (traits) appear in successive generations" (Mendel, 1865). Mendel used *Pisum sativum* (pea) plants as an experimental model organism to generate monohybrid crosses and observe patterns of inheritance of natural variation in seven traits including: flower colour, colour and form of the mature seed coat, colour and form of the mature pod, colour of the unripe pods, and overall plant height. Mendel's core discovery was that for each trait the hybrid took the characteristic of a single parent. Yet the progeny of the hybrid expressed the characteristics of both parents, suggestive that the hybrid possessed the determinants of both parental lines. Mendel termed the characteristics that were present in the hybrid as dominant, and those that were absent until the subsequent generation as recessive. This led to the observation that in the second generation the dominant and recessive traits appeared at an approximate ratio of 3:1. Heritable variation was apparent when observing the progeny of a single plant, yet the 3:1 ratio became more apparent when collating results. This led Mendel to conclude that large sample sizes accounted for errors induced by natural variation.

From the understanding that the dominant phenotype can be the result of the parental or hybrid type, Mendel suggested that the two observed phenotypes in the second generation were actually three distinct classes of dominant, recessive and hybrid, with the characteristic of the dominant type always apparent in the hybrid. In contemporary annotation, this ratio would be denoted as homozygous AA, heterozygous Aa and aA, and homozygous aa. In dihybrid

crosses with parental lines possessing two separate phenotypes, Mendel observed the progeny exhibited four phenotype classes of a ratio of 9:3:3:1, the exact product of two 1:2:1 ratios. In trihybrid crosses Mendel observed 27 phenotype classes in a ratio consistent with three 1:2:1 ratios. Backcrossing the hybrids to the parental lines again generated ratios consistent with the observed patterns of inheritance. The significance of Mendel's discovery remained unrealized until the subsequent rediscovery by the German botanist Carl Correns and Dutch botanist Hugo de Vries in the 1890's. The discoveries laid the foundation for the discipline termed genetics by William Bateson in 1906 (Bowler, 1989).

The first demonstration that genetics could be used to harness natural variation for economic benefit came through the breeding of resistance of yellow rust resistance in wheat (*Puccinia striiformis* f.sp. *tritici*) (Biffen, 1907). By crossing the partially resistant variety Rivet with the susceptible variety Red King, Biffen observed the susceptibility of Red King in the F_1 hybrids. The F_2 progeny displayed a Mendelian pattern of inheritance, segregating for susceptible and resistant at a ratio of 3:1. Although Biffen further observed "In the following (F_2) generation the relatively immune individuals bread true for this character, though not necessarily to other characters as well", implying that the introgression of resistance had come at physiological cost. Both the F_1 and the F_2 observations were suggestive of recessive resistance, or a dominant allele that conferred susceptibility. Furthermore, the F_3 progeny of the susceptible lines either conferred complete susceptibility, or were segregating for resistant. This confirmed the Mendelian prediction that two thirds of the susceptible progeny would be heterozygous at the resistance locus. In a separate study (Biffen, 1912), recessive resistance to yellow rust in the variety American Club was observed, with the F_2 progeny of crosses to susceptible lines again segregating for susceptible and resistant at a ratio of 3:1. Biffen also noted that no morphological defects were present in the resistant F_2, consequently making the lines ideal for selection of parents for the breeding of commercial varieties.

The first application of genetics to study both the plant and pathogen was performed on the *Linum usitatissimum* (flax) and *Melampsora lini* (flax rust)

pathosystem (Flor, 1942; Flor, 1947; Flor 1955). Flor generated hybrids from cross-fertilizing different races of the fungus, and evaluated the segregation of pathogenicity in F_2 generation following inoculation of flax varieties each containing a different R-gene. In many cases, he observed a 3:1 ratio of avirulent to virulent F_2 progeny, suggesting that a single *avr* allele confers an incompatible phenotype in a specific correspondence with each *R*-gene. From this, Flor proposed an interaction between paired cognate genes in the pathogen and host. This provided a theoretical basis for the gene-for-gene hypothesis (Flor, 1971), where an avirulence gene in the pathogen produces an avirulent phenotype on a host possessing the corresponding *R* gene. There were exceptions. For example, two flax varieties were resistant and six were susceptible to all F_2 fungal cultures, and one culture showed segregation ratios suggestive of two recessive loci inducing pathogenicity. However, Flor's simple gene-for-gene hypothesis has provided a general explanation for major gene resistance commonly used to develop new resistant varieties, and as described below, provided the theoretical basis for the molecular identification of host and pathogen genes encoding matching pairs of recognition (*R*) and avirulence (*avr*) proteins, respectively.

1.3 Molecular basis of induced host defence

Active defence following pathogen infection results in a hypersensitive (HR) response, the initiation of rapid cell death surrounding the site of infection to arrest pathogen development (Jones and Dangl, 2006). The recognition initiating a HR has common features to the defence strategies deployed by vertebrates. In both cases, host recognition of pathogen derived proteins results in the induction of signalling cascades that generate a defence response. In mammalian pathosystems, initial recognition is triggered by antigens, and in plants by elicitors (Keen, 1990). Elicitors translated from avirulence genes only initiate an HR response in plants possessing the corresponding *R* gene. Pathogens vary in their avirulence gene repertoire and consequently their compatible (virulent) and incompatible (avirulent) host interactions. Consequently, avirulence genes generate specific recognition by certain plant species.

The isolation of the first pathogen effectors was achieved through the generation of a genomic library of *Pseudomonas syringae* pv. *glycinea* (Staskawicz *et al.*, 1984). Out of 680 race 6 cosmid clones mobilised in individual conjugations to a race 5 strain, a single clone changed the race 5 isolate from virulent to avirulent on susceptible soybean cultivars. Similar work on *P. syringae* pv. tomato race 0 led to the discovery of *avrPto* (Ronald *et al.*, 1992). This avirulence gene interacts with the tomato *Pto R* gene, which in turn was the first *R* gene to be cloned that fits the gene-for-gene hypothesis (Martin *et al.*, 1993). Different resistance genes are often clustered with alternate functional alleles being present at some *R* gene loci. A high recombination frequency has also been observed between tightly linked *R* genes, allowing for the selection of new recognition specify (Crute and Pink, 1996, Bennetzen *et al.*, 1988, Pryor, 1987, Shepherd and Mayo, 1972).

Adapted pathogens secrete further effectors that evade or suppress pattern recognition receptor dependant responses. Effector repertoires have been documented in a diverse range of pathogens. *Pseudomonas syringae* can suppress PAMP triggered immunity through the injection of type III effector (T3E) proteins into the host cell. Immune suppression is achieved through altering organelle function, blocking RNA pathways and interfering with receptor signalling (Block and Alfano, 2011).

The first layer of plant immunity is perception by extracellular pattern recognition receptors. Typically plasma membrane-localized receptor-like kinases or receptor-like proteins, pattern recognised receptors recognise conserved microbial or pathogen associated molecular patterns (PAMPs) such as peptidoglycans, bacterial lipopolysaccharides or fungal chitin (Monaghan and Zipfel, 2012). A binding of the effector and *R* protein initiates physiological changes in the host that attempt to suppress infection. These include intracellular signalling and a transcriptional reprogramming through the activation of calcium-dependant and mitogen-dependant protein kinases, which is sufficient to inhibit colonisation in the majority of cases (Monaghan and Zipfel, 2012).

Intracellular immune receptors provide the second layer of defence by recognition of secreted pathogen effector proteins. Such resistance is often correlated with an HR. These cytoplasmic *R* genes typically encode proteins with a central nucleotide binding (NB) domain, flanked by leucine rich repeats (LRRs) on the C- terminal and an N-terminal signalling domain. Two modes of recognition of NB-LRR genes have been documented: non-self-recognition through directly interacting with effector proteins, and modified self-recognition through recognising host target modification by effector proteins (Chisholm *et al.*, 2006, Jones and Dangl, 2006).

The HR induced necrosis does not directly inhibit the pathogen, but enables the build-up of metabolites that act as secondary elicitor for the activation of defence response in neighbouring cells. Therefore, *R* genes encode two functions: the specific recognition of pathogen elicitors and HR initiation (Keen, 1990).

NB-LRR genes are highly polymorphic owing to the specialisation induced through pathogen host coevolution. Positive selection for adaptation following the loss of immunity is apparent through allelic sequence variation. Divergence within NB-LRR genes has developed two subclasses distinguished by an N-terminal Coiled-coil (CC) or Toll and human interleukin receptor (TIR) domain. This divergence has led to differences in signalling networks, with TIR-NB-LRR genes requiring a functional allele of a lipase-like enhanced disease susceptibility *(EDS)1* gene, and CC-NB-LRR genes often requiring the Non-race-specific Disease Resistance *(NDR)1* gene to confer resistance (Aarts *et al.*, 1998b, McHale *et al.*, 2006). New specificities can also evolve owing to the complexity of NB-LRR genes, where transposon insertions, gene rearrangements and duplications allow for the selection of mutations conferring resistance to different isolates. However, the limited number of NB-LLR repertories in any given plant genome indicate the original virulence factors were conserved across different pathogen classes (Goritschnig *et al.*, 2016).

1.4 The *Brassicaceae*

The angiosperm family *Brassicaceae* is currently known to include 338 genera and 3709 species (Warwick *et al.*, 2006). Many are important to both research and agriculture, the foremost being the *Brassica* crops and *Arabidopsis thaliana*. *A. thaliana* has become the experimental model system of choice in plant biology owing to its comparatively small genome and chromosome number (157 Mb, n = 5) (Johnston *et al.*, 2005), rapid generation time, ease of crossing and fecundity (Meyerowitz and Somerville, 1994). Over three decades since 1985, more than 50,000 research articles have been published covering 406 biological disciplines which contain Arabidopsis in the title, abstract or keywords (Provart *et al.*, 2016). This research has developed the field of plant biology creating new ways address food insecurity through crop improvement.

Brassicas are highly diverse in both speciation and morphology. This diversity has led to the worldwide cultivation of brassica crops that are adapted to different climates with different harvestable components that are used as culinary vegetable, oilseeds and condiments (Cheng *et al.*, 2014). Three diploid brassicas (*B. rapa,* (AA), n = 10; *B. nigra,* (BB), n = 8; and *B. oleracea,* (CC), n = 9) and three allotetraploids originating from each pair of diploids (*B. juncea,* (AABB), n = 18; *B. napus,* (AACC), n = 19; and *B. carinata,* (BBCC), n = 10) are the most widely cultivated. The relationship between the species was first observed by U (1935), and is widely referred to as 'The Triangle of U' (Figure 1.1).

The diversity of cultivated subspecies within each of the *Brassica* denoted by 'U's' triangle is again extensive. *B. rapa* includes turnip, cok choy and field mustard. *B. nigra* is primarily cultivated for seeds used as a spice. *B. oleracea* cultivars include cabbage, broccoli, kale, cauliflower and Brussels sprouts. *B. juncea* is a mustard variety, with cultivars including English, Indian and Chinese mustard. *B. carinata* is predominantly cultivated as an oilseed crop, although comparatively high seed glucosinolate and erucic acid levels make *B. napas* the primary source of oilseed production. Turnip cultivars of *B. napus* are also grown, and are predominantly used for animal fodder (Schmidt and Bancroft, 2011).

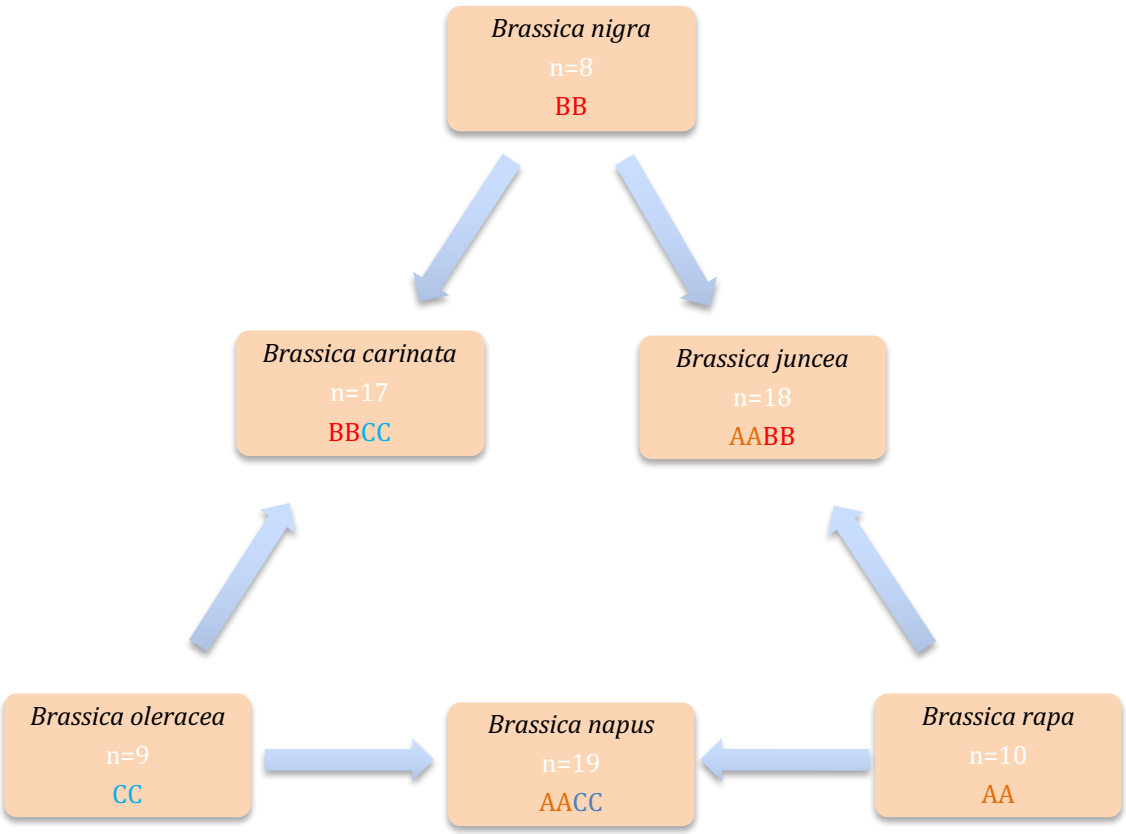


Figure 1.1 U's triangle (U, 1935) representing the relationship between the three diploid Brassicas *Brassica rapa, Brassica nigra* and *Brassica oleracea,* and the three allotetraploids *Brassica carinata, Brassica juncea* and *Brassica napus*. The haploid genomes of each diploid species are labelled A (*B. rapa*), B (*B. nigra*) and C (*B. oleracea*).

Polyploidy has been identified a driving force for evolution and speciation, through the creation large levels of gene redundancy, enabling evolutionary adaptation through subfunctionalization or neofunctionalization (Cusack and Wolfe, 2007, Blanc and Wolfe, 2004). The higher chromosome number and increased genome size are indicative of time since the polyploidy. Neopolyploids such as *B. napus* will evolve to mesopolyploids and ultimately paleopolyploids through diploidization and genetic rearrangement (Wolfe, 2001, Kagale *et al.*, 2014). Parental subgenomes are discernible in both neopolyploids and mesopolyploids, yet are more disguised in paleopolyploids owing to integration through genome restructuring over time (Kagale *et al.*, 2014). The *A. thaliana* genome reveals evidence of three distinct polyploidy events, α, β and γ (Bowers *et al.*, 2003) that are shared with the crucifer taxa (Haudry *et al.*, 2013). The genus *Brassica* shows evidence of a further whole genome triplication event

(Wang *et al.*, 2011, Beilstein *et al.*, 2010, Haudry *et al.*, 2013, Kagale *et al.*, 2014, Liu *et al.*, 2014, Parkin *et al.*, 2014) that has driven the diversification now utilised by agriculture (Cheng *et al.*, 2014).

The *Arabidopsis* and *Brassica* genomes have been shown to share a minimum of 21 conserved syntenic blocks (Lysak *et al.*, 2007). The phylogeny between *A. thaliana*, the *Brassica* A, B and C genomes and the three allotetraploids denoted by U's triangle enable comparative analysis between the model organism and its important crop relatives. Such analysis enables an understanding of how speciation caused the loss, retention and adaptation of *R* genes that has driven pathogen adaptation (Peele *et al.*, 2014, Yu *et al.*, 2014).

The diversity across the *Brassica* species is reflected by the diversity in pathogenic organisms that have adapted to exist on *Brassica* hosts, causing economically significant crop damage and increasing the agrochemical requirement for production. In many cases, much progress has been made in understanding the genetic basis for resistance in wild relatives of crop species and/or crop varieties, enabling the introgression of resistant traits through conventional breeding aided by marker assisted selection or transgenic approaches. The airborne fungal pathogen *Leptosphaeria maculans* causes the disease Phoma stem canker, inducing serious losses in *B. napus* in North America, and Australia, with UK losses in excess of £48 million per annum (Fitt *et al.*, 2006). Major gene resistance has now been discovered in *B. napus* (Delourme *et al.*, 2006). Light leaf spot is caused by the airborne fungus *Pyrenopeziza brassicae* and is one of the most significant diseases of *B. napus* in continental Europe and the UK, and also causes severe losses in *B. oleracea* (Majer *et al.*, 1998). QTL's contributing to field resistance have now been identified in *B. napus* (Pilet *et al.*, 1998). Downy mildew, caused by the terrestrial oomycete *Peronospora parasitica* is both soil and airborne, has a worldwide distribution and is one of the most damaging foliar diseases of *B. oleracea.* Resistant sources have been identified in *B. oleracea* diversity collections, and generated into double haploid lines to determine the genetic basis for resistance (Vicente *et al.*, 2012). Vascular wilt of *B. napus* caused by soil borne fungus *Verticillium longisporum* was first confirmed in the UK in 2007 and can causes crop losses ranging from 12-24%. Significant losses in *B. oleracea* have also been

documented (Klosterman *et al.*, 2009). Differential responses to infection have now been discovered in *B. napus,* providing a means for determining the genetic basis of resistance (Eynck *et al.*, 2007). Turnip yellows virus is one of the most important viruses effecting Brassicas in the United Kingdom and is transmitted through the mouth parts of *Myzuz persicae* (peach potato aphid). Oil content and yield of *B. napus* and vegetative production of *B. oleracea* are seriously affected by infection. Introgression of resistance into commercial *B. napus* has now been achieved (AHDB, 2015). Black spot caused by different species of the soil an airborne fungus *Alternaria* is infectious across the *Brassicaceae*, with crop losses as high 86% in *B. oleracea* being documented (Nowicki *et al.*, 2012). Major effect resistance has not yet been identified in cultivated *Brassica* species, although differing degrees of tolerance have been reported (Nowicki *et al.*, 2012). Sclerotinia stem rot caused by the soil borne fungus *Sclerotinia sclerotiorum* is again damaging across the *Brassicaceae,* causing yield losses in oilseed production ranging from 10-80% (Mei *et al.*, 2011). Loci associated with partial resistance (Zhao and Meng, 2003) and major resistance (Wu *et al.*, 2013) have now been discovered in *B. napus.* Clubroot caused by *Plasmodiophora brassicae* has a world-wide distribution and is pathogenic across the *Brassicaceae*. Resistant *B. rapa* accessions have now been developed using European turnip cultivars as a source of resistance (Hirai, 2006). *Xanthomonas campestris* is the causal agent black rot of crucifers and *Brassica,* and is the most important disease of *B. oleracea* worldwide. Resistance to different physiological races of *X. campestris* has now been discovers in both *B. oleracea* and *B. campestris* (Vicente *et al.*, 2002).

1.5 *Albugo candida*

The eukaryotic oomycete order *Albuginales* consists entirely of obligate biotrophic pathogens of plants. The largest genus is *Albugo,* the causal agent of white blister rust (white rust or staghead), with a worldwide distribution and comprised of over 50 species that collectively parasitize over 400 domesticated and wild host species (Biga, 1955, Choi and Priest, 1995, Walker and Priest, 2007). Three species cause economically significant crop damage including *A. candida* in oilseed and vegetable brassicas (Choi *et al.*, 2009), *A. ipomoeae-*

panduratae in sweet potato (Sato *et al.*, 2009), and *A. tragopogonis* (syn. *Pustula helianthicola*) in sunflower (Lava *et al.*, 2013, Thines *et al.*, 2006).

White blister rust causes crop yield and quality reductions in different ways. The development of rust pustules on foliage restricts the photosynthetic capacity of the host, whilst simultaneously downgrading the aesthetic and nutritional quality of the harvestable component. Pustule growth on the stem causes structural weakness, impairing transpiration and nutrient transport whilst increasing the likelihood of lodging. Malformations such as abnormal stem twisting and impaired floral development cause the majority of seed yield reductions through pod abortion and frequently host death (Figure 1.2). This phase of disease is especially destructive in seed crops such as Indian mustard (*B. juncea*). Symptoms are influenced planting time, nutrient availability, environmental conditions and genotype-pathotype interaction (Saharan and Verma, 1992).

A. candida can infect over 200 host species in 63 genera within the *Brassicaceae*. Pathogen development post infection is favoured by canopy temperatures ranging from 12-24°C with relative humidity in excess of 70% (Liu and Rimmer, 1990, Chattopadhyay *et al.*, 2011). Regions where such conditions commonly present during crop development include India and Pakistan where large areas of arable land are dedicated to the cultivation of susceptible *B. juncea* and oilseed crucifers. Yield reductions following white rust have been calculated to be as high as 89.8%, but more commonly range from 40-60% (Saharan and Verma, 1992). Such losses impact regional food security and are economically devastating to smallholder farmers. In United Kingdom high rates of infection can be observed in commercially grown *B. oleracea, B. juncea* and wild *Capsella bursa-pastoris*.

A

B

Figure 1. White blister rust caused by *Albugo candida* growing on **A,** leaf tissue of *Brassica oleracea* (cabbage) in Warwickshire, UK; **B,** Inflorescence tissue of *B. juncea* (English mustard) growing in Peterborough, UK.

A. candida is a diploid organism that reproduces both sexually and asexually. Generally, symptoms become apparent above ground through the development of white pustules on the leaf surface and aerially exposed regions through the enzymatic digestion of epidermal cell walls. The pustules contain asexual sporangiospores, situated for dispersal by air currents or rain droplets following the rupturing of the sorus. Each sporangiospore contains 4-6 biflagellate zoospores, which are released upon re-hydration and swim into the stomatal opening of the subsequent host. Following infection, the zoospores encyst through the loss of the flagella and the production of cell walls; allowing the development of a germ tube which enters the sub-stomatal chamber before penetrating the palisade mesophyll cells. Intercellular hyphae then migrate through the leaf tissue where they encounter veins, allowing access to the hypocotyl causing the infection to become systemic. Epidermal emergence and subsequent sporulation conclude the asexual life cycle (Holub *et al.*, 1995). *A. candida* also exists in the soil or as a seedbourne contaminant as dormant oospores developed through induced hypertrophy associated with the sexual phase of reproduction. Male (antheridia) and female (oogonia) sex organs develop on the hyphal tips deep within the inflorescence tissue (Holub *et al.*, 1995).

A. candida is currently subdivided into 17 physiological races determined by host specificity across the Brassicaceae. The first six races were described by Pound and Williams (1963) as race 1 of *Raphanus stivas,* race 2 of *B. juncea,* race 3 of *Armoracia rusticana,* race 4 of *Capsella bursa-pastoris,* race 5 of *Sisymbrium officinale* and race 6 *Rorippa islandica.* Race 7 was later reported on *B. rapa* (syn. *B. campestris; (Verma et al., 1975)*), race 8 on *B. nigra* (Delwiche, 1976), race 9 on *B. oleracea* and race 10 *Sinapis arvensis* (Hill *et al.*, 1988), race 11 on *B. carinata* (Williams, 1985) and races 12 – 17 providing a host differential between Indian varieties of *B. rapa* and *B. juncea* (Verma *et al.*, 1999, Gupta and Saharan, 2002). In addition, sub-characterisation of race 2 and 7 identifies virulence (V) or avirulence (A) on cultivars of *B. juncea* and *B. rapa* (Petrie, 1994).

It remains unclear how the races of *A. candida* evolved. Evolutionary and population genetic theory suggest that the natural selection and trade-offs of specialisation can lead to adaptive radiation and speciation of a pathogen (Abbott *et al.*, 2013, Stukenbrock *et al.*, 2012). However, *A. candida* is a distinct species that maintains multiple physiologically specialised races (described above), which are each adapted for a different range of host species. Yet there are examples of races able to cause disease on *Brassicaceae* other than their corresponding host (Pound and Williams, 1963, Downey and Rimmer, 1993, Rimmer *et al.*, 2000).

Albugo spp. are highly effective in suppressing host defence with the consequence of enhancing susceptibility to secondary infection by otherwise avirulent pathogens (Cooper *et al.*, 2008). Thus, it is possible that suppression of innate immunity facilitates the coexistence of multiple physiological races in the same host tissue, enabling sexual-crossing that enables the emergence of distinct species (Hedrick, 2013). Comparative analysis of whole genome sequences of five *A. candida* confirmed molecular divergence of three distinct races (McMullan *et al.*, 2015). However, there was a mosaic structure of sequence variation indicating genetic recombination between the different subgroups had indeed occurred.

However, this mechanism promotes intraspecific competition for resources from different pathogens such as *Hyaloperonospora parasitica* (Cooper *et al.*, 2008). In addition, the gene-for-gene hypothesis predicts that pathotypes will commonly occur within physiological races that have the ability to overcome or break resistance conferred by a specific R-gene. Sexual reproduction between races would potentially create an evolutionary disadvantage, as hybrids will inherit effector repertoires from both parental races, enabling recognition by host immunity that would otherwise go undetected. Yet under the extreme selection pressure presented in monoculture cropping systems, functional recombinants may ultimately be generated, with asexual reproduction allowing rapid colonisation once a functional hybrid has been selected.

1.6 White rust resistance in *Brassicaceae* hosts

Much has been achieved in understanding the genetics of white rust resistance across the *Brassicaceae*. In *A. thaliana*, resistance to race 4 isolate AcEm1 with phenotype of small necrotic flecks on the upper leaf surface was mapped on chromosome 1 in Ksk-1 ((Crute *et al.*, 1993, Holub *et al.*, 1995). Defined as resistance to *A. candida (RAc)1*, the locus was further characterised as a TIR-NB-LRR (Borhan *et al.*, 2004) requiring functional expression of the lipase like *(EDS)1* (Aarts *et al.*, 1998b). Recessive resistance *RAc2* has been mapped to a 6 cM interval at the bottom arm of chromosome 3 in Ksk-2 and *RAc3* has been identified as linked to the *RPP8/HRT* on chromosome 5 in Ksk-1 (Borhan *et al.*, 2001). The white rust resistance gene, *(WRR)4* has been characterised in *A. thaliana* Col-0. Located on the bottom arm chromosome 1, WRR4 is an *EDS1* dependent TIR-NB-LRR, and confers major effect resistance to *A. candida* races 2, 4, 7 and 9 (Borhan *et al.*, 2008). Col-0 is now known to possess three other WRR alleles, the functional gene pair *WRR5* and *6*, and *WRR7*. All are situated on chromosome 5 (Holub and Cevik, unpublished).

Pound and Williams (1963) observed a 3:1 segregation ratios of resistant to susceptible from separate self-pollinated heterogeneous accessions of *B. oleracea* resistant to a race 1 isolate, suggestive of resistance being conferred by a single dominant gene. In *B. napus*, inheritance studies on independent crosses between the Canadian race 7 resistant cultivar Regent with two Chinese susceptible cultivars predicted three independent dominant genes at three loci, Ac7-1, Ac7-2 and Ac7-3 (Fan *et al.*, 1983, Liu *et al.*, 1996). Inheritance studies in *B. napus* have also suggested a single recessive gene *(wpr)* conferring a partially resistant phenotype with pinhead size pustules developing on the leaf surface at the site of infection in *B. napus,* (Bansal *et al.*, 2005).

In *B. juncea*, *A. candida* resistance *(Acr)*, has been mapped to a 6.3 cM region chromosome 7 in resistant line J90-2733 (Cheung *et al.*, 1998). J90-2733 was also used to define *A. candida Ac2(t)* locus with RADP markers (Prabhu *et al.*, 1998). This was subsequently narrowed with two PCR-based cleaved amplified polymorphic (CAPS) markers at distances of 3.8 cM and 67 cM from the R gene (Varshney *et al.*, 2004). Here, the use of PCR based genotyping greatly improved that ability for breeders to apply marker assisted selection (MAS). Two further

independent loci have been mapped in *B. juncea* in partially resistant Eastern European line Heera on A4 (AcB1-A4.1), and fully resistant Eastern European line Donskaja-IV on A5 (AcB1-A5) (Panjabi-Massand *et al.*, 2010). Evolving in the absence of pathotypes found in South Asia, these sources of resistance may prove durable in Indian oilseed production.

In *B. rapa,* the development of 144 restriction fragment length markers on a susceptible 'R500' x race 2 and 7 resistant 'Per' recombinant identified a single major effect locus *(Aca1)* on A4 in the same location as resistance previously characterised against race 2. This suggested either a single gene conferring resistance to both races or two tightly linked genes (Kole *et al.*, 2002). In addition, a second a minor effect QTL was discovered on A02, syntenic to *Aca2* conferring resistance to race 9 in *B. oleracea* which we define in this work (Chapter 2).

1.7 Genetics of avirulence in *Albugo candida*

The inheritance of virulence in the pathogen has also been studied. By generating hybrids of race 2 and race 7 through co-infection on a common susceptible host and examining virulence of the F_2 on cultivars of *B. rapa* cv. Torch, a 3:1 segregation ratio was observed, suggestive of a single dominant avirulent gene *(AvrAc1)* (Adhikari *et al.*, 2003). This provided further evidence of a gene for gene relationship in the *Albugo-Brassica* pathosystem. In addition, systemic resistance to *A. candida* has shown to be induced by pre inoculation with incompatible isolates prior to inoculation with otherwise virulent isolates, or inoculation with both virulent and a virulent isolate in *B. juncea* (Singh et al, 1999) (Singh *et al.*, 1999).

1.8 Aims and objectives

Broad spectrum resistance to *A. candida* race 9 has previously been identified in the *B. oleracea* double haploid accession EBH527, whilst the accession A12DH was identified as being broadly susceptible to an extensive collection of *A. candida* isolates derived from UK *B. oleracea* production (DEFRA, 2003). However, interestingly, this accession is resistant to an *A. candida* isolate obtained from Australia (AcAus) (Holub, unpublished). Thus, the primary aim of this study was to generate an F5 recombinant inbred population of A12DH x EBH527, and apply next generation sequencing to develop markers and define the genetic basis for resistance in EBH527. Secondly, to define a locus conferring resistance to AcAus in A12DH using the same population.

Borhan et al (2008) proposed that *A. candida* contains a highly conserved effector present in at least four races (2, 4, 7 and 9) to explain the apparent broad spectrum resistance conferred by the *WRR4-Col* allele from *A. thaliana*. If so, then Col-virulent pathotypes may arise in natural populations due to mutations in this predicted Avr gene. Two *WRR4-Col* virulent isolates, AcExeter and AcCarlisle, have been identified ((Fairhead, 2012) Master's thesis). The third aim of this study was to use AcExeter and AcCarlisle to identify and map a new and potentially broader spectrum source of white rust resistance in *A. thaliana* that could be used as a transgene in commercial *Brassica* production.

Although *A. candida* has been useful to investigate disease resistance in *A. thaliana* under controlled environment conditions, it is important to know whether *A. thaliana* is a potential reservoir for this pathogen under field conditions. Thus, the fourth aim was to determine whether *A. candida* could be readily detected as the cause of white rust in natural populations of *A. thaliana,* particularly in floral stem and leaf tissue of plants growing in close proximity to *Capsella bursa-pastoris*, which is the most prolific source of inoculum. The identification AcExeter provided the first evidence that *A. candida* can overcome WRR4-mediated resistance, and was the first evidence of a previously race non-specific resistance to *A. candida* breaking down. Thus, the fifth aim was to determine whether additional Col-virulent isolates could be collected from Arabidopsis under field conditions. Association genetics can provide a means for identifying virulence determinants in microbial pathogens (Bart *et al.*, 2012).

Thus, the sixth aim was to determine whether a collection of Col-0-virulent isolates could be used to search genome-wide effector sequences and identify candidates for *avrWRR4-Col*. As described in this study, the Norwegian accession of *A. thaliana* Oy-0 carries resistance to AcExeter and AcCarlisle. Thus, the final aim was to use the effector database to search for candidate effectors that trigger *WRR4* and *WRR-OyC1* resistance.

Chapter 2

Use of a recombinant inbred *Brassica oleracea* population and modern genotyping technology to determine the genetic basis of broad-spectrum white rust resistance

2.1 INTRODUCTION

Albugo candida (causal agent of white rust) is a major threat to production of vegetable and oilseed brassicas in the UK and across the world (Saharan *et al,* 2014). The range of approved crop protection products for use on *B. oleracea* (cabbage, broccoli, cauliflower, kale, etc) is heavily restricted, as it is a vegetable crop where the harvested biomass is directly consumed (Alford, 2008). Fungicide applications of metalaxyl and Chlorothalonil are approved for use on *B. oleracea* in the UK, yet only two sprays up to the maximum permitted dosage are recommended (Alford, 2008). With few available control options the likelihood of selecting for chemical resistant pathogens is greater. Disease resistant varieties provide an essential alternative control of white rust, increasing the potential durability of integrated disease management whilst reducing the environmental impact of production and the consumer's exposure to agrochemical residues. White rust resistance is an important target for commercial breeding of brassicas.

Single dominant *R* genes are available in varieties of many crop species, and have typically been deployed individually in cropping systems. Consequently, selection pressure on monocultures has enabled the pathogen in many cases to rapidly overcome the resistance. Well documented examples include late blight of potato caused by *Phytophthora infestans*, stem rust of wheat caused by *Puccinia graminis*, blackleg of oilseed rape caused by *Leptosphaeria maculans* and downy mildew of lettuce caused by *Bremia lactucae* (Ballini *et al.*, 2013, Zhang *et al.*, 2009, Sivasithamparam *et al.*, 2005, Crute and Norwood, 1981). Strategies such as pyramiding of *R* genes or the use of multi-parental populations (*e.g.,* variety mixtures) have been considered for longer-lasting disease control, and have worked in some crops (Joshi and Nayak, 2010). However, such approaches have been impractical in most crops due the time and cost required to introgress multiple target loci into a new variety. Whilst mixing cultivars does have potential application in cereals and other combinable crops, the morphological uniformity desired by retailers and the consumer makes this impractical for vegetable production.

The advent of new crop improvement technologies is greatly improving the feasibility of introgressing multiple resistant genes into a cultivar. For

example, progress has been achieved in model organisms such as *Arabidopsis thaliana* where research has revealed the molecular basis of *R* genes conferring resistance to a wide spectrum of pathogens. These genes are potentially useful in transgenic crops, as demonstrated with a white rust resistance gene (*WRR4*) in oilseed brassicas (Borhan *et al.*, 2010), or more generally as precedents for the development of molecular markers in conventional breeding (Speulman *et al.*, 1998, Aarts *et al.*, 1998a, van der Linden *et al.*, 2004). Combining multiple` resistance specificities is theoretically achievable using conventional marker-assisted breeding to pyramid resistance alleles from several loci and/or 'stacking' of alleles in a single DNA construct for GM application. Recent advances in affordable DNA sequencing have enabled rapid development of marker technologies such as Genotyping-by-Sequencing (GBS) (Elshire *et al.*, 2011) and Resistant Gene Enrichment Sequencing (RenSeq) (Jupe *et al.*, 2013).

With these advances in technology, *B. oleracea-A. candida* is emerging as a suitable crop pathosystem for improvement through the sustainable deployment of *R* genes. Nine sources of white rust and downy mildew resistance have been identified in a European diversity collection of 400 *B. oleracea* accessions (Leckie *et al.*, 1994, Vicente *et al.*, 2012, DEFRA, 2003), with the aim of establishing a breeding resource for combined resistance to both diseases. The *B. oleracea* ssp. *alboglabra* accession A12DH was identified as being broadly susceptible to an extensive collection of *A. candida* and *H. brassicae* isolates derived from UK *B. oleracea* production (DEFRA, 2003). However, interestingly, this accession is resistant to an *A. candida* isolate obtained from Australia (AcAus) (Holub, unpublished). Doubled haploid (DH) lines were produced for each source of resistance to facilitate genetic studies and more reproducible pathotyping of pathogen isolates (DEFRA, 2003). Specifically for white rust resistance, the DH lines were assessed for reaction to 18 *A. candida* isolates sampled from *B. oleracea* across major vegetable growing regions across the UK. Three lines (EBH527, 516 and 535) were resistant to all isolates, whilst two lines (EBH508 and 553) exhibited differential reactions.

The broad spectrum resistant line EBH527 was chosen for further investigation in the current study, to determine the genetic basis for resistance through the use of a recombinant inbred mapping population (A12DH x

EBH527) and the latest marker technologies (GBS and RenSeq). A candidate *R*-gene was identified, and subsequently enabled use of the C genome diversity set (Walley *et al.*, 2012) to investigate allelic variation relative to the resistance phenotype. Work was also initiated to map alternative resistance specificity in both EBH527 and A12DH to the *A. candida* isolate from Australia (AcAus).

2.2 MATERIALS AND METHODS

2.2.1 Maintenance of *Albugo candida* isolates

A standard isolate of *A. candida* race 9 (AcBoWells) was obtained by myself and Dr. Joana Vicente in September 2014 from a cabbage field experiment at the University of Warwick Crop Centre. A genetically refined culture was generated by inoculating *B. oleracea* var. Maris Kestrel with zoosporangia harvested from a single small pustule site with a pipette tip. These were suspended in sterile water and then transferred to the underside of nine-day-old cotyledons. The resulting isolate was then bulked and maintained on cotyledons of *B. oleracea* var. Marris Kestrel. Similarly, a second isolate obtained by Professor Eric Holub from researchers in Australia (AcAus) was bulked and maintained on *B. oleracea* var. Senna.

For preparation of plants for inoculation, seeds were grown in a peat-based Levington M2 compost in P180 plug trays (2.5 x 2.5 cm cells) cut to 10 x 6 grids and placed in propagator trays. Holes of a depth of approximately 0.5 cm were made in the compost. A single seed was placed in each cell and covered with vermiculite. The trays were covered with aluminium foil and placed in a fridge at 4° C for 48h to break dormancy and promote even germination. Lids were placed on the propagators, which were subsequently placed in a Conviron plant growth cabinet at 20±2°C with a 10h photoperiod.

For quantification of inoculum, a haemocytometer was used and the inoculum concentration was adjusted to approximately concentration of $4x\ 10^4$ zoosporangia ml^{-1}. Small droplets (ca. 10µl) were applied to the upper and lower surfaces of cotyledon leaves attached to the plant using a pipette. Propagators were sealed and placed back in the growth cabinet at 20±2°C for an initial 12h

period in darkness followed by a 10h photoperiod. Pustules would emerge 7-10 days later. The isolates were subcultured every two weeks for use on experimental lines

2.2.2 Recombinant inbred mapping population and diversity collection of *Brassica oleracea*

The initial A12DH - EBH527 cross was performed by Professor Eric Holub. F_1, F_2 and F_4 seeds were provided as starting material of this project. To establish a mapping population, F_5 recombinant inbreds were generated in a polytunnel from 70 F_4 lines. One plant from each parental line was grown in individual 3L pots containing M2 compost. Flowering plants from each generation were covered with air permeable bags and agitated to assist self-pollination. Each plant was desiccated at maturity by severing the stem with secateurs. The harvested plant material was placed in a drying room for ten days, prior to seed cleaning. A working stock of clean seed samples was sealed in 1.5 ml Eppendorf tubes and stored at -20°C, with the remainder stored in 5 ml Falcon tubes.

In addition, 126 lines from the wild species and diversity fixed foundation set and 120 lines from the *B. oleracea* diversity fixed foundation set (Appendix 1) were obtained from the Genetic Resources Unit at Warwick Crop Centre (Walley *et al.*, 2012). Transcriptome sequence from the wild species diversity fixed foundation set was obtained through the University of Warwick (Barker et al, unpublished).

2.2.3 Inoculation of experiments

Experimental lines of *B. oleracea* were tested using the same method as used for the maintenance of *A. candida* isolates. To test the F_5 A12DH x EBH527 mapping population, three replicates of each inbred, ten susceptible (A12) and ten resistant (EBH527) controls were randomly established in five propagators. To test the diversity set, 3-5 replicates of each line were established in individual columns across 30 propagators. In each test, nine-day-old cotyledons were inoculated with a single 10 μl drop of *A. candida* inoculum to the upper surface of the leaf at an approximate concentration of $4x\ 10^4$ zoosporangia ml^{-1}. Propagators containing the experimental lines were sealed and placed back in

the growth cabinet at 20°C for an initial 12 h period of shading, followed by a 10h photoperiod. Symptoms were assessed ten days' post inoculation using a quantitative phenotype type scale. For the mapping population, a one-way analysis of variance was performed across the five propagators.

2.2.4 DNA extraction

True leaf tissue from each of the F_4 parental lines, EBH527 and A12DH was harvested using a sterile scalpel and tweezers and placed in 1.5 ml Eppendorf tubes. The samples were freeze dried for seven days, sealed in tubes and stored at -20°C.

A sterile pestle was used to crush lyophilized tissue within the tubes. Approximately 50 mg of lyophilized tissue was transferred into 2 ml Eppendorf tubes. Each sample was disrupted by adding two tungsten beads and placing the sealed Eppendorf tubes in the TissueRuptor® for a total time of 1.5 m at 60 Hz. After 45 s of disruption, the TissueRuptor® containing plates were inverted, swapping samples situated closer to the machine with those situated on the edge of the containing plate which consequently reverberated at a greater frequency.

DNA was extracted using a DNeasy® Plant Mini Kit. Quantities of buffer were amended in order to achieve a higher DNA yield. A 500 µl aliquot of lysis buffer AP1 and 5 µl of RNase A were added to each sample. The samples were incubated for ten minutes at 65°C, and agitated four times during incubation. A 162.5 µl volume of neutralisation buffer P3 was added to each sample, which was then incubated on ice for ten minutes. The lysate was centrifuged for 5 m at 14,000 rpm. The flow through was transferred into a new 2 ml Eppendorf tube and 1 ml of Buffer AW1 was added. The samples were mixed by pipetting and centrifuged through DNeasy mini spin columns at 8000rpm in three separate aliquots of approximately 556µl. Three washes were performed using 500 µl of wash buffer AW2. The first two were centrifuged through the DNeasy mini spin columns for 1 minute at 8000 rpm. The final wash was centrifuged at 14000 rpm for 2 minutes. The samples were eluted into two separate 1.5 ml Eppendorf tubes, using 100 µl of buffer AE for each. A Qubit flurometric quantifier was used to measure the yield achieved thought the extractions. The first elution achieved

concentrations ranging from 300 – 140 ng/µl, and the second elution from 60 - 16 ng/µl. The first elution was used was used for next generation sequencing applications and the second elution was used for conventional PCR techniques.

2.2.5 Molecular genotyping

For Genotyping-by-Sequencing (GBS), DNA samples from the first elution of the 70 F_5 RILs, EBH527 and A12DH were reduced to 10 µg of DNA at ≥50 ng/µl. The samples were sent to the Genomic Diversity Facility at Cornell University. The library was prepared through restriction digests at ApeKI recognition sties as described in Elshire et al. (2011). Single-end reads were generated by 48-plex sequencing through 101 cycles on an Illumina HiSeq 2500.

For Resistance Gene Enrichment Sequencing (RenSeq), DNA samples from the first elution of the EBH527 and A12DH were reduced to 10 µg of DNA at ≥50 ng/µl. The samples were sent to The Sainsbury Laboratory in Norwich for Resistance Gene Enrichment and Sequencing (RenSeq). The library was prepared as described in Jupe et al. (2013). Paired-end reads were generated by 96-plex sequencing through 150 cycles on an Illumina HiSeq 2500.

Bioinformatics support for GBS and RenSeq data analysis was provided by Dr. Jonathan Moor and Dr. Yi-Fang Wang. Quality control was performed through FastQC (Andrews, 2010). Adaptors and low quality reads were removed using TrimGalore (Krueger, 2015). The trimmed reads from both datasets were aligned to the *B. oleracea* reference genome (accession TO1000; (Parkin *et al.*, 2014)) by the Bowtie2 (version 2.2.3; (Langmead and Salzberg, 2012)) with default settings. The aligned .sam files were converted into .bam files using SAMtools (version 0.1.19; (Li *et al.*, 2009)).

For the reference-based SNP calling, the SAMtools mpileup and bcftools were applied with phred score>=20 (Li, 2011). The results were converted from vcf format to bcf format. The Discosnp++ (Uricaru *et al.*, 2015) was applied with default settings for the reference free SNP calling between A12DH and EBH527. The reference-based SNP markers for the GBS parental lines were converted into a genotype score matrix for the progeny by a custom PERL script written by Wang (2014).

2.2.6 QTL analyses

All QTL analyses were performed using RQTL (Broman *et al.*, 2003). Three statistical approaches were applied to interval mapping. Standard interval mapping applied a maximum likelihood estimation under a mixture model (Lander and Botstein, 1989). The Haley-Knott regression used approximations of the mixture model (Haley and Knott, 1992). And the multiple imputation used the mixture model with multiple imputations as appose to maximum likelihood estimation (Sen and Churchill, 2001).

2.2.7 Reference based marker development and use in genotyping

The *B. oleracea* reference genome TO1000 (Parkin *et al.*, 2014) was used for designing markers in candidate genes (Table 2.2). Primer 3 through Geneious version 9.1.4 (Kearse *et al.*, 2012) was used for all primer design.

All PCR reactions for genotyping were conducted using Phusion® high-fidelity DNA polymerase. Reaction volumes of 25 µl were created containing 5 µl of 5 x GC buffer, 0.5 µl of 10 mM dNTPs, 1.25 µl of 10 µm forward and reverse primers, 1 µl of template, 15.75 µl nuclease-free water and 0.25 µl of DNA polymerase added last. All reactions were prepared on ice. All PCRs were performed using a standard touchdown program (Table 2.1) with extension times adapted to 30 seconds per kilo base of product and annealing temperature adapted to each primer pair (Table 2.2). A 5 µl sample of PCR product was added to 2 µl of loading dye for assessment using electrophoresis.

Table 2.1. Standard touch down polymerase chain reaction (PCR) program used for the amplification of regions of genomic DNA within a major effect QTL conferring resistance to *Albugo candida* in *Brassica oleracea.* D: denaturation, A: annealing, E: extension.

	Start	**Touchdown (8 Cycles)**			**(28 Cycles)**			**Finish**
	D	D	A	E	D	A	E	E
Temperature (°C)	98	98	65 – 57	72	98	56	72	72
Time (minutes)	3.00	0.30	0.30	1.30	0.30	0.30	1.30	7.00

Genomic DNA and PCR products were assessed through electrophoresis, using agarose gels ranging from 0.5 – 4% volume to weight agarose/ Tris Borate buffer (TBE) according to fragment size. Separation was through applying 80-160V over 1-14 h depending on fragment size and gel concentration. Invitrogen 1 kb or 1 kb plus ladders were used for all samples. Centrifuge driven PCR purification was performed using the Qiagen PCR purification kit according to protocol.

A 5 µl aliquot of PCR product and 5 µl of primer at a concentration of 5 mM were pipetted into 1.5 ml Eppendorf tubes and sequenced through the LightRun service provided by GACT Biotech. For the sequencing of PCR products greater than 500 bp, internal primers were designed from the TO1000 reference genome and used to sequence the samples as described (Appendix 2).

Sequence analysis was performed *in silico*. Geneious version 9.1.4 (Kearse et al, 2012) was used for trimming, quality control, reference based alignment and polymorphism detection for all sequenced PCR products.

Table 2.2. Markers designed from *Brassica oleracea* var. TO1000 reference genome (Parkin *et a*l., 2014) within a major effect QTL conferring resistance to *Albugo candida* isolate AcBoWells. QTL identified in an F_5 recombinant inbred mapping population derived from a cross of parental lines A12DH (susceptible) and EBH527 (resistant).

Marker	Type	Gene ID	Primer sequence 5' To 3'	Anneling Temp (°C)	Product length (bp)	Gene type
Bo-01	SNP	Bo2g016470	F: CTTTGAATCTGACGGATGAAGGAAG	63.07	2990	TIR-NBS-LRR class disease
			R: GATTCAATTTCTCCGTCAAGGCTAG	68.93		resistance protein
Bo-02	Whole	Bo2g016480	F: TCAGAACAACTTGAGTTATTTCATTCTCAT	61.50	1872	GDSL esterase/lipase
	gene		R: ATATTACCAAAGTATTTGTGGTCCAAGAAA	61.91		
Bo-03	Whole	Bo2g016490	F: TTTTGAGGGTTTGTAGTCAGGAGAA	61.10	1384	Ethylene-responsive
	gene		R: GTCTCTTTGTTACACGAAAACGACT	58.82		transcription factor
Bo-04	Whole	Bo2g016500	F: TTGGACTAGCTATTGTAAATCTTTCTTGAG	60.58	3731	Protein of Unknown
	gene		R: AAACTTTCAATACACGTTGGAAGTTTATTC	61.76		Function (DUF239)
Bo-05	Whole	Bo2g016510	F: GAAACAGTTGGGAGTGTAAAAGGAG	60.22	1325	G-type lectin S-receptor-like
	gene		R: TGATTCTGTGAAGTGATGTGTTGTG	60.82		protein kinase
Bo-06	SNP	Bo2g016520	F: CTTCAAAGACATGTAATAACTTCCTCTCTT	59.84	452	Pentatricopeptide repeat-
			R: CATAGCCTCATCAACTTGACCATTC	61.79		containing protein
Bo-07	SNP	Bo2g016550	F: TCTACTCCTTCTTCTTCCTCTGCTA	58.07	447	Lung seven transmembrane
			R: GTTGATATTATCTGTTGCTGTGCGA	61.14		receptor family protein
Bo-08	SNP	Bo2g016610	F: CCTTGCAATAAACAAAGACACATGC	62.58	500	TIR-NBS-LRR class disease
			R: TTTGCCGCTCAGATTTTCTTGAGCC	68.93		resistance protein

2.3 RESULTS

2.3.1 Reference-based SNP identification from Genotyping-by-Sequencing data of recombinant inbred lines

Approximately four million single-end 100 bp reads were generated from Illumina sequencing of the pooled barcoded samples for each recombinant inbred. The average alignment rate of reads per sample to the TO1000 reference genome was 90.14%, with an average single alignment rate of 44.23% and an average multiple alignment rate of 44.92% (Figure 2.1). The read data was filtered for each sample of parent and recombinant inbred to remove examples with a read depth of less than ten. The filtered reads from parents A12DH and EBH527 were then compared to identify 10152 single nucleotide polymorphisms (SNPs) relative to nine chromosomes and 465 unassigned scaffolds in the reference genome. SNPs were removed from this list if data was missing in three or more RIL's, and 19 were removed that were monomorphic between A12DH and EBH527, leaving 5335 SNPs (5184 aligned to nine chromosomes and 151 to scaffolds) as genetic markers for constructing an A12DHxEBH527 linkage map.

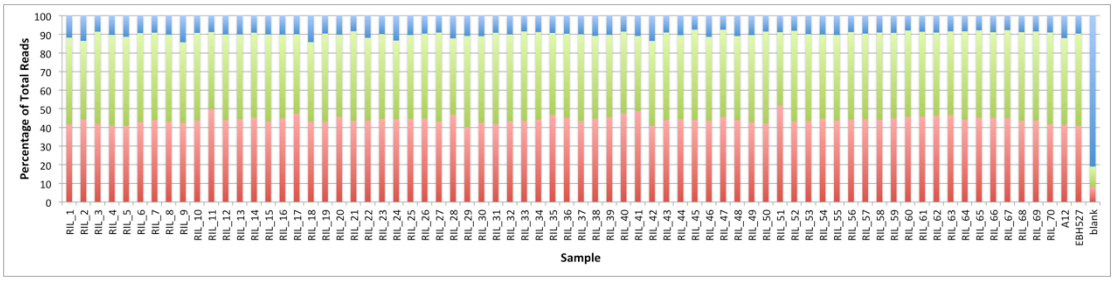


Figure 2.1. Alignment of reads to the reference genome of *Brassica oleracea* (TO1000) generated through Genotyping-by-Sequencing (GBS) of 70 F_5 A12DH x EBH527 recombinant inbred lines. Colours indicate the percentage of sequence reads that align to the reference once (RED), more than once (GREEN), or no alignment (BLUE).

2.3.2 Construction of a linkage map for the A12DH x EBH527 mapping population

A preliminary linkage map was constructed using the Haldane mapping function through MapDisto (Lorieux, 2012) with a majority of markers being assigned to nine linkage groups. Markers that had initially mapped at this stage to scaffolds in the reference genome were subsequently positioned in one of the nine linkage groups of the A12DHxEBH527 map. This provided 1616 usable markers that were distributed amongst 387 unique positions (separated by at least one recombinant in the mapping population) and a total map distance of 577.31 cM. A single marker from each position was chosen which retained a total map distance to 556.08 cM. This representative set of 387 markers was used to generate a new core map of markers in nine linkage groups using the Lander-green algorithm (Figure 2.2). No re-ordering of markers was observed compared with the preliminary map, however expansion of each linkage group was apparent resulting in a total map distance of 857.90 cM.

The error logarithm of the odds (LOD) score (Lincoln and Lander, 1992) was calculated for both maps, and identified 17 genotyping errors in the Haldane map, and eight in the Lander-Green map (Table 2.3). Throughout the core set of 387 markers, this gave genotyping error rates of 4.39% and 1.80% retrospectively. Given the phred score > 20 should have yielded a base score accuracy of 99%, the map generated through the Lander-green algorithm was used for all subsequent analysis. A final calculation of the pairwise recombination fractions was performed on the map to ensure all markers were correctly positioned. A low recombination score was identified for each linkage group (Figure 2.3).

Table 2.3. Genotyping errors identified by error logarithm of the odds (LOD) score (Lincoln and Lander, 1992) in maps generated for the A12DHxEBH527 mapping population of *Brassica oleracea* using the Haldane mapping function in MapDisto (left), and the Lander-Green algorithm in rQTL (right).

Haldane mapping function			Lander-Green algorithm		
Chr	Marker	Error LOD	Chr	Marker	Error LOD
1	C1_9185330	4.399104	9	C9_49680324	5.008935
1	C1_39339432	4.354891	7	C7_36665321	4.389245
7	C7_22774980	4.338783	3	C3_55672048	4.317023
1	C1_34895284	4.338533	8	C8_31557217	4.292683
8	C8_31557217	4.321148	3	C3_55672048	4.263668
8	C8_30557258	4.308828	1	C1_9185330	4.218287
5	C5_2764043	4.307996	5	C5_2764043	4.112729
4	C4_6777362	4.303254	9	C9_53266711	4.069941
7	C7_45047380	4.298430			
7	C7_36665321	4.274911			
3	C3_55672048	4.266879			
2	C2_44720126	4.146484			
3	C3_55672048	4.137833			
2	C2_21261609	4.097974			
4	C4_43251317	4.097540			
9	C9_53266711	4.087518			
7	C7_30485203	4.031587			

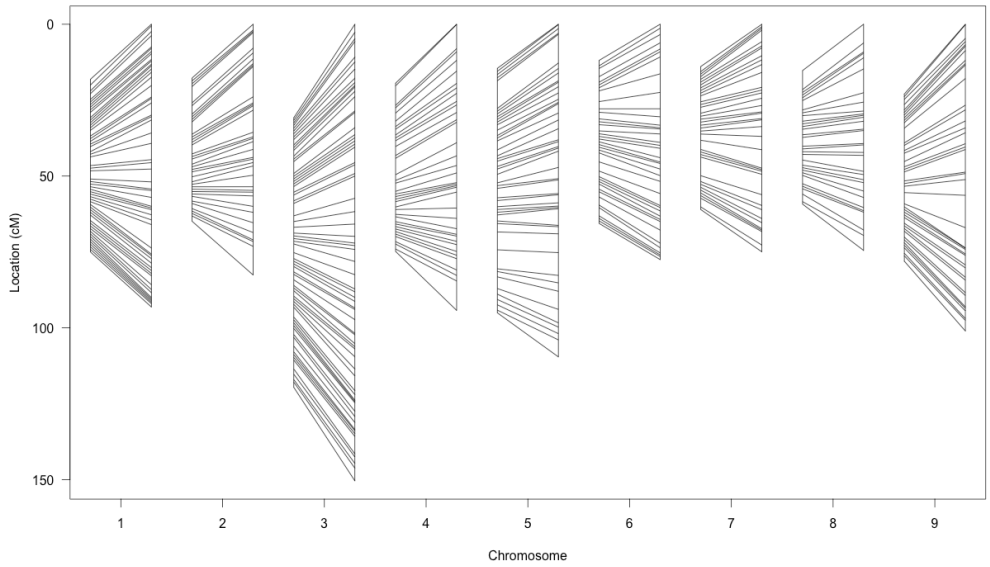


Figure 2.2. Comparison of *Brassica oleracea* linkage maps. Markers developed through genotyping-by-sequencing of F_5 recombinant inbred population of A12DH x EBH527. The Haldane mapping function in MapDisto (left-side of each chromosome) created a total map distance of 556.08 cM. The Lander-Green algorithm in rQTL (right-side) created a total map distance of 857.90 cM.

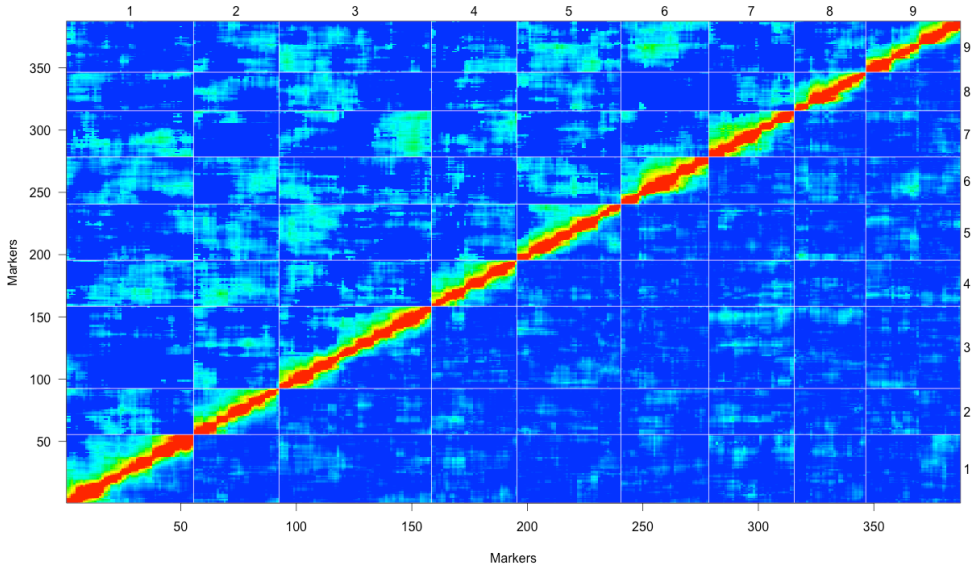


Figure 2.3. Pairwise recombination fractions and LOD scores of all markers in the linkage map generated for the A12DHxEBH527 mapping population of *Brassica oleracea* through the Lander-Green algorithm in rQTL. The colour spectrum from RED to BLUE indicates the recombination fraction from small to large, respectively.

2.3.3 Characterisation of white rust phenotypes in *Brassica oleracea* mapping population

An eight-class disease severity scale (Figure 2.4) was used to determine the phenotypic interaction of experimental lines following inoculation of cotyledons with a standard isolate of *A. candida* race 9. The phenotype classes represent a visual interpretation of the ability of the pathogen to develop within the host. A class '0' phenotype represents the greatest degree of inhibition, with no evidence of pathogen presence or host response. Phenotype classes 1-3 indicate differing host responses that have inhibited pathogen development: class 1 shows small necrotic lesions and no visible sporulation; class 2 shows a yellowing surrounding the site of infections with sporulation inhibited; and class 3 shows flaccid grey host tissue surrounding the site of infection with sporulation partially inhibited. Differing degrees of sporulation are apparent in phenotype classes 4-7. Class 7 showed the greatest amount of unrestricted sporulation on the lower cotyledon surface, and often sporulation on the upper surface. In class 4, the pathogen is unable to develop through the leaf tissue or away from the site of infection, although small white blisters could often be observed in the immediate vicinity to the site of inoculation. In class 5, sporulation is visible on both the upper and lower surface of the leaf but is confined to the site of infection, suggesting the pathogen has penetrated through the cotyledon but is restricted from extending hyphae laterally. This was similar to class 6, yet here no sporulation was detectable on the upper surface.

Figure 2.4. An eight-class scale of phenotypic variation observed in cotyledons of *Brassica oleracea* following inoculation with *Albugo candida* race 9 (AcBoWells) on the upper (left), and lower leaf surfaces (right) ten days post inoculation. The degree of host response and pathogen reproduction observed on upper and lower surfaces was recorded ten days post inoculation ranging from: **0** = no visual presence of the pathogen or host response; **1** = discrete necrotic lesions of host tissue confined to the site of inoculation on the upper surface of the leaf and no visible rust pustules; **2** = yellowing of host tissue on the upper and lower surfaces of the leaf with occasional small rust pustules; **3** = flaccid grey host tissue confined to site of infection on the upper and lower surface of the leaf with inhibited pathogen development; **4** = no host response but numerous large pustules visible that are confined to the site of infection on the upper surface; **5** = small pustules visible on upper and lower surfaces in tissue surrounding the site of inoculation; **6** = no host visible host response but large pustules on the lower surface and confined in tissue below the site of inoculation; and **7** = large pustules visible mostly on the lower surface and of the leaf migrating out from the site of infection.

Figure 2.5. Phenotypes observed on the upper (top row) and lower (bottom row) cotyledon surfaces of *Brassica oleracea* accessions A12DH (left), EBH527 (centre left), F_1 hybrid of A12DH x EBH527 (centre right), and F_1 hybrid EBH527 x A12DH right, ten days post inoculation with *Albugo candida* race 9.

2.3.4 Inheritance of white rust resistance in cotyledons to *Albugo candida* race 9

The phenotypes of parent accessions A12DH and EBH527 were observed across ten replicates. A12DH consistently showed a phenotype of 5, whereas EBH527 consistently showed a phenotype of 0 (Figure 2.5). Three F_1 lines of A12xEBH527 and the reciprocal EBH527xA12DH were tested, and all showed phenotype of 5 (Figure 2.5), suggesting that the resistance inherited from EBH527 is recessive.

A sample of 600 F_2 individuals was tested for response to inoculation with *A. candida* and a ratio of 357 resistant versus 243 susceptible lines was observed. The 600 F_2 individuals were tested as subsamples grown in nine different propagators. A one-way analysis of variance was conducted across the mean phenotypes of each propagator. No significant variation between the means of each phenotype across the nine subsamples (Figure 2.6). Chi-squared tests were performed for several genetic models, and an expected ratio of 9 resistant to 7 susceptible provided the best fit of the observed data (chi-square=2.575; p=0.109), suggesting two genes with either complementary or recessive epistatic interactions. An additive interaction for resistance can be excluded, because a 7:9 of resistance to susceptible would be expected in this case. However, recessive epistasis is a possibility, where the recessive allele of one gene masks the effect of either allele of the second gene.

Amongst 70 F_5 recombinant inbreds, 27 were resistant and 43 were susceptible. By this generation, 97% of the inbreds will be homozygous for either allele of a given gene (i.e., .48.5% AA : 48.5% BB : 3% AB). For a single gene model with the residual heterozygous class having a susceptible phenotype (48.5% resistant : 51.5% susceptible), the observed data had a chi-square of 2.763 with 1 degrees of freedom (p= 0.01).

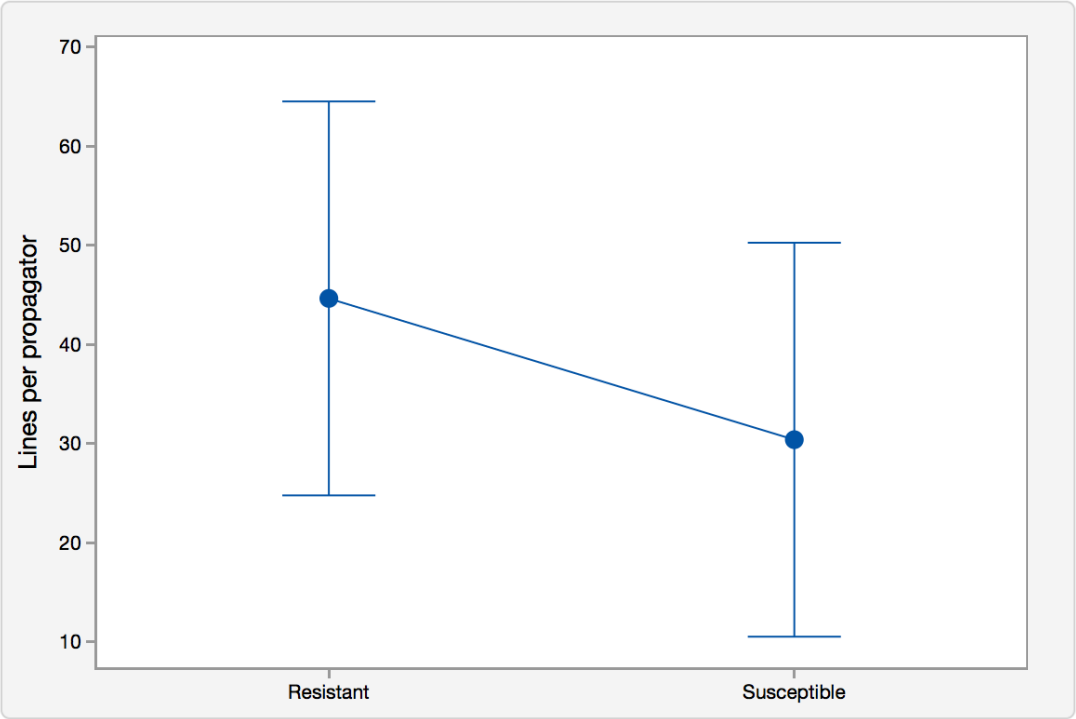


Figure 2.6. One-way analysis of variance of phenotypes observed across nine propagators containing 600 F_2 lines of *Brassica oleracea* A12xEBH527 ten days after inoculation with *Albugo candida* race 9. The mean of the resistant group was 44.63 with a standard deviation of 33.15, giving a 95% confidence interval between 24.76 and 64.49. The mean of the susceptible group was 30.38 with a standard deviation of 16.54, giving a 95% confidence interval between 10.51 and 50.24. p>0.005.

2.3.5 Standard interval mapping of a major locus for white rust resistance in EBH527

For QTL mapping, three replicates of the F_5 inbred population were tested using a randomised block design. A linear mixed model was applied to the phenotype data to calculate the effect of genotype and experimental position. Most of the variation (86%) was attributed to the genotype, with 0% attributed to position and the remaining 14% identified as residual. Average phenotypes were subsequently calculated from the three replicates and used for QTL mapping (Appendix 3).

Conditional genotype probabilities were calculated on a grid with a density of 1 cM assuming a genotyping error probability of 0.001. Standard interval mapping was performed in RQTL (Broman *et al.*, 2003) using three methods: 1) a maximum likelihood estimation under a mixture model; 2) a Haley-Knott regression performed using approximations of the mixture model;

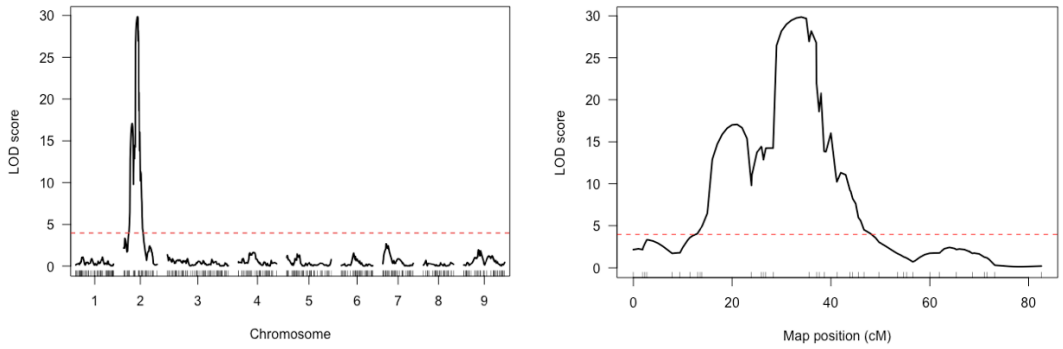
and 3) a multiple imputation method that was applied instead of the maximum likelihood estimation (Broman and Sen, 2009).

For both maximum likelihood estimation and Haley-Knott regression, 1000 permutations of the genotype probabilities were conducted to calculate genome wide LOD significance threshold of 3.96 for a 5% confidence interval. For the multiple imputation method, 100 imputations were performed, calculating the genome wide LOD significance threshold as 4.09 for a 5% confidence interval. The maximum likelihood estimation detected single QTL with a LOD score of 29.9, located on chromosome 2 at 34.00 cM between markers C2_3711838 and C2_4567864, the Haley-Knott regression detected a single QTL with a LOD score of 23.3, situated on chromosome 2 at 36.00 cM between markers C2_3711838 and C2_4978392, and the multiple imputation method detected single QTL with a LOD score of 28.7, situated on chromosome 2 at 33.00 cM between markers C2_3711838 and C2_4567864 (Figure 2.7).

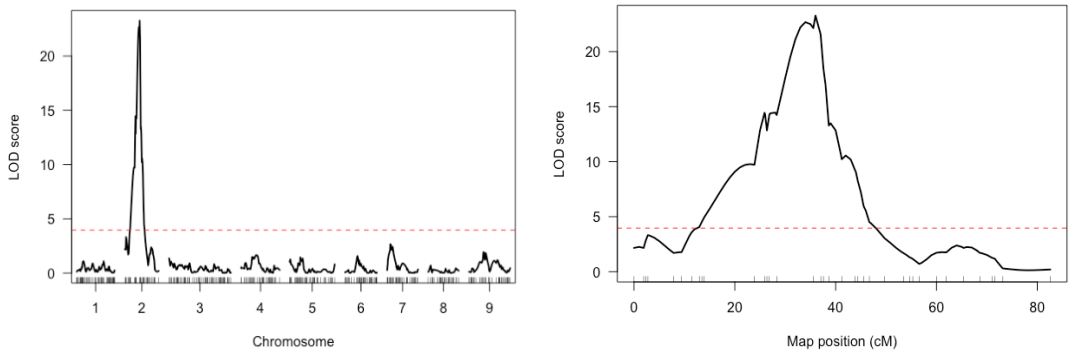
2.3.6 Multiple QTL mapping

A multiple QTL scan was conducted on each model controlling the primary locus on chromosome 2 to make any minor effects more apparent. The genotype probabilities were calculated with an error of 0.01 and a step size of 1 cM for maximum likelihood and Haley-Knott regression, and 100 imputations were performed for the multiple imputation method. The marker closest to the peak locus identified the physical position of 4567846. This position was used as an additive covariant in the single QTL scans shown in Figure 2.8, and no further significant QTLs could be detected across the rest of the genome.

A



B



C

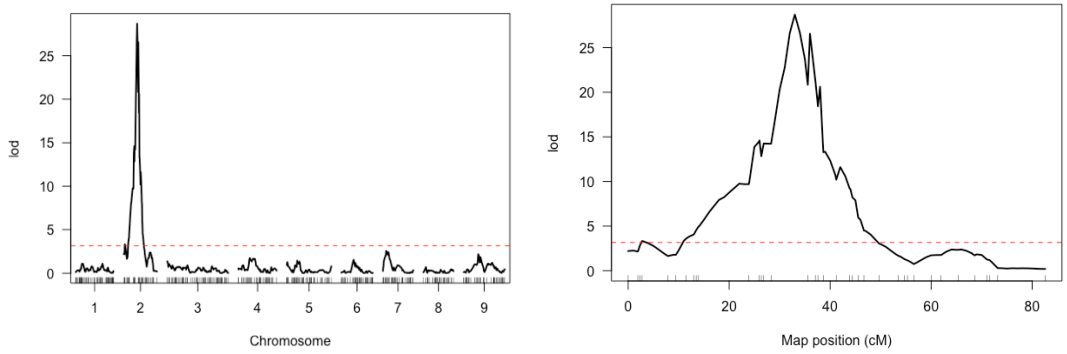
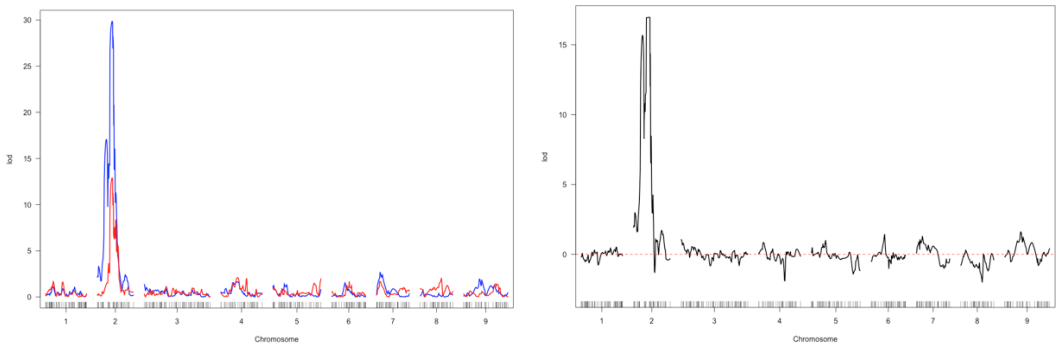
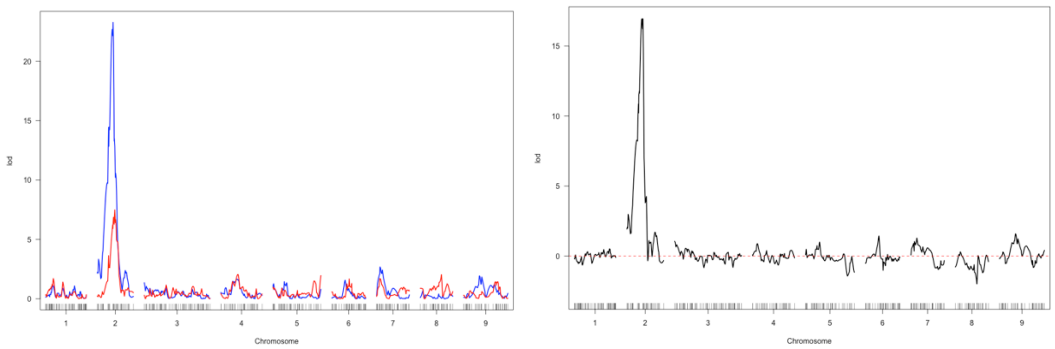


Figure 2.7. Mapping of resistance to *Albugo candida* (AcBoWells) in *Brassica oleracea* using an F_5 A12DH (susceptible) x EBH527 (resistant) mapping population. **A**, shows the genome scan using the maximum likelihood algorithm (logarithm of the odds (LOD) significance threshold of 3.96 for a 5% confidence interval); **B,** shows the scan using Haley-Knott regression (LOD significance threshold of 3.96 for a 5% confidence interval); and **C,** shows the scan using multiple imputations (LOD significance threshold of 4.09 for a 5% confidence interval). Phenotypes were recorded ten days post inoculation of cotyledons and scored using a eight class phenotype scale of resistance and differing phenotypes of susceptibility. **Left:** whole genome. **Right:** chromosome 2.

A



B



C

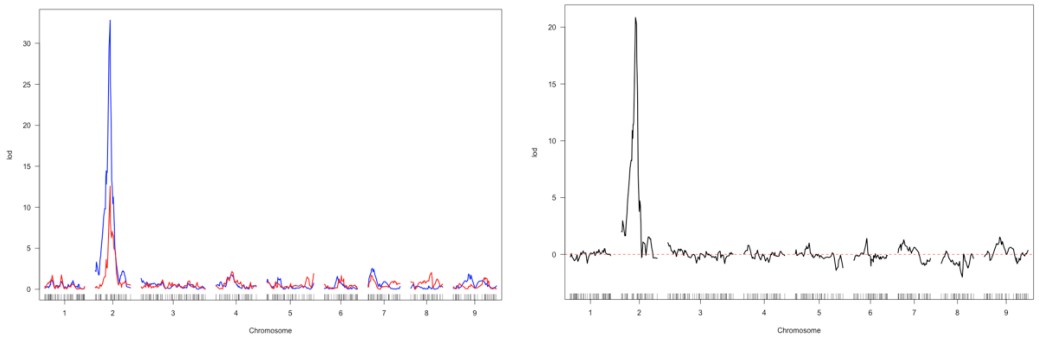


Figure 2.8. Composite interval mapping of resistance to *Albugo candida* (AcBoWells) in *Brassica oleracea* using an F_5 A12DH (susceptible) x EBH527 (resistant) mapping population, controlling the marker most tightly linked to the major effect on chromosome 2 (C2_4567846). **A**, shows the genome scan using the maximum likelihood algorithm; **B,** shows the scan using Haley-Knott regression; **C,** shows the scan using multiple imputations. Phenotypes were recorded ten days post inoculation of cotyledons and scored using a five phenotype scale of resistance and differing phenotypes of susceptibility. Left: single QTL genome scan (**blue**) and single genotype scan with C2_4567846 used as an addative covarient (**red**). Right: effect of controling C2_4567846 ilistrated by subtracting the the single QTL with the additive covarient from the standard model.

2.3.7 Two-dimensional QTL genome scan

A two-dimensional genome scan was performed to begin searching for multiple linked or interacting QTLs underling resistance to *A. candida* in the mapping population. By taking into account large effect QTLs, the multidimensional approach reduces residual variation, allowing the detection of modest QTLs. The comparison of single and two-QTL models also enables a better separation of linked QTLs. In addition, epistasis can only be detected through multiple QTL models.

By considering each pair of positions across the genome as the putative locations for two QTLs, the maximum LOD for the full model *(Sf)* and the maximum LOD for the additive model *(Sa)* were calculated (Table 2.4). Maximisation of both models was allowed at different positions. A LOD score for a test of epistasis *(Si)* could be calculated *(Si = Sf – Sa)*. Two LOD scores that indicate evidence of a second QTL were also calculated; the comparison of the full model to the best single-QTL model *(Sfv1)*, and the comparison of the additive model to the best single-QTL model *(Sav1).* These five thresholds were calculated using the multiple imputation method as the model accounts for missing genotype the putative QTL locations (Broman and Sen, 2009). Pairs of chromosomes were reported for which one or both of the following conditions hold true: Sf ≥ Sf and (Sfv1 ≥ Sfv1 or Si ≥ Si); or Sa ≥ Sa and Sav1 ≥ Sav1.

Table 2.4. Predicted second gene effects for resistance to *Albugo candida* in *Brassica oleracea* A12DH x EBH527 F_5 mapping population calculated using a multiple imputation genome scan. Significance thresholds for the LOD of the full model *(Sf),* LOD of the additive model *(Sa),* test for epistasis *(Si),* the comparison of the full model to the best single-QTL model *(Sfv1)*, and the comparison of the additive model to the best single-QTL model *(Sav1)* were calculated using 100 permutations.

	Position 1 full	Position 2 full	*Sf*	*Sfv1*	*Si*	Position 1 additive	Position 2 additive	*Sa*	*Sav1*
Significance threshold	-	-	9.29	7.36	6.06	-	-	6.38	3.13
C2:C2	30	35	40.4	7.56	6.2756	32.5	35	34.1	1.28
C2:C5	35	95	42.2	9.38	6.2840	35.0	95	35.9	3.10
C2:C8	35	40	38.9	6.08	2.1708	35.0	50	36.7	3.91
C2:C9	35	75	37.8	5.02	-0.0315	35.0	75	37.6	5.05

When comparing the full two-QTL model with the best single fit QTL model, there is evidence of a second QTL on chromosome 2 and another on chromosome 5 with allowance for epistasis. When comparing the full additive QTL model without allowance for epistasis with the best single QTL model, there is evidence of additional QTLs with an additive effect on chromosomes 8 and 9. The effects of putative linked loci on chromosome 2, and unlinked loci on chromosomes 5, 8 and 9 were assessed by analysing phenotype as a function of the genotype of the most tightly linked markers (Figure 2.9).

The results reveal recombinants homozygous for EBH527 at the inferred positions 3711838 and 4567864 on chromosome 2 that exhibit a resistant phenotype. A heterozygous allele at either marker combined with a homozygous allele at the other creates a susceptible phenotype, suggesting a single or two closely linked recessive genes. This is conferred by all recombinants heterozygous at both positions showing a susceptible phenotype. Recombinants homozygous for the A12DH allele at position 3711838 and either heterozygous, or homozygous for the EBH527 allele at position 4567864 appear to be segregating for resistance. This may indicate recombination between the two markers. All lines homozygous for the A12DH allele at both positions confer resistance.

The interaction between QTLs on chromosomes 2 and 5 again show all recombinants resistant when homozygous for the EBH257 allele at both loci. Apart from one outlier, all lines homozygous for the EBH527 allele at the chromosome 2 locus and heterozygous at the C5 locus are resistant, whilst either a heterozygous or homozygous A12DH allele on chromosome 2 confers susceptibility in all cases apart from two outliers. This may indicate that a recessive gene on chromosome 2 and a dominant gene on chromosome 5 both contribute to the full resistant phenotype.

The additive effect of the locus on chromosome 8 requires a homozygous EBH527 allele at the chromosome 2 locus for a resistant phenotype. A heterozygous allele on chromosome 8 creates the most notable phenotypic variation from the additive and interactive effects with a homozygous EBH527 allele on chromosome 2. This may indicate that the additive effect on chromosome 8 is recessive. However, a resistant phenotype is observed when a

homozygous EBH527 chromosome 2 allele is combined with a homozygous A12DH allele on chromosome 8, indicating the gene may be monomorphic between the two lines.

The additive effect of the chromosome 9 locus requires a homozygous EBH527 allele at the chromosome 2 locus for a resistant phenotype. A heterozygous allele at the chromosome 9 combined with a homozygous EBH527 allele on chromosome 2 creates a resistant phenotype in all cases, which may indicate that the chromosome 9 allele is dominant.

The predicted locations of the putative QTLs were defined as a model in order estimate the effects of each locus. For this purpose, the multiple imputation method was used to compensate for missing genotype data. By pulling out the imputed genotypes at each location, a five QTL model was generated allowing for interactions between the two QTLs on chromosome 2, and between the primary peak on chromosome 2 with the chromosome 5 QTL. The overall fit of the model provided a LOD score of 51.6 relative to null model, with 96.64% of the phenotypic variance being accounted for. Each locus was dropped and reintroduced in succession, allowing comparison to be made between the full model and the model with the term omitted. The results provided strong evidence for loci on chromosomes 2, 5 and 9, and for interactions between the two loci on chromosome 2 and between chromosomes 2 and 5. The chromosome 8 locus was found to be insignificant (Table 2.5).

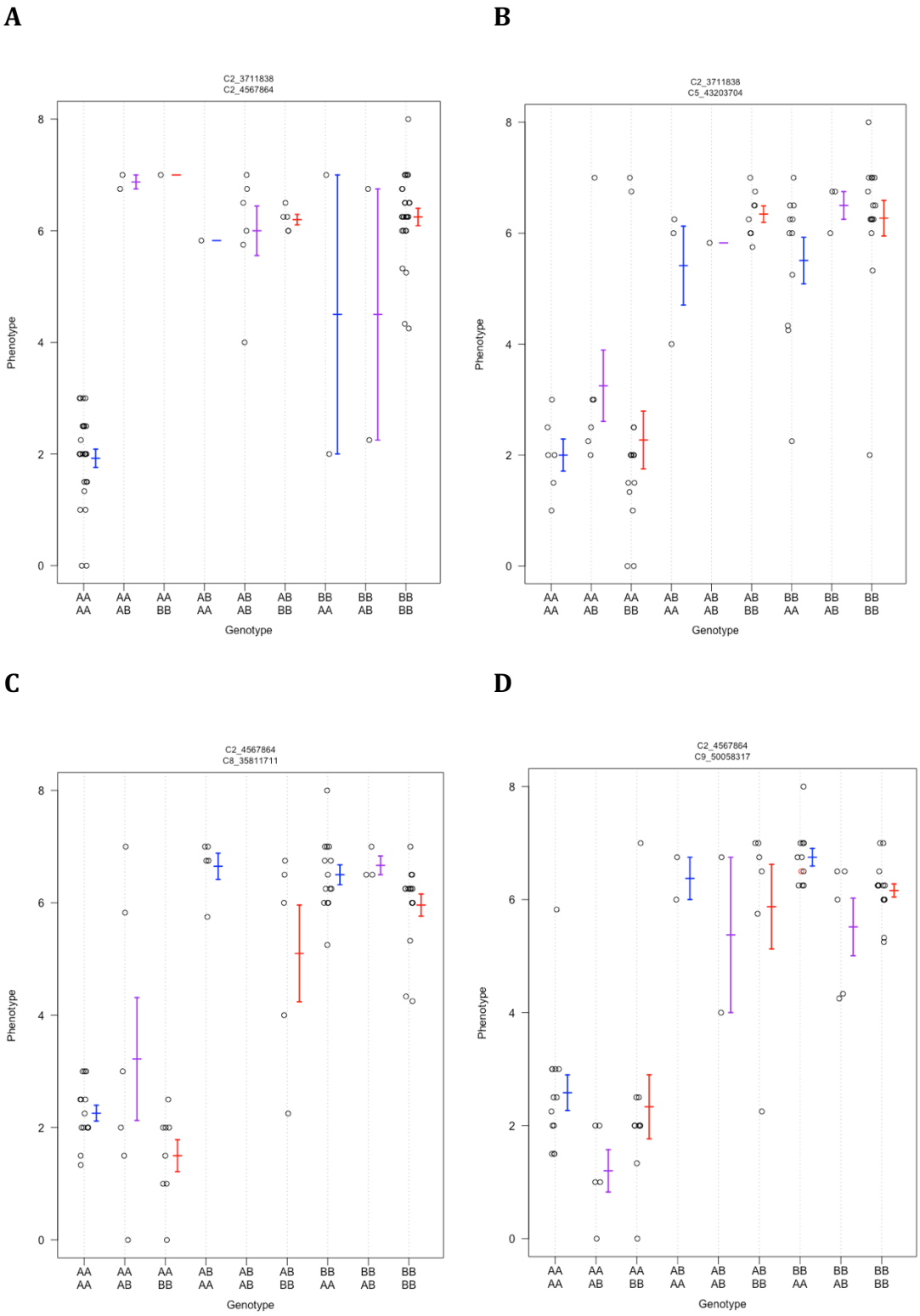


Figure 2.9. Dot plots of *Brassica oleracea* F_5 recombinant inbred lines with white rust resistant phenotype as a function marker genotypes identified as interactive, **A** and **B,** and additive **C** and **D,** using a two dimensional two QTL genome scan. Black dots correspond to observed genotypes and red dots correspond to missing, consequently imputed genotypes.

Table 2.5. Analysis of the effect of dropping each QTL independently from a five QTL model generated by identifying fixed locations of QTLs contributing to resistance to *Albugo candida* AcBoWells in *Brassica oleracea* A12DH (susceptible) x EBH527 (resistant) F5 mapping population.

Position	Type 3 sum of squares	LOD	% variance	F value	P value (F stastic)
C2 30cM	12.73	11.10	3.61	9.13	8.47e-07
C2 35cM	56.62	26.69	16.06	24.42	4.44e-16
C5 95cM	9.08	8.66	2.58	6.53	3.72e-5
C8 50cM	0.64	0.80	0.18	1.37	0.26
C9 75cM	6.90	7.00	1.99	14.88	8.11e-06
C2 30cM & C2 35cM	13.85	11.78	3.93	14.93	3.83e-08
C2 35cM & C5 95cM	5.08	5.43	1.44	5.48	9.55e-04

2.3.8 Fine-mapping eliminates NB-LRR genes as candidate genes conferring recessive white rust resistance in EBH527 at the Bo-ACA2 locus

The locus for a major effect QTL on chromosome 2 defined using the Haley-Knott regression was designated Bo-*ACA2*. Br-*ACA1* was previously mapped by Kole *et al.* (2002) as a white rust resistance locus on chromosome 4 of *B. rapa.*

A physical interval of 1,266,554 bp was estimated for Bo-*ACA2* by relating markers C2_3711838 and C2_4978392 to their physical locations on the TO1000 reference genome. This interval was wider than the 856,026 bp defined by both standard interval, and multiple imputation genome scans, encompassing more potential candidate genes and both putative QTLs identified in the two dimensional two QTL genome scan (Table 2.6). Four recombinants (RIL_22, RIL_38, RIL_49 and RIL_59) were identified which reduced the interval to 410,528 bp between markers C2_4567864 and C2_4978392 (Figure 2.10). This fine map interval spans 67 annotated genes in the TO1000 reference including seven that encode TIR-NBS-LRR proteins.

Importantly, all RILs that are homozygous for EBH527 DNA across the interval had a resistant phenotype, all RILs that are homozygous for A12DH DNA had a susceptible phenotype, and all heterozygous RILs are susceptible. These

results therefore confirmed the predictions made from the F_1 phenotype (Figure 2.5) that the resistance is recessive.

A new marker technology called Resistance gene enrichment and sequencing (RenSeq) (Jupe *et al.*, 2013, Jupe *et al.*, 2014) was used to investigate whether Bo-*ACA2* resistance is conferred by one of the TIR-NBS-LRR genes. This method enables capturing of high quality sequence of NB-LRR genes across the whole genome of a plant accession. Following quality control, RenSeq of genomic DNA from A12DH yielded a total of 283,535 paired end reads including 15251 (5.38%) that did not map to the TO1000 reference, 88253 (31.13%) that aligned once and 180031 (63.50%) aligned more than once to the reference. Sequence comparison of matching alleles identified 7114 SNPs between A12DH and TO1000, supported by a depth of coverage of ten or greater. Similarly, RenSeq of genomic DNA from EBH527 yielded a total of 1141966 paired end reads; including 26753 (2.34%) that did not map to the reference, 273202 (23.92%) that aligned once and 842011 (73.73%) that aligned more than once to the reference. Sequence comparison identified 10788 SNP's between EBH527 and TO1000, supported by a depth of coverage of ten or greater. By combining the datasets, 9723 unique polymorphisms were identified between A12DH and EBH527 in 514 predicted gene models.

Of the seven NB-LRR genes within the *ACA2* interval of TO1000, five possessed unique polymorphisms between both parents supported by a depth of coverage greater than ten. In Bo2g014340 a polymorphism was detected between EBH527 and TO1000, but an insufficient depth of coverage in A12DH meant this could not be identified as a certain polymorphism between the two parents. In Bo2g016440, no polymorphisms could be detected between EBH527 and TO1000, and again the A12DH locus was not supported by a sufficient depth of coverage to accurately identify any mutations (Table 2.6). Of the genes identified as polymorphic, Bo2g016470 and Bo2g016610 were used for further analysis owing comparatively high mean depth of coverage and ability to exclude the remaining candidate NB-LRRs if the parental mutations did not cosegregate with the phenotype across the recombinants.

In Bo2g016470 a SNP at position 4,840,023 from thymine in A12DH to guanine in EBH527, and in Bo2g016610 a SNP at position 4,914,224 from

thymine in A12DH to adenine in EBH527 was used to genotype the key recombinants. Neither cosegregated the phenotype, narrowing the interval to 64,445 bp containing 13 predicted genes in the reference genome.

Table 2.6. Mean depth and breadth of sequence coverage for Nucleotide Binding Site-Leucine-Rich Repeat (NBS-LRR) genes in *Brassica oleracea* reference genome TO1000 within the fine mapping interval of Bo-*ACA2.* Sequence was generated from parent accessions A12DH (susceptible) and EBH527 (resistant) using a new exome capture method called Resistant Gene Enrichment Sequencing (RenSeq). (Juper et al 2013; Jupe et al, 2014) Corresponding homologs were identified for three genes on the top arm of chromosome 5 in *Arabidopsis thaliana*.

Gene ID	A12DH depth	EBH527 depth	A12DH breadth	EBH527 breadth	Supported Mutations
Bo2g014110	19.3	49.5	42.3	51.7	Yes
Bo2g014320	25.6	185.0	64.3	100	Yes
Bo2g014340	7.3	185.4	58.3	88.6	No
Bo2g014350	11.3	443.5	52.9	100	Yes
Bo2g016440	1.1	10.6	37.1	42.5	No
Bo2g016470	80.7	126.0	99.2	76.0	Yes
Bo2g016610	63.2	119.4	100	91.0	Yes

Accession	Phenotype	C2_4567864	Bo2g016470	Bo2g016480	Bo2g016490	Bo2g016500	Bo2g016510	Bo2g016520	Bo2g016550	Bo2g016610	C2_4978392
EBH527	1.0	AA	AA	AA	MM	MM	MM	AA	AA	AA	AA
A12	7.0	BB	BB	BB	MM	MM	MM	BB	BB	BB	BB
RIL_59	1.5	AA	AA	AA	MM	MM	MM	BB	BB	BB	BB
RIL_49	2.3	AB	AB	AA	MM	MM	MM	AA	AA	AA	AA
RIL_38	6.5	AB	BB	BB	MM	MM	MM	BB	BB	BB	BB
RIL_22	5.8	AA	AB	AB	MM	MM	MM	AB	AB	AB	AB

Figure 2.10. Summary of phenotype and genotype information for F_5 A12DH (susceptible) x EBH527 (resistant) recombinant inbred lines (RIL) of *Brassica oleracea*. The inbreds were phenotyped for white rust resistance following cotyledon inoculation with *Albugo candida* race 9 (AcBoWells). Genotype information spans the map interval on chromosome 2 for Bo-*ACA2* defined using a Haley-Knott regression mapping of Genotyping-by-Sequencing markers (C2_4567864 and C2_4978392). Allelic variation in eight genes within the interval was used to generate new markers including two genes that encode TIR-NBS-LRR proteins (Bo2g016470 and Bo2g016610). A marker generated within Bo2g016480 co-segregated with inbred phenotypes; this gene encodes a GDSL lipase-like protein. For genotype scores: AA indicates homozygous for an EBH527 allele; BB indicates homozygous for an A12DH allele; AB indicates heterozygous alleles; and MM indicates monomorphic alleles. Interaction phenotypes were 'resistant' (ranging in a score of 1 to 3) or 'susceptible' (ranging in a score from 6 to 8).

2.3.9 Comparative sequence analyses identifies a single candidate resistance gene

The centre point of the interval was taken for further reference-based marker design to search for polymorphisms between the parents. Bo2g016550, encoding a lung seven transmembrane receptor family protein, was found to be polymorphic and was used to genotype the recombinants. The marker did not cosegregate with phenotype, therefore excluding this gene as a candidate and narrowing the interval to 25,637 bp that contains seven predicted genes in TO1000. Within this fine interval, two genes were found to be monomorphic including Bo2g016540 (encoding a kinase family protein) and Bo2g016530 encoding a glycogenin-1 protein).

Reference based marker design and genotyping revealed Bo2g016520, encoding a pentatricopeptide repeat–containing protein, also does not

cosegregate with phenotype and narrowing the interval to 18,033 bp that contains four predicted genes in the reference genome. These were Bo2g016510 (a G-type lectin S-receptor-like serine/threonine-protein kinase), Bo2g016500 (a protein of unknown function), Bo2g016490 (an ethylene-responsive transcription factor) and Bo2g016480 (a GDSL lipase).

Primers flanking each of the four remaining candidate genes were designed from the reference genome and whole gene amplicons were obtained from A12DH and EBH527. Multiple sequencing reactions were performed on each sample and the reads were aligned back to the reference genome. Three genes (Bo2g016510, Bo2g016500 and Bo2g016490) were found to be monomorphic; whereas a single SNP was found in Bo2g016480 at position 4,846,936 (with adenine in A12DH and TO1000, and thymine in EBH527). This marker was used to genotype the recombinants. The genotype of the recombinants did cosegregate with their corresponding phenotype, thus suggesting that Bo2g016480 is the only remaining candidate gene (Figure 2.10).

The SNP in Bo2g016480 encodes a non-synonymous change in the predicted codon of the first reading frame, translating into an isoleucine in EBH527 and a lysine in A12DH and TO1000. The EBH527 protein is predicted to have a hydrophobic side chain as opposed to an electrically charged side chain, thus affecting the hydrophobicity and isoelectric point. (Kearse et al, 2012). The secondary structure of the A12DH protein also appears to have an additional beta strand in place of an elongated turn in EBH527, thus reducing the size of a predicted antigenic region in A12DH by two residues (Table 2.7).

Table 2.7. Predicted changes in the secondary structure of the *ACA2* protein (encoded by Bo2g016480) caused by the translation of either a lysine in the transcript from the A12DH allele or an isoleucine in the EBH527 allele at amino acid position 300.

Allele	Type of change	Minimum	Maximum	Length
A12DH	Coil	292	294	3
	Turn	295	304	10
	Antigenic region	301	310	10
	Beta Strand	305	307	3
	Turn	308	309	2
	Coil	310	310	1
EBH527	Coil	292	294	3
	Turn	295	309	15
	Antigenic region	299	310	12

2.3.10 Characterisation of Bo2g016480 diversity with respect to white rust resistance.

Ten replicates of 246 lines from the C genome diversity collection were phenotyped following cotyledon inoculation with *A. candida* isolate AcBoWells to search for additional sources of resistance (Appendix 1). The set was representative of the morphological, physiological and genotypic variation across the C genome (Walley *et al.*, 2012). Twelve lines were scored as class 2, suggestive of resistance with a similar phenotype to EBH527. Transcriptome sequence was available for the set of 112 lines (Barker, unpublished). Nucleotide sequence of Bo2g016480 was extracted from the data. The A12 residue was highly conserved, with all accessions apart from EBH527 having a lysine at the mutated position. Phylogenetic analysis of both the nucleotide and translated protein sequence indicate that EBH527 has a unique allele of Bo2g016480 amongst the line assessed (Figure 2.11 and 2.12). Interestingly, EBH527 did not group with four other accessions (HRIGRU007514, C04044, C044045 and HRIGRU011555) that exhibited resistance to *A. candida*, suggesting that they contain alternative genes conferring white rust resistance.

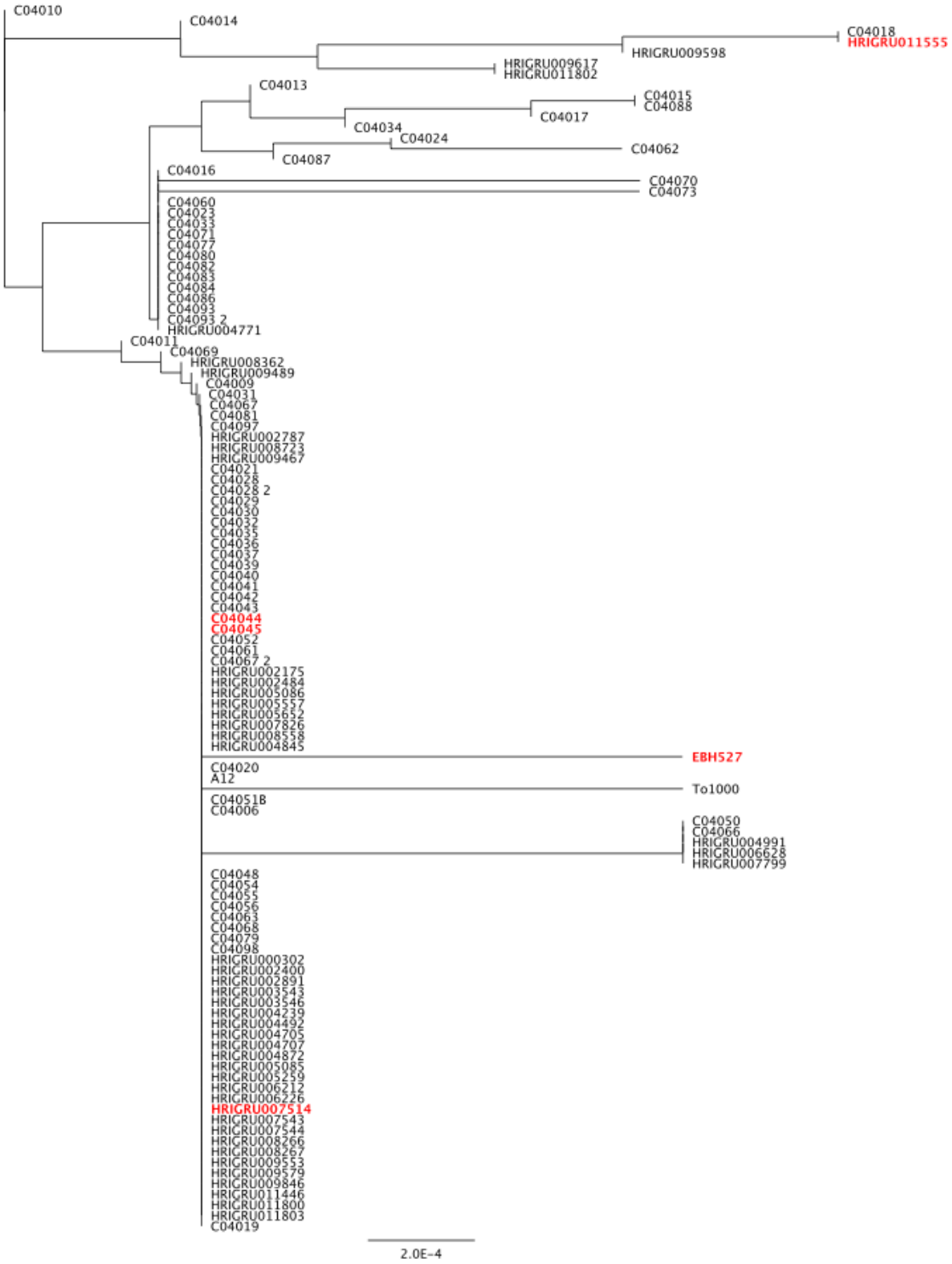


Figure 2.11. Phylogeny of coding nucleotide sequence variation in alleles of Bo2g016480 from a C genome diversity collection of *Brassica oleracea* and related species generated using Jukes-Cantor genetic distance model and Neighbor-Joining build method with no outgroup. This gene was identified as a candidate determinant for Bo-*ACA2* resistance to *Albugo candida* race 9 (AcBoWells). **Red** labels indicate alleles from accessions that are resistant to AcBoWells; and **Black** labels indicate alleles from susceptible accessions.

Figure 2.12. Phylogeny of protein sequence variation in alleles of Bo2g016480 from a C genome diversity collection of *Brassica oleracea* and related species generated using Jukes-Cantor genetic distance model and Neighbor-Joining build method with no outgroup. This gene was identified as a candidate determinant for Bo-*ACA2* resistance to *Albugo candida* race 9 (AcBoWells). **Red** labels indicate alleles from accessions that are resistant to AcBoWells; and **Black** labels indicate alleles from susceptible accessions.

2.3.12 Mapping of a QTL conferring resistance to an Australian isolate of *Albugo candida* in A12DH

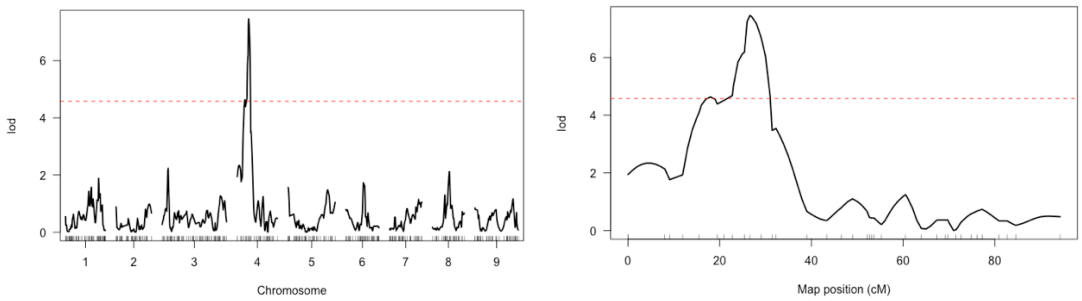
An Australian isolate of *A. candida* isolate (AcAus) was used to phenotype ten plants of both A12DH and EBH527. Both lines consistently displayed a resistant phenotype of class 1. The same isolate was used to phenotype a subset of the F_5 A12DH x EBH527 mapping population. The inbreds segregated 26 susceptible to 12 resistant (based on a mean phenotype score across three replicates), indicating that resistance is conferred in each parent by a different resistance gene. F_1 or F_2 seed was not available for testing with AcAus.

As described above for Bo-*ACA2,* initial mapping of resistance to AcAus was performed using the available GBS data for the mapping population. Conditional genotype probabilities were calculated on a grid with a density of 1 cM assuming a genotyping error probability 0.001. Standard interval mapping was performed using a maximum likelihood estimation under a mixture model, Haley-Knott regression was performed using approximations of the mixture model, and a multiple imputation method was applied to model in place of the maximum likelihood estimation (Broman and Sen, 2009). For both maximum likelihood estimation and Haley-Knott regression, 1000 permutations of the genotype probabilities were conducted to calculate genome wide LOD significance threshold of 4.58 for a 5% confidence interval. For the multiple imputation method, 100 imputations were performed, calculating the genome wide LOD significance threshold as 4.14 for a 5% confidence interval. All three methods identified a single major effect locus on chromosome 4 (Figure 2.13): the maximum likelihood estimation and multiple imputation both detected a locus at 26.60 cM between markers C4_5615273 and C4_10354410 (LOD score of 7.5), and Haley-Knott regression detected a locus at 27.00 cM between markers C4_5615273 and C4_10354410 (LOD score of 7.5).

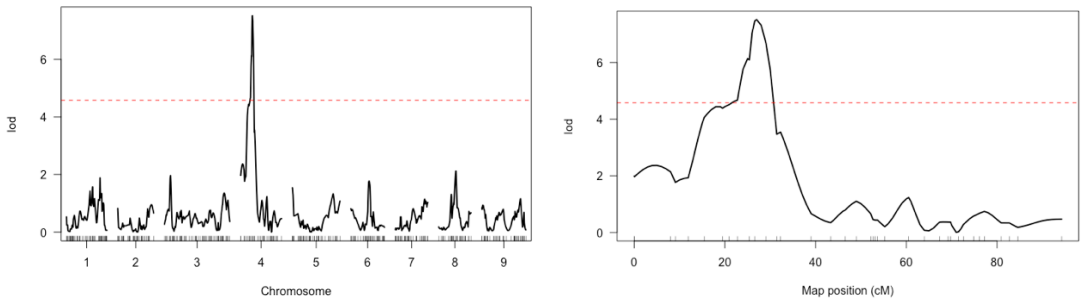
A multiple QTL scan was then conducted on each model for the primary locus on chromosome 4 to determine if any minor effects are involved. The genotype probabilities were calculated with an error probability of 0.01 and a step size of 1 cM for maximum likelihood and Haley-Knott regression, and 100 imputations were performed for the multiple imputation method. The genotype of the marker closest to the peak of the locus for each output was identified at

the physical position 8059884, and was subsequently removed and used as an additive covariant in the single QTL scans (Figure 2.14). No further significant QTLs could be detected across the rest of the genome using this method.

A



B



C

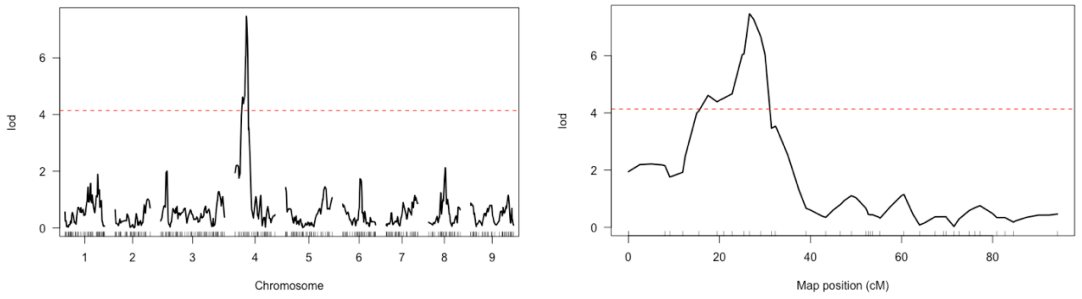
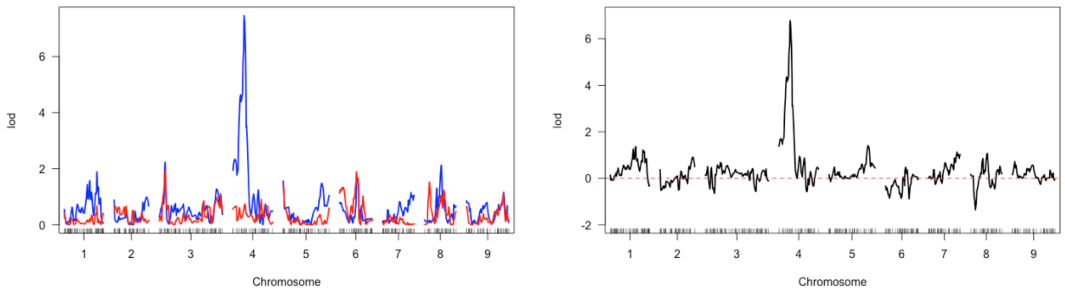
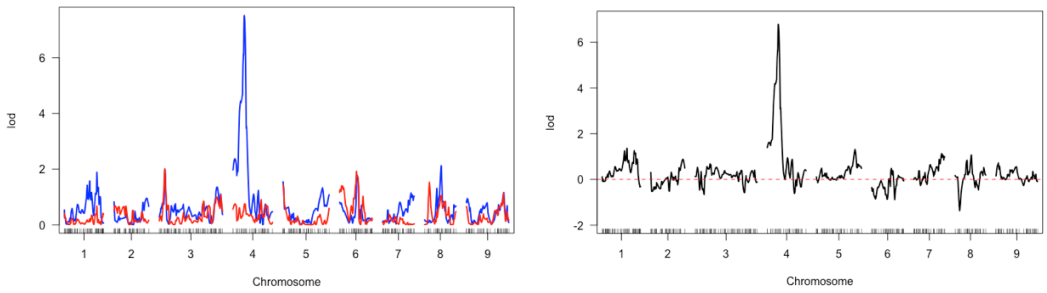


Figure 2.13. Mapping of resistance to *Albugo candida* (AcAus) in *Brassica oleracea* using an F_5 A12DH (resistant) x EBH527 (resistant) population. **A**, shows the genome scan using the maximum likelihood algorithm (logarithm of odds (LOD) of 4.58 for a 5% confidence interval); **B,** shows the scan using Haley-Knott regression (LOD of 4.58 for a 5% confidence interval); and **C,** shows the scan using multiple imputations (LOD of 4.14 for a 5% confidence interval). Phenotypes were recorded ten days post inoculation of cotyledons and scored using an eight class phenotype scale. **Left:** whole genome. **Right:** chromosome 4.

A



B



C

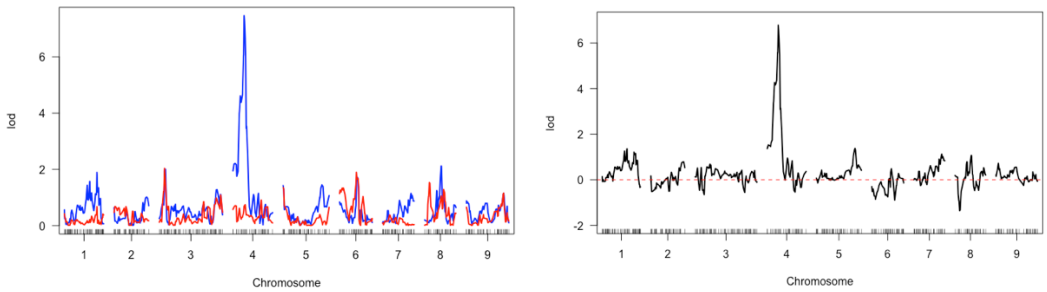


Figure 2.14. Composite interval mapping of resistance to *Albugo candida* (AcAus) in *Brassica oleracea* using an F_5 A12DH (resistant) x EBH527 (resistant) mapping population, controlling the marker most tightly linked to the major effect on chromosome 4 (C4_8059884). **A**, shows the genome scan using the maximum likelihood algorithm; **B,** shows the scan using Haley-Knott regression; **C,** shows the scan using multiple imputations. Phenotypes were recorded ten days post inoculation of cotyledons and scored using a five class phenotype scale of resistance and differing phenotypes of susceptibility. Left: single QTL genome scan (**blue**) and single genotype scan with C4_8059884 used as an addative covarient (**red**). Right: effect of controling C4_8059884 illustrated by subtracting the the single QTL with the additive covarient from the standard model.

A two-dimensional genome scan was performed to begin searching for multiple linked or interacting QTLs that may underlie resistance to *A. candida* (AcAus) in the mapping population. A multiple imputation genome scan was used to obtain values for the maximum LOD of the full model *(Sf)*, maximum LOD of the additive model *(Sa)*, a LOD score for a test for epistasis *(Si)*, the comparison of the full model to the best single-QTL model *(Sfv1)*, and the comparison of the additive model to the best single-QTL model *(Sav1)*. With a 5% confidence interval and 100 permutations these values were calculated as 12.6, 7.51, 8.51, 9.86, and 4.05 retrospectively. No pairs of loci matched the criteria of $Sf \geq Sf$ and $(Sfv1 \geq Sfv1 \text{ or } Si \geq Si)$, or $Sa \geq Sa$ and $Sav1 \geq Sav1$, implying no interactive or additive QTL's could be detected. The predicted location of the putative QTL on chromosome 4 was subsequently defined as a model and compared to the null, concluding that it accounted for 57.70% of the phenotypic variance.

The interval between markers C4_5615273 and C4_10354410 was used as a focus for further analyses. All inbred lines that are either homozygous for A12DH DNA or heterozygous exhibited a resistant phenotype, indicating a dominant A12DH allele conferring resistance to AcAus. Two out of 13 inbreds that are homozygous for EBH527 across the interval also had a resistant phenotype (RIL_50 and RIL_67), which could be explained by an expected resistance inherited from EBH527 at a different location. These inbreds would therefore provide a resource for mapping resistance to AcAus in EBH527 through the development of mapping populations from the progeny. As both parents were resistant and no second QTL was detected, it is possible all or some of the lines phenotyped as resistant had both the A12DH and EBH527 *R* genes. Fine mapping within the interval is consequently problematic, as it not possible the rule out that an EBH527 locus in producing the resistant phenotype. Nevertheless, three recombinants, which if lacking the EBH527 resistance allele, suggest the interval could be narrowed between markers C4_8059884 and C4_9033847 (Figure 2.15).

Accession	Phenotype	C4_5615273	C4_6777362	C4_8059884	C4_9033847	C4_10354410
16	1	BB	BB	BB	BB	BB
28	1	BB	BB	BB	BB	BB
37	1	BB	BB	BB	BB	BB
48	1	BB	BB	BB	BB	BB
66	1	BB	AB	AB	AB	AB
54	1.03	AB	AB	AB	AB	AB
30	1.5	AA	AB	AB	AB	AB
65	1.7	BB	BB	AB	AB	AB
58	1.05	BB	BB	BB	BB	AA
34	1.1	BB	BB	BB	BB	AA
46	4.1	AA	AA	AA	BB	BB
23	4.2	AA	AA	AA	AA	AA
36	5	AA	AA	AA	AA	AA
5	5.5	AA	AA	AA	AA	AA
40	5.5	AA	AA	AA	AA	AA
50	1.2	AA	AA	AA	AA	AA
67	1.25	AA	AA	AA	AA	AA

Figure 2.15. Summary of genotype and white rust resistance phenotype for F_5 A12DH x EBH527 recombinant inbreds of *Brassica oleracea* following cotyledon inoculation with an *Albugo candida* isolate from Australia (AcAus). Genotype information spans a putative QTL for a resistance locus on chromosome 4 (centered on boxed position), identified using a maximum likelihood, Haley-Knott regression, and multiple imputation genome scan of a genetic map constructed from markers generated through Genotyping by sequencing (GBS). Genotype scores include**:** AA= homozygous for the EBH527 allele, BB= homozygous for the A12DH allele, and AB=heterozygous. Phenotypes were using a seven-class severity scale including classes 1-3 (resistant with no blisters) or classes 4-8 (with increase amounts of blisters).

One broad mapping interval spans 145 predicted gene models in the TO1000 reference genome (Appendix 3), and only includes one example of a resistance gene that encodes a CC-NBS-LRR protein. Bo4g038670 is located between 8,267,696 and 8,271,861 bp, and is a ortholog of a known CC-NBS-LRR downy mildew resistance gene in *A. thaliana* designated *RPP7* (Holub, 2007). The RenSeq data from both parents was analysed to search for polymorphisms. A12DH had a mean depth of coverage of 10.6 reads a 36.9% breadth of coverage, and EBH527 had a mean depth of coverage of 24.1 reads with a 74.8% breadth of coverage. This data was insufficient for reliable comparison of alleles, so primers flanking the gene were designed from the reference genome and used to obtain whole gene amplicons from both parents. Multiple Sanger sequencing reactions were performed on both samples and the reads were aligned back to the reference genome, revealing that EBH527 shares the same allele as TO1000 whereas the A12DH allele has 15 SNPs, a 35 bp deletion and a 7 bp insertion. The nucleotide sequence for both parents was translated into protein coding

sequence, and revealed a SNP at position 8,268,802 (a cytosine in EBH527 and TO1000, and an adenine in A12) causes a premature stop codon in A12DH, shortening the length of the predicted protein from 714 to 510 amino acids. Two additional non-synonymous mutation are also present in the A12DH allele prior to the stop codon: an adenine at position 8,271,245 in A12DH to encode aspartic acid instead of a thymine to encode valine in EBH527 and TO1000; and a thymine at position 8,271,115 in A12DH to encode alanine instead of a cytosine to encode valine in EBH527 and TO1000. A protein BLAST search was conducted on both alleles. Both contained a conserved N-terminal coiled coil domain, and a AAA ATPase domain within the NB-ARC although EBH527 and TO1000 contained an additional topology modulation protein caused by the two non-synonymous SNPs. Most notable was the lack of the leucine rich repeat domain in the A12DH allele caused by the premature stop codon. The 35 bp deletion, 7 bp insertion and remaining 13 SNPs upstream of the premature stop codon were evidence of redundancy.

2.4 DISCUSSION

The discovery of a single candidate gene that cosegregates with the phenotype of key recombinants provides strong evidence that a GDSL lipase (Bo2g016480) at the *ACA2* locus explains the broad spectrum resistance to *A. candida* race 9 in EBH527. The recessive nature of this resistance is intriguing because it may indicate that the alternative allele in accession A12DH encodes a protein that is required for compatibility with *A. candida.* In addition, the identification of additive effects of other loci indicates that alleles of additional genes are required in the genetic background to enhance resistance (or disrupt compatibility).

The application of new genotyping technology made it possible to progress rapidly in defining a narrow interval for the *ACA2* locus. RenSeq has been promoted recently as a powerful means of identifying *R*-genes in major crop species (Witek *et al.*, 2016, Andolfo *et al.*, 2014, Jupe *et al.*, 2013, Jupe *et al.*, 2014). However, *R*-genes are typically associated with dominant resistance, so GBS was critical in this current work to characterise an example of the recessive

resistance. Once a locus had been established, then high quality sequence of NB-LRR genes within the map interval was useful as a means of narrowing the interval.

Although this genetic mapping suggests that an NB-LRR gene is not involved at the *ACA2* locus, this important class of R-protein may still play a role at a secondary locus. For example, this has been demonstrated for a recessive resistance in barley to the wheat stem rust pathogen *Puccinia graminis*, in which tightly linked NB-LRR genes were found to be essential components of the resistance (Wang *et al.*, 2013). Interestingly, copies exist in both *A. thaliana* and *B. rapa* within regions associated with white rust resistance. In *A. thaliana* Col-0, the homologous GDSL lipase At5g18430 lies approximately 0.28 Mb from *WRR5* and *WRR6*; whereas in *B. rapa*, the Bra023673 homolog on chromosome 2 sits within a putative QTL conferring a minor effect to resistance to *A. candida* race 2 (Kole *et al.*, 2002). This suggests that syntenic blocks of genes are evolving resistance to *A. candida* across the *Brassicaceae*, and we can speculate that *R* gene variation within these blocks are driving the specialisation apparent within the coevolving pathogen (McMullan *et al.*, 2015).

To initiate verification that the candidate GDSL lipase is required for compatibility with *A. candida,* experiments were initiated using TDNA insertion knock-out mutants of the homologous gene in *A. thaliana* (At5g18430). These were generated in the Col-0 background (reference genome) and available from a public resource (Alonso *et al.*, 2003) supplied by the Arabidopsis Stock Centres. A knock-out or loss-of-function mutation would be expected to exhibit enhanced resistance to a virulent isolate of *A. candida.* Five independent mutant lines were obtained, and ten of each was tested with a Col-virulent isolate AcExeter. All of the lines were heterozygous for the TDNA insert. Two lines (SALK_116756 and SALK_079740) with insertions affecting the first exon had at least two individuals that exhibited no blisters and small necrotic lesions surrounding the site of inoculation. No seed germinated for one line, and results were inconclusive for the other two lines. Homozygous lines will be selected for each TDNA insertion for testing in a follow-up experiment to confirm a non-segregating resistant phenotype. Nevertheless, the preliminary results suggest that a GDSL lipase is required for susceptibility to a virulent isolate of *A. candida*.

Homologs of the candidate ACA2 lipase in *A. thaliana* and *B. rapa* have been shown to be co-expressed with At4g34760 encoding a SAUR-like auxin-responsive protein, At4g25010 encoding an integral membrane protein, and At3g20710 encoding an F-box family protein (Figure 2.16A) (Lee *et al.*, 2015). The co-expression with an integral membrane protein adds to the evidence that the ACA2 lipase is localised to the cell wall or extra cellular matrix ECM. Expression analysis in *A. thaliana* has shown the highest level of expression of At5g18430 in the stomata (Figure 2.16B) (Yang *et al.*, 2008). Since the stomata are the primary site of *A. candida* penetration, this fits with At5g18430 being involved in defence response. In addition, expression has been shown to be upregulated following treatment with abscisic acid (ABA) (Figure 1B). ABA is known to play an important role in both abiotic and biotic stress (Zhu, 2002, Seo and Koshiba, 2002).

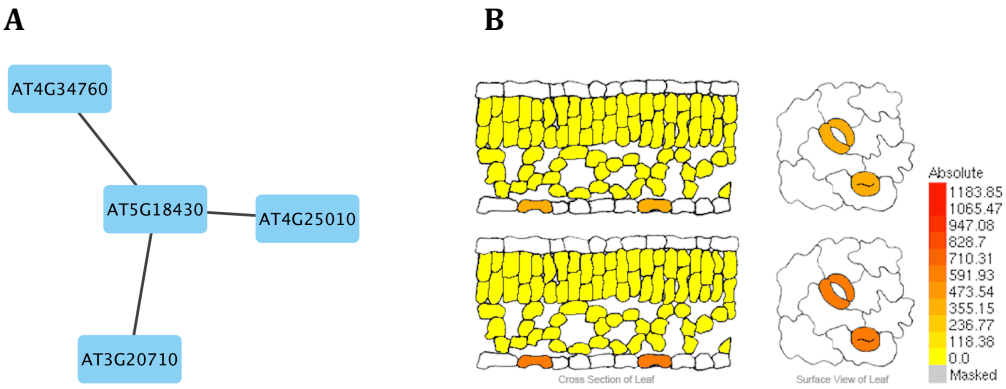


Figure 2.16. Summary of gene expression information for the ortholog in *Arabidopsis thaliana* (At5g18430) of the *Brassica oleracea* candidate white rust resistance gene *ACA2* including: **A**) a predicted functional gene network in *Arabidopsis thaliana* (Lee *et al.*, 2015); and **B)** expression in *A. thaliana* Columbia-0 following treatment with water (top) and with abscisic acid (bottom)(Yang *et al.*, 2008).

Pathogen and host lipids and lipid metabolites are expected to play a fundamental role in pathogenesis and host defence responses and elicitation of systemic acquired resistance (Shah, 2005). Induced defence mechanisms are generally thought to be activated by a signal that is generated at the primary infection site and translocated throughout the host tissue (Dangl *et al.*, 2013). GDSL lipase proteins have been shown to induce both local and systemic resistance against necrotrophic pathogens in plants (Kwon *et al.*, 2009). The

ECM contains active components that regulate cell – cell interactions, including secreted proteins that play a role in physiological processes such as growth and defence responses. For instance, the application of two-dimensional (2D) gel electrophoresis and matrix-assisted laser desorption ionization time-of-flight mass spectrometry has been applied to the secretome of cultured *Arabidopsis* cells to identify the secreted proteins involved in plant defence (Oh *et al.*, 2005). Salicylic acid (SA) induced changes in the secretome revealed 13 different proteins species expression were altered by SA treatment.

A secreted lipase with a GDSL motif called GLIP1 (GDSL LIPASE1) has been identified which is essential for non-host resistance in *A. thaliana* to the necrotrophic fungus *Alternaria brassicicola* (Oh *et al.*, 2005). Subcellular localisation experiments indicate that the GLIP1 protein is confined to the extracellular matrix. Inoculation of *glip1* mutants with *A. brassicicola* exhibit spreading lesions due to fungal growth, as compared with small necrotic lesions that confine the infection in the wild type. Given a similar domain architecture of the candidate ACA2 lipase, it is possible it is also a secreted protein found in the extracellular matrix. A subcellular localisation experiment would be needed to explore this possibility further. If this were the case, then it is possible that it could be directly involved in antimicrobial activity and/or required for defence related signalling.

GLIP1-mediated resistance to *A. brassicola* appears to involve salicylic acid (SA) and ethylene induced signalling of systemic acquired resistance (SAR) in surrounding tissues, which confers resistance to biotrophic pathogens such as *Pseudomonas syringae* (Kim *et al.*, 2013). *A. brassicicola* initiates an SAR response that induces defence related genes including *PDF1.2, PR-3, and PR4* in Col-0 (Penninckx *et al.*, 1996, Thomma *et al.*, 1999) Oh et al (2005) Investigated the systemic spread of *A. brassicicola* in *glip1* mutants not observed in Col-0 with SAR through infiltrating the mutants with recombinant GLIP1 protein followed by *A. brassicicola* inoculation and trypan blue staining. These results suggest that GLIP1 activated systemic resistance when challenged by the pathogen. If ACA2-mediated resistance is mechanistically similar, then it is possible that the candidate lipase may inhibit pathogen development via salicylic acid defence signalling and SAR. However, an issue remains that these responses have only

been observed in response to infection with a necrotrophic fungus, leading to the possibility that GLIP1 and candidate ACA2 lipase are mechanistically distinct.

A GDSL lipase has been implicated as a component of non-host resistance to a biotrophic fungus in *A. thaliana.* Comparative genome wide transcription profiling of *A. thaliana* identified the gene UDP-glucosyltransferase 8482 *(BRT1)* that is required for resistance to *Phakopsora pachyrhizi,* the causal agent of Asian soybean rust (Langenbach *et al.*, 2016). Ten genes were found to be transcriptionally co-regulated with *BRT1*, and one of these was a GDSL lipase *(PING7).* Expression as a transgene in soybean was found to significantly enhance resistance to *P. pachyrhizi*, demonstrating how GDSL lipases could potentially to be used to engineer resistance to biotrophic pathogens in major crops.

A GDSL lipase also appears to play a role in the haustorial interaction between *A. thaliana* and another co-evolving biotrophic oomycete *Hyaloperanospora arabidopsidis* (Caillaud *et al.*, 2014). At this interface the pathogen is separated from the host by the extrahaustorial matrix, which contains cell wall material from both the plant and pathogen, and the extra haustorial membrane (EHM). The EHM and plasma membrane are continuous yet differ in protein composition. The plasma membrane that lines the plasmodesmata is proposed to be a specialised domain, and has shown to contain specialised receptors. Plasmodesmata Located Proteins (PDLPs) have been shown to upregulate and localise to the extra-haustorial membrane following penetration by *HpA*. The GDSL lipase At4g28780 has been shown by co-immunoprecipitation to interact with PDLP1, possibly implying a role in defence response (Caillaud *et al.*, 2014)

Effector proteins are pathogen molecules that are translocated into the plant cell and serve as virulence factors to promote pathogenicity. Some effectors elicit race-specific resistance (aka avirulence proteins) in the presence of the corresponding *R* gene. Recent evidence suggests the host's plasma membrane is the active site of effector proteins. For example, *P. syringae* effector proteins including AvrRpm1, AvrPto, AvrRpt2 and AvrPhB have been shown to localise to plasma membranes. Membrane localisation of secreted effectors has been shown to be shown to be important for both virulence and avirulence

function, and has been shown to coregulate with host lipid modification myristoylation and/or palmitoylation. It has been proposed the GDSL-motif lipase At2g04020 is myristoylated in *A. thaliana* (Boisson et al., 2003). If the candidate Aca2 lipase was involved in myristoylation in *B. oleracea,* then the membrane binding properties of cytoplasmic proteins and consequently signal transduction cascades required for defence response would be altered through a non-functional allele.

Perception of pathogen derived effector molecules is fundamental to race-specific resistance. R proteins including Cf2, Cf4, Cf5 and Cf9 from tomato (Dixon *et al.*, 1996, Jones *et al.*, 1994, Munnik, 2001, Van der Hoorn *et al.*, 2001) and *Xa21* from rice (Song *et al.*, 1995) are integral membrane proteins. Arabidopsis R proteins *RPM1* and *RPS2* are also plasma membrane associated (Boyes *et al.*, 1998). Membrane localization of *RPM1* is mediated through lipid modification, and is fundamental and to conferring resistance to *P. syringae*. The additional requirement of *RIN4*, also target to the plasma membrane (Axtell and Staskawicz, 2003) is suggestive that this is the location of *RPM1*-mediated defence related signalling.

Further work is needed for a theoretical understanding of the role of ACA2 in host defence. Yet this would not necessarily be required for application in crop production. The selection of parental lines both homozygous for the EBH527 allele, or the induction of the EBH527 allele into crop types through gene editing would ensure the trait was present in the resulting variety. Yet if it is the case the A12DH allele encodes a protein that is required for compatibility with *A. candida,* this may suggest the EBH527 allele is non-functional. The effect of this allele on other traits of agronomic importance would need to be investigated further prior to any commercial application.

Chapter 3

Mapping of white rust resistance to *Albugo candida* variants that can break broad spectrum *WRR4*-mediated resistance in *Arabidopsis thaliana* Columbia

3.1 INTRODUCTION

Arabidopsis thaliana accessions have been identified that possess varying degrees of resistance to *Albugo candida* race 4, allowing their use as a tool to assess the genetic basis of pathogen host interactions (Borhan *et al.*, 2008, Links *et al.*, 2011). For example, *A. thaliana* accession Wassilewskija (Ws-0) is fully susceptible to this pathogen, which produces profuse asexual sporulation visible as white blisters that form without inducing a host response; whereas accession Columbia (Col-0) has several White Rust Resistance (*WRR*) genes that provide combined layers of induced resistance

Three genes have been molecularly characterised in Col-0 which all encode proteins with an N-terminal Toll/Interleukin-1 Receptor domain, a nucleotide-binding site and a C-terminal leucine-rich repeat domain (TIR-NBS-LRR). *WRR4* (At1g56510) induces a rapid defence that restricts the pathogen from being able to extend beyond the first penetrated cell (Borhan *et al.,* 2008), and effectively masks phenotypes conferred by other resistance genes in Columbia. The hyperstatic defence layers included resistance from a pair of TIR-NBS-LRR genes (*WRR5a,* At5g17880; and *WRR5b,* At5g17890 which contains an additional LIM domain at the C-terminus) that induce a yellowing host response surrounding a patch of hypha without sporulation (Cooper, Cevik & Holub, unpublished). And, a third layer of defence is conferred by *WRR7* (a gene located on the bottom arm of chromosome 5 encoding a small LIM domain protein) that exhibits a flaccid response of the cotyledon and occasional production of small, restricted pustules (Holub & Cevik, unpublished).

A selection of F_9 Col-0 x Ws-0 recombinant inbreds provides a means to compare microscopic differences in phenotype conferred independently by each *WRR* gene (Figure 3.1). For example, inbred CW20 has a Columbia *WRR4* allele and Ws-0 alleles of *WRR5a, 5b* and *7* and exhibits full immunity following inoculation with the *A. candida* race 4 isolate AcEm2; inbred CW5 contains the Columbia alleles of *WRR5a* and *5b* and a Ws-0 allele of *WRR4,* and exhibits a yellow patch phenotype following inoculation; and inbred CW14 has Ws-0 alleles of *WRR4*, *5a* and *5b* and exhibits the flaccid grey phenotype associated with *WRR7* following inoculation. In the susceptible control Ws-0, the pathogen develops extensively across the entire cotyledon within five days after

inoculation, resulting in sporulation without inducing a host response. In CW20 (*WRR4* alone), pathogen development is arrested in a stomatal chamber cell. In CW5 (*WRR5a* and 5b), pathogen development is restricted with no sporulation, and a host response is visible by the retention of trypan blue stain in mesophyll cells surrounding the hyphae. And, in CW14 (*WRR7*) more extensive trypan blue staining is visible, however, the host response is less successful at impairing pathogen development with more extensive sporulation surrounding infection sites.

A. thaliana ecotype Ws-3 has been shown to have varying degrees of susceptibility to *A. candida* race 4 (AcEm2), race 2 (Ac2v), race 7 (Ac7v) and inhibited susceptibility to race 9 (Ac9v), where restricted colonies were able to form without the development of pustules (Borhan et al, 2008). Transformation of Col-0 *WRR4* into Ws-3 induces resistance to the representative isolates of all four races, indicating a broad-spectrum resistance and consequent applicability as a transgene in commercial brassica production (Borhan et al, 2008).

Borhan et al (2008) proposed that *A. candida* contains a highly conserved effector present in at least four races (2, 4, 7 and 9) to explain the apparent broad spectrum resistance conferred by the *WRR4-Col* allele. If so, then Col-virulent pathotypes may arise in natural populations due to mutations in this predicted Avr gene. As described in Chapter 4, three such isolates of *A. candida* isolates (AcExeter, AcCarlisle and Ac167) have been identified as used to identify candidate *avrWRR4-Col* genes. The aim of this chapter is to use two of these Col-virulent isolates (AcExeter and AcCarlisle) to identify and map a new and potentially broader spectrum source of white rust resistance in *A. thaliana*.

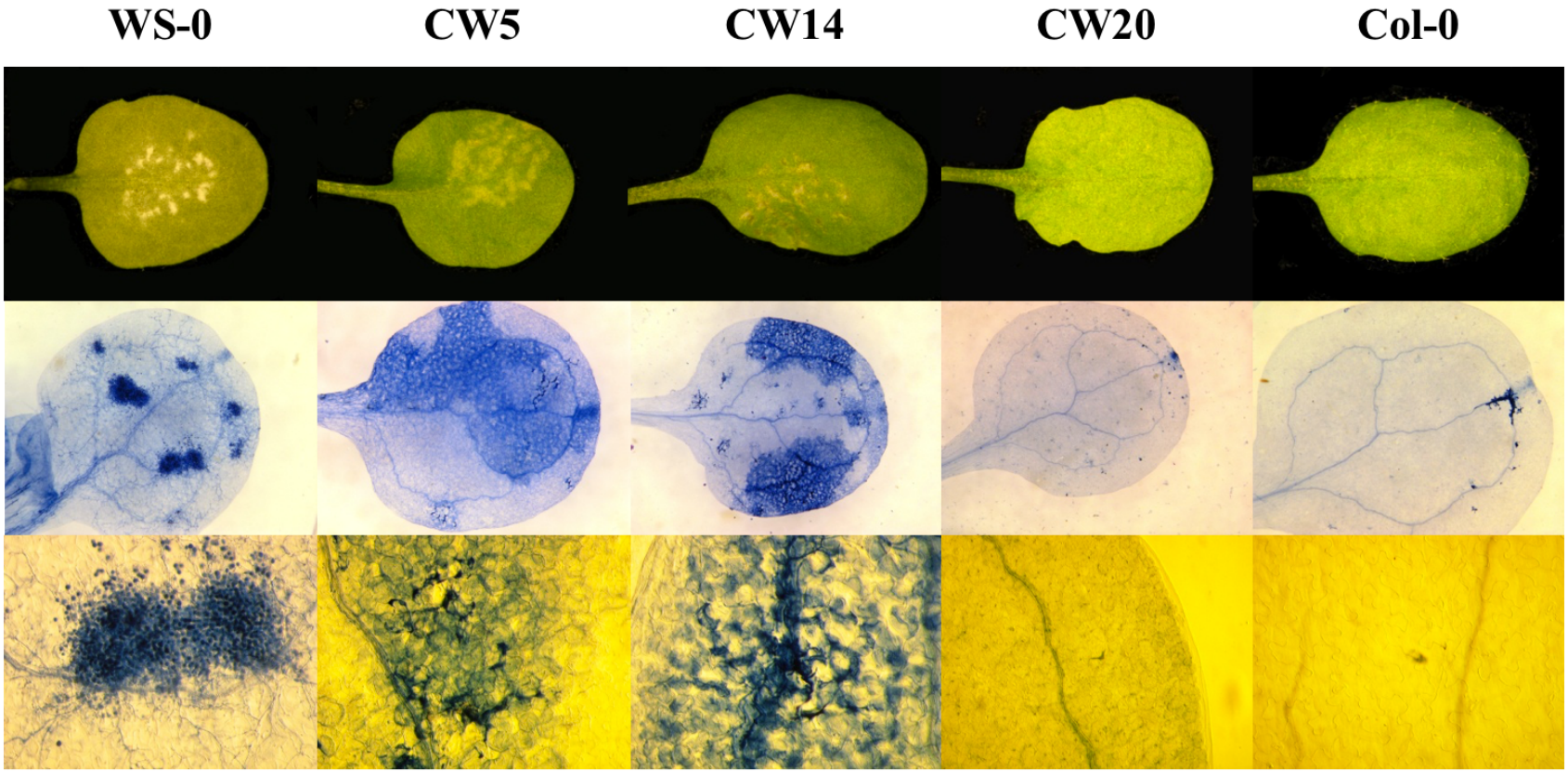


Figure 3.1. Cotyledons of *Arabidopsis thaliana* accessions Col-0 (resistant), Ws-0 (susceptible) and F_9 recombinant inbred lines identified as each having a single white rust resistance (WRR) specificity that induces differing degrees of resistance to *Albugo candida* race 4. Cotyledons were photographed seven days post inoculation with isolate AcEm2, and stained with trypan blue. Upper images are 10× magnification and lower images are 40× magnification. Inbred CW5 has the Col-0 alleles of *WRR5a* and *5b* (At5g17880 and At5g17890) from the top arm of chromosome 5; inbred CW14 has the Col-0 allele of *WRR7* from the bottom arm of chromosome 5, and inbred CW20 has the Col-0 allele of *WRR4* (At1g56510) from the bottom arm of chromosome 1. (Figure prepared from photographs taken during MSc research project, Fairhead, 2012).

3.2 MATERIALS AND METHODS

3.2.1 Maintenance of *Albugo candida* isolates

The Col-virulent *A. candida* isolates AcExeter and AcCarlisle were bulked and maintained on juvenile leaves of *A. thaliana* Col-0, and the Col-avirulent isolate AcEm2 was bulked and maintained on juvenile leaves of *A. thaliana* Ws-*eds1*. Seeds were grown on an Arabidopsis compost mix (6:2:2 Levington F2, sand and vermiculite) in P180 plug trays (2.5cm x 2.5cm cells) cut to 10 x 6 and placed in propagator trays. Two to five seeds were sown per cell prior to being reduced to single plant following germination. The trays were covered with aluminium foil and placed in a fridge at 4°C for 48h to promote even germination. The propagators were covered with transparent lids and placed in a Conviron plant growth cabinet at 20±2°C with a 10h photoperiod.

All three isolates were revived from leaf tissue that was stored at -80 C and containing mature pustules. The leaves were submerged and agitated in sterile water. A haemocytometer was used to estimate the inoculum concentration in order to adjust it to an approximate concentration of 4x 10^4 zoosporangia per ml. Small droplets (ca. 10µl) were applied to the upper and lower surfaces of cotyledon and juvenile leaf tissue using a repetitive pipette. Propagators were sealed and placed back in the growth cabinet at 20±2 C for an initial 12h period in darkness followed by a 10h photoperiod. Pustules would emerge 7-10 days later. The isolates were sub-cultured every two weeks for use on experimental lines

3.2.2 Characterising phenotypic variation of *Arabidopsis thaliana* germplasm in response to *Albugo candida*

A diversity collection consisting of 19 accessions (Bur-0, Can-0, Col-0, Ct-1, Edi-0, Hi-0, Kn-0, Ler-0, Mt-0, No-0, Oy-0, Po-0, Rsch4-4, Sf-2, Tsu-0, Wil-2, Ws-0, Wu-0 and Zu-0) was used to identify examples that are resistant to each of the Col-virulent isolates of *A. candida.* These accessions had previously been used to generate a Multiparent Advanced Generation Inter-Cross (MAGIC) inbred mapping population (Kover *et al.*, 2009), and have been used to generate a

database of high resolution genotypes with more than single nucleotide polymorphisms (SNPs) relative to the reference Col-0 genome (Gan *et al.*, 2011). As described below, genetic mapping of resistance was conducted using phenotype scores for 450 MAGIC inbred lines, and 470 F_8 recombinant inbreds from a bi-parental cross of Oy-0 (resistant) x Col-0 (susceptible) obtained from INRA (Simon *et al.*, 2008).

A ten-class disease severity scale (Figure 3.2) was used to determine the phenotypic interaction of the wild accessions and experimental lines of *A. thaliana* following inoculation of cotyledons with *A. candida,* specifically with isolates that can overcome white rust resistance in the accession Columbia. The phenotype classes represent a visual interpretation of the ability of the pathogen to develop within the host. A '0' phenotype indicates full resistance, with no visible presence of the pathogen or host response. A class 1 phenotype shows small necrotic lesions at the immediate site of infection with no blisters being apparent, also indicating rapid host response that completely restricts the infection. Phenotypes classes 2-4 exhibit increasing degrees of necrosis at the site of infection with no apparent presence of the pathogen. A class 5 phenotype shows yellowing patches that develop on the upper and lower leaf surface, typically with no sporulation in the majority of cases, yet minute blisters can occasionally be detected indicating that hyphae have developed within the leaf. Phenotypes 6-7 exhibit a grey flaccid response in the cotyledon, with development of minute pustules can be detected on the upper and lower leaf surface. And, phenotype classes 8 and 9 are fully susceptible with unrestricted sporulation on the lower leaf surface, with no visible host response in class 8 but flaccidity of host tissue in class 9.

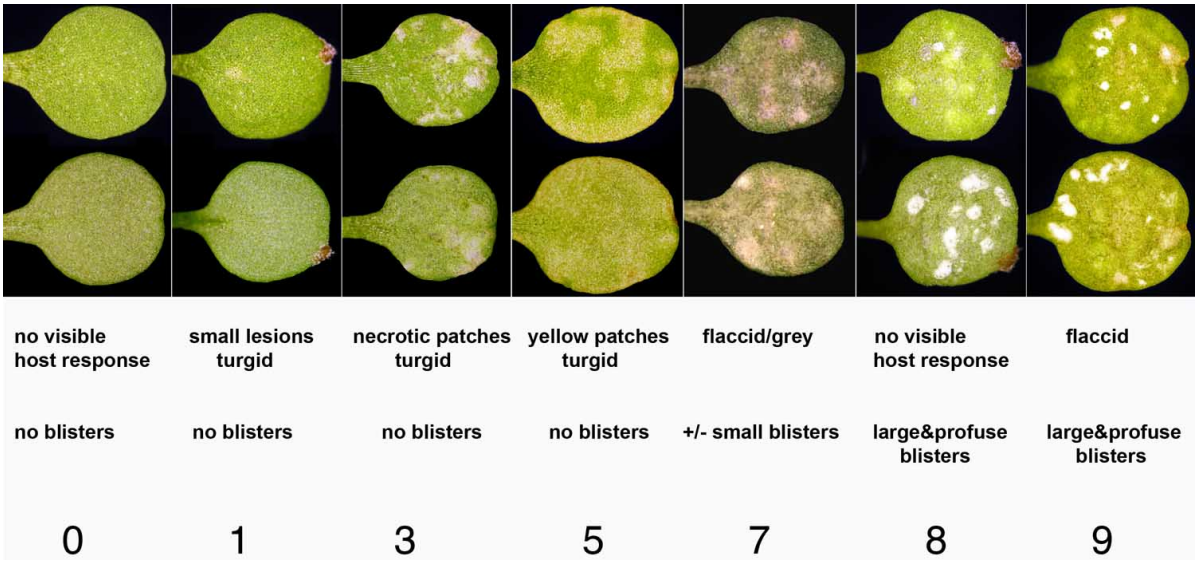


Figure 3.2. A ten-class scale of interaction phenotypes in cotyledons of *Arabidopsis thaliana* following inoculation with *Albugo candida* race 4. The degree of host response (top row of descriptions) and pathogen reproduction (bottom row of descriptions) observed on upper and lower surfaces was recorded ten days post inoculation.

The diversity collection and mapping populations of *A. thaliana* were phenotyped for response to *A. candida* using the same method applied for the maintenance of the pathogen isolates. Ten-day-old cotyledons were inoculated with a single 10µl drop of *A. candida* inoculum to the upper surface of the leaf at an approximate concentration of $4x10^4$ zoosporangia per ml. Propagators containing the experimental lines were sealed and placed back in the growth cabinet at 20°C for an initial 12 h period of shading followed by a 10 h photoperiod. Three replicates of each experimental line were tested with following inoculations with AcExeter and AcCarlisle. Symptoms were assessed ten days post inoculation using the phenotype scale described above, and average phenotypes across replicates were calculated for further analysis.

3.2.3 QTL analysis

QTL mapping in the MAGIC population was conducted using a multipoint method through a HAPPY software package in R (Kover *et al.*, 2009). QTL mapping of the bipartite F_8 Oy-0 x Col-0 population was performed on a genetic map of 85 markers across the five linkage groups that accompanied the population (Simon *et al.*, 2008) using RQTL (Broman *et al.*, 2003). Three statistical approaches were applied to interval mapping including: 1) standard interval mapping using a maximum likelihood estimation under a mixture model (Lander and Botstein, 1989); 2) Haley-Knott regression using approximations of the mixture model (Haley and Knott, 1992); and 3) a multiple imputation method using the mixture model with multiple imputations as opposed to maximum likelihood estimation (Sen and Churchill, 2001).

3.3 RESULTS

3.3.1 Phenotypic characterisation of *A. thaliana* diversity collection

Phenotypic variation was observed amongst the 19 MAGIC parents following separate inoculation with AcExeter and AcCarlisle and the Columbia-avirulent control isolate AcEM2 (Table 3.1). Importantly, three accessions (Mt-0, Oy-0 and Rsch-4) were resistant to all three isolates. Four accessions (Bur-0, Ct-1, Edi-0 and Ler-0) exhibited the same pattern of phenotypes as Col-0, indicating the presence of a functional *WRR4* allele. Three accessions (Can-0, Ws-0 and Kn-0) were broadly susceptible to all three isolates. An interaction phenotype was not obtained for eight accessions, however, six of these (Hi-0, No-0, Sf-2, Wil-2, Wu-0 and Zu-0) were susceptible to both of the Col-virulent isolates.

Table 3.1. Summary of interaction phenotypes in cotyledons of 19 *Arabidopsis thaliana* accessions recorded at ten days after inoculation with three isolates of *Albugo candida,* race 4: AcEm2, Ac Exeter and AcCarlisle. The accessions are the parents of a Multiparent Advanced Generation Inter-Cross (MAGIC) recombinant inbred population (Kover et al, 2009).

	A. candida isolate		
A. thaliana accession	AcEm2	AcExeter	AcCarlisle
Can-0, Ws-0, Kn-0	Susceptible	Susceptible	Susceptible
Bur-0, Col-0, Ct-1, Edi-0, Ler-0	Resistant	Susceptible	Susceptible
Mt-0, Oy-0, Rsch-4	Resistant	Resistant	Resistant
Hi-0, No-0, Sf-2, Wil-2, Wu-0, Zu-0	NA	Susceptible	Susceptible
Po-0, Tsu-0	NA	NA	NA

3.3.2 Mapping of resistance to Columbia-virulent *Albugo candida* isolates using a MAGIC inbred population

The 450 MAGIC recombinant inbreds were used for mapping resistance to both AcExeter and AcCarlisle. A major effect allele was identified on the bottom arm of chromosome 1 between markers PERL0132168 and NMSNP1_24 at physical positions 16134927 and 24734584, respectively, (P≤0.001) (Figure 3). Interestingly, this interval spans the region that contains *WRR4.* The haplotype of Can-0, Mt-0, Oy-0 and Rsch-4 at the centre of the QTL are predicted to confer resistance to AcExeter and AcCarlisle, indicating they may share an allele of the same broad-spectrum *R* gene.

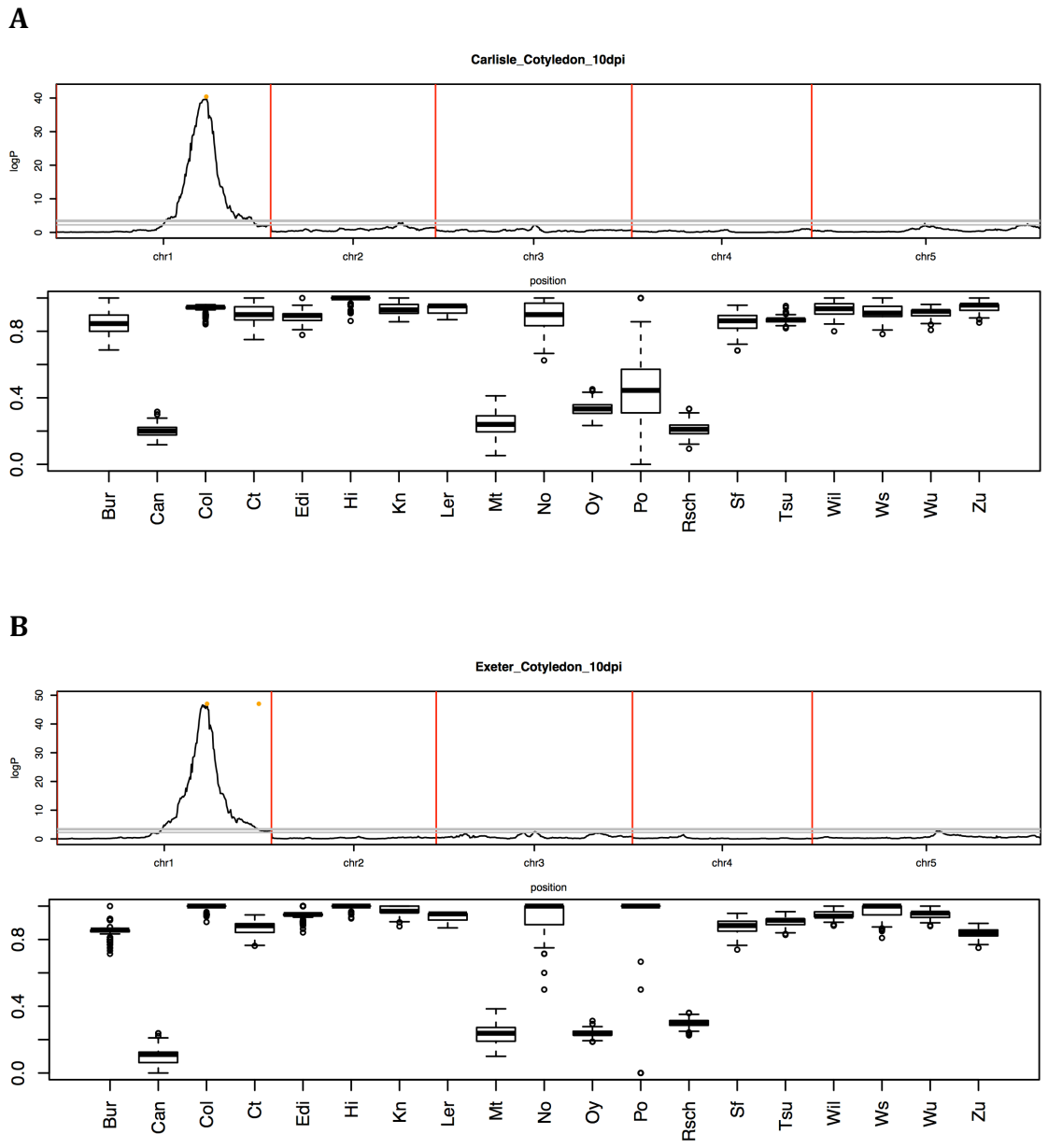


Figure 3.3. QTL analysis of white rust resistance in 470 Multiparent Advanced Generation Inter-Cross (MAGIC) inbred lines of *Arabidopsis thaliana* recorded ten days after inoculation with *Albugo candida* isolates AcCarlisle (**A**) or AcExeter (**B**). Whisker plots represent the predicted contribution of each genotype to the corresponding phenotype, ranging from 0.0 (resistant) to 1.0 (susceptible).

3.3.3 QTL interval mapping of resistance in *Arabidopsis thaliana* accession Oy-0 to AcExeter

The Norwegian accession Oy-0 was chosen as an example of resistance to AcExeter for fine-mapping using an F_9 Col-0 x Oy-0 recombinant inbred population. For this QTL analysis, conditional genotype probabilities were calculated on a grid with a density of 1 cM assuming a genotyping error probability 0.001. Standard interval mapping was performed using 1) a maximum likelihood estimation under a mixture model, 2) a Haley-Knott regression that was performed using approximations of the mixture model, and 3) a multiple imputation method that was applied instead of the maximum likelihood estimation.

For the maximum likelihood estimation and Haley-Knott regression (Figure 4.5A and 4.5B, respectively), 1000 permutations of the genotype probabilities were conducted to calculate genome wide LOD significance threshold. These were 2.5 and 2.4 for a 5% confidence interval, respectively. For the multiple imputation method (Figure 4.6C), 1000 imputations were performed to calculate the genome wide LOD significance threshold as 2.3 for a 5% confidence interval. All three methods detected three QTLs above the corresponding threshold. A major effect QTL on chromosome 1 changed slightly in the position and LOD score according to method. The maximum likelihood estimation gave a LOD of 61.95 at position 73.0 cM, the Haley-Knott regression gave a LOD of 55.5 at 73.0 cM, and the multiple imputation scan gave a LOD of 53.72 at 74.0cM. A second minor QTL was mapped on chromosome 3 to position 67 cM (LOD of 8.0) with the maximum likelihood estimation and Haley-Knott regression; and to position 68 cM (LOD 7.6) with the multiple imputation scan. And, a third minor QTL was detected on chromosome 5 at position 16.7 cM (LOD of 3.89) with all three methods.

A multiple QTL scan was conducted on each model controlling the primary locus on chromosome 1 to see whether any minor effects following inoculation with AcExeter would become more significant (Figures 4.5 and 4.6). The genotype probabilities were calculated with an error probability of 0.01 and a step size of 1 cM for maximum likelihood and Haley-Knott regression, and 100 imputations were performed for the multiple imputation method. The genotype

for the marker closest to the peak locus for each output at identified the physical position C1_4567846, and was subsequently removed (Figure 4.5A,B, and C) or used as an additive covariant in the single QTL scans (Figure 4.6A,B and C). The LOD score of the two minor QTLs increased in all cases. For the maximum likelihood genome scan, the chromosome 3 locus had a LOD of 12.4, and the chromosome 5 locus had a LOD of 5.64. For the Haley-knot regression the chromosome 3 locus had a LOD of 12.6 and the chromosome 5 locus had a LOD of 5.67. For the multiple imputation genome scan the chromosome 3 locus had a LOD of 13.4 and the chromosome 5 locus had a LOD of 5.42. No additional QTLs were detected with any method.

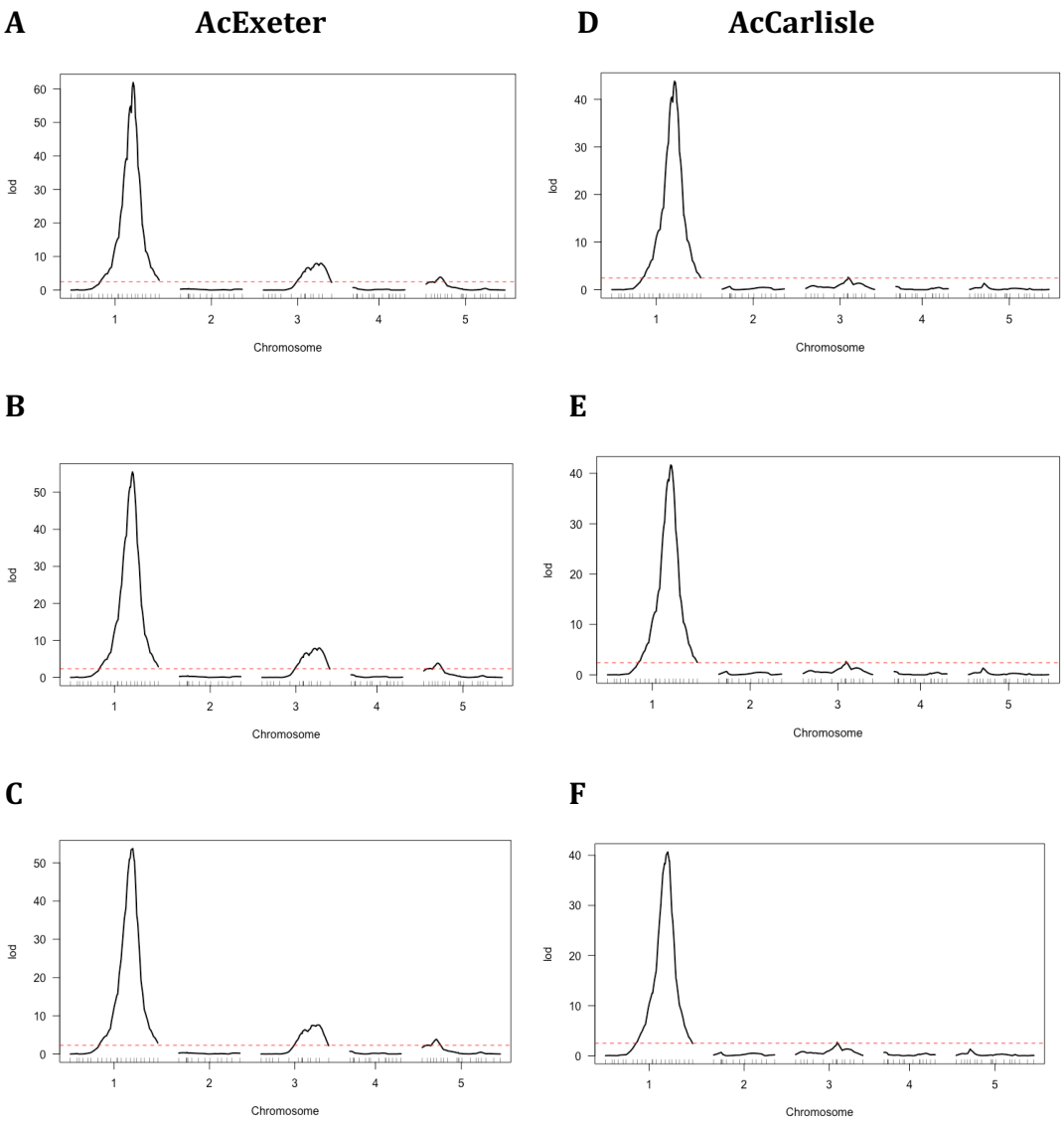


Figure 3.4. Mapping of resistance to *Albugo candida* isolate AcExeter (**A**,**B**,**C**) and AcCarlisle (**D**,**E**,**F**) in *Arabidopsis thaliana* using an F_9 Col-0 (susceptible) x Oy-0 (resistant) mapping population. **A** and **D**, show genome scans using the maximum likelihood algorithm (logarithm of odds (LOD) of 2.5 for both isolates at a 5% confidence interval); **B** and **F,** show scans using Haley-Knott regression (LOD of 2.4 for both isolates at a 5% confidence interval); and **C** and **F,** show scans using multiple imputations (LOD of 2.3 and 2.5 for each isolate, respectively at a 5% confidence interval). Phenotypes were recorded ten days post inoculation of cotyledons and scored using a ten class phenotype scale of resistance (see Figure 1).

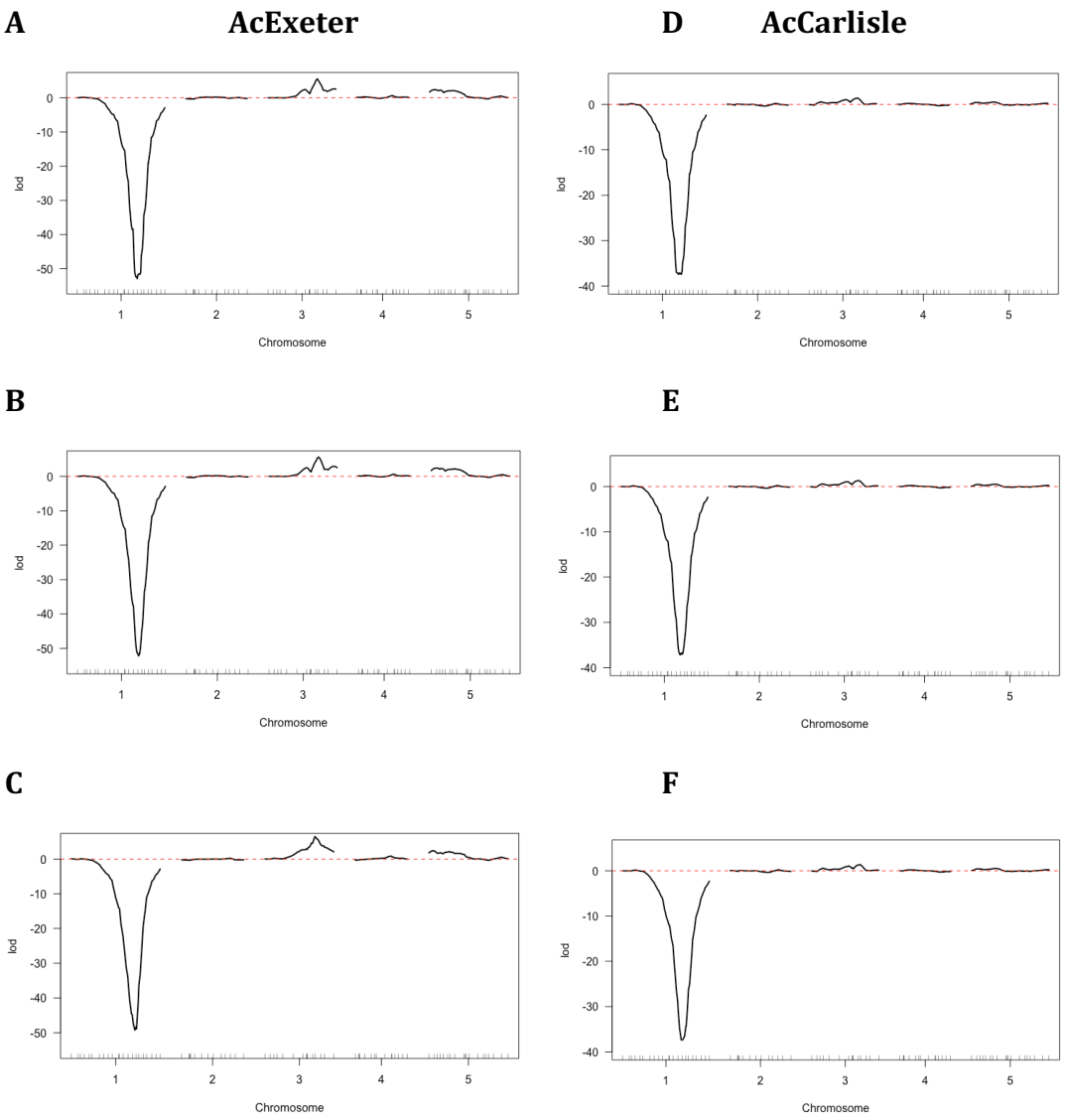


Figure 3.5. Composite interval mapping of resistance to *Albugo candida* isolate AcExeter (**A**,**B**,**C**) and AcCarlisle (**D**,**E**,**F**) in *Arabidopsis thaliana* using an F_9 Col-0 (susceptible) x Oy-0 (resistant) mapping population, with the effect of subtracting the marker (C1_20384) that is most tightly linked to the major effect locus on chromosome 1. **A** and **D**, show genome scans using the maximum likelihood algorithm; **B** and **E,** show scans using Haley-Knott regression; and **C** and **F,** show scans using multiple imputations. Phenotypes were recorded ten days post inoculation of cotyledons and scored using a five class phenotype scale of resistance and differing phenotypes of susceptibility.

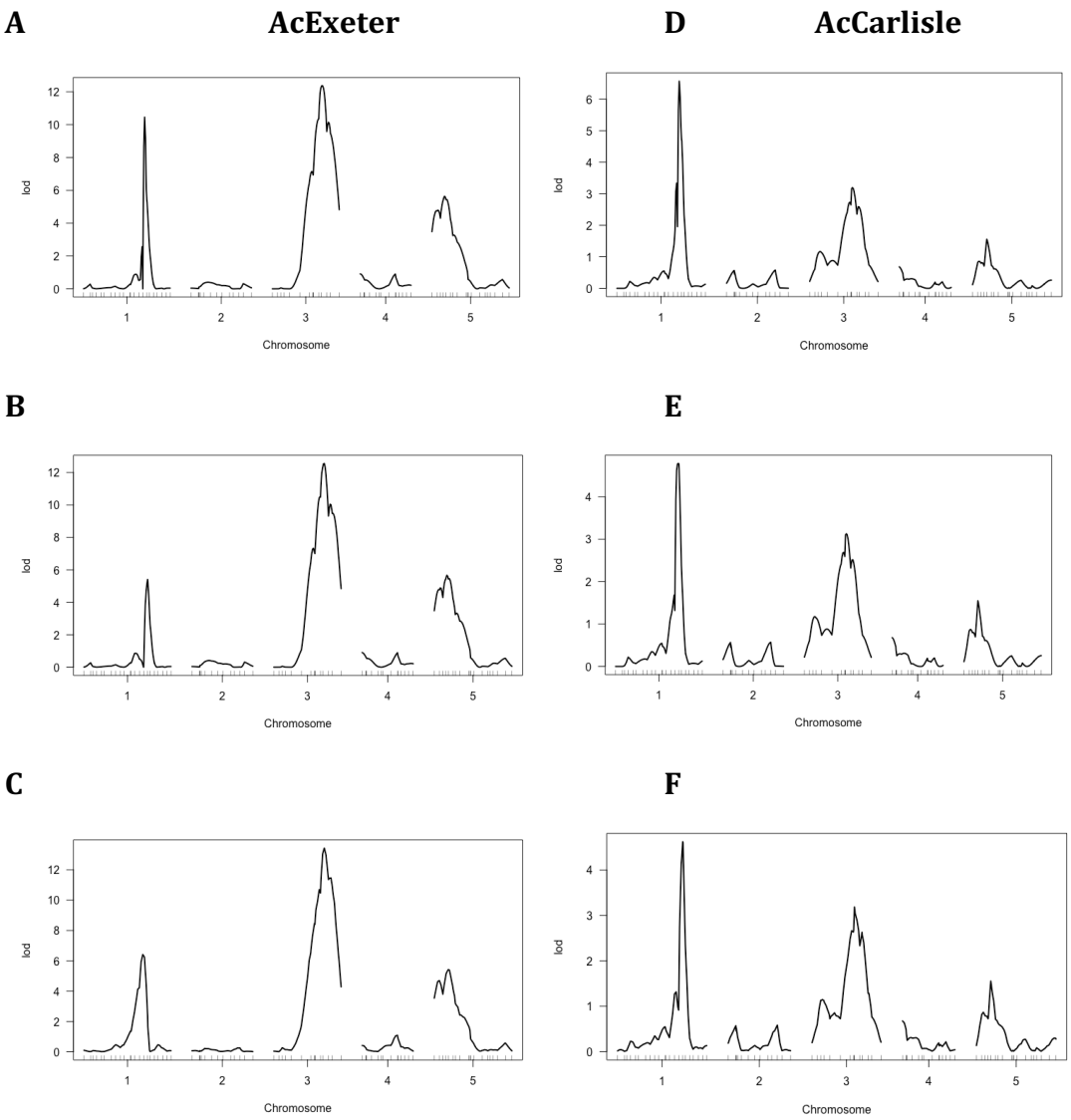


Figure 3.6. Composite interval mapping of resistance to *Albugo candida* isolates AcExeter and AcCarlisle in *Arabidopsis thaliana* using an F_9 Col-0 (susceptible) x Oy-0 (resistant) mapping population, by using the marker most tightly linked to the major effect on chromosome 1 (C1_20384) as an additive covariate. **A** and **D**, show genome scans using the maximum likelihood algorithm; **B** and **E,** show scans using Haley-Knott regression; and **C** and **F,** show scans using multiple imputations. Phenotypes were recorded ten days post inoculation of cotyledons and scored using a five class phenotype scale of resistance and differing phenotypes of susceptibility.

A two-dimensional genome scan was performed to begin searching for multiple linked or interacting QTLs underlying resistance to *A. candida* isolate AcExeter (Table 3.2). When comparing the full two-QTL model with the best single fit QTL model, there is evidence of the major effect QTL on chromosome 1, and the secondary QTLs on chromosomes 3 and 5 with allowance for epistasis (Figure 3.7). The significance thresholds for each of these interactions is greater than when comparing the full additive QTL model without allowance for epistasis with the best single QTL model, indicating that they are interactive. However, there is evidence of a second QTL on chromosome 1 when comparing the full additive QTL model without allowance for epistasis with the best single QTL model. The effects of putative linked loci on chromosome 1, and unlinked loci on chromosomes 3 and 5 were assessed through analysing phenotype as a function of the genotype of the most tightly linked markers, and by assessing the phenotype averages of each of the four two locus genotype groups for each pair of contributing effects.

Table 3.2. Predicted second gene effects for resistance to *Albugo candida* isolate AcExeter in *Arabidopsis thaliana* Col-0 x Oy-0 F_9 mapping population calculated using a multiple imputation genome scan. Significance thresholds for the LOD of the full model *(Sf),* LOD of the additive model *(Sa),* test for epistasis *(Si),* the comparison of the full model to the best single-QTL model *(Sfv1)*, and the comparison of the additive model to the best single-QTL model *(Sav1)* were calculated using 100 permutations.

	Position 1 full	Position 2 full	*Sf*	*Sfv1*	*Si*	Position 1 additive	Position 2 additive	*Sa*	*Sav1*
Significance threshold	-	-	5.10	3.66	3.27	-	-	3.47	1.18
C1:C1	68	74	56.3	2.60	0.565	70	76	55.8	2.03
C1:C3	74	58	70.2	16.49	1.775	74	58	68.4	14.71
C1:C5	72	16	58.9	5.22	0.182	72	16	58.8	5.04
C3:C5	60	16	12.1	4.46	0.333	60	16	11.8	4.13

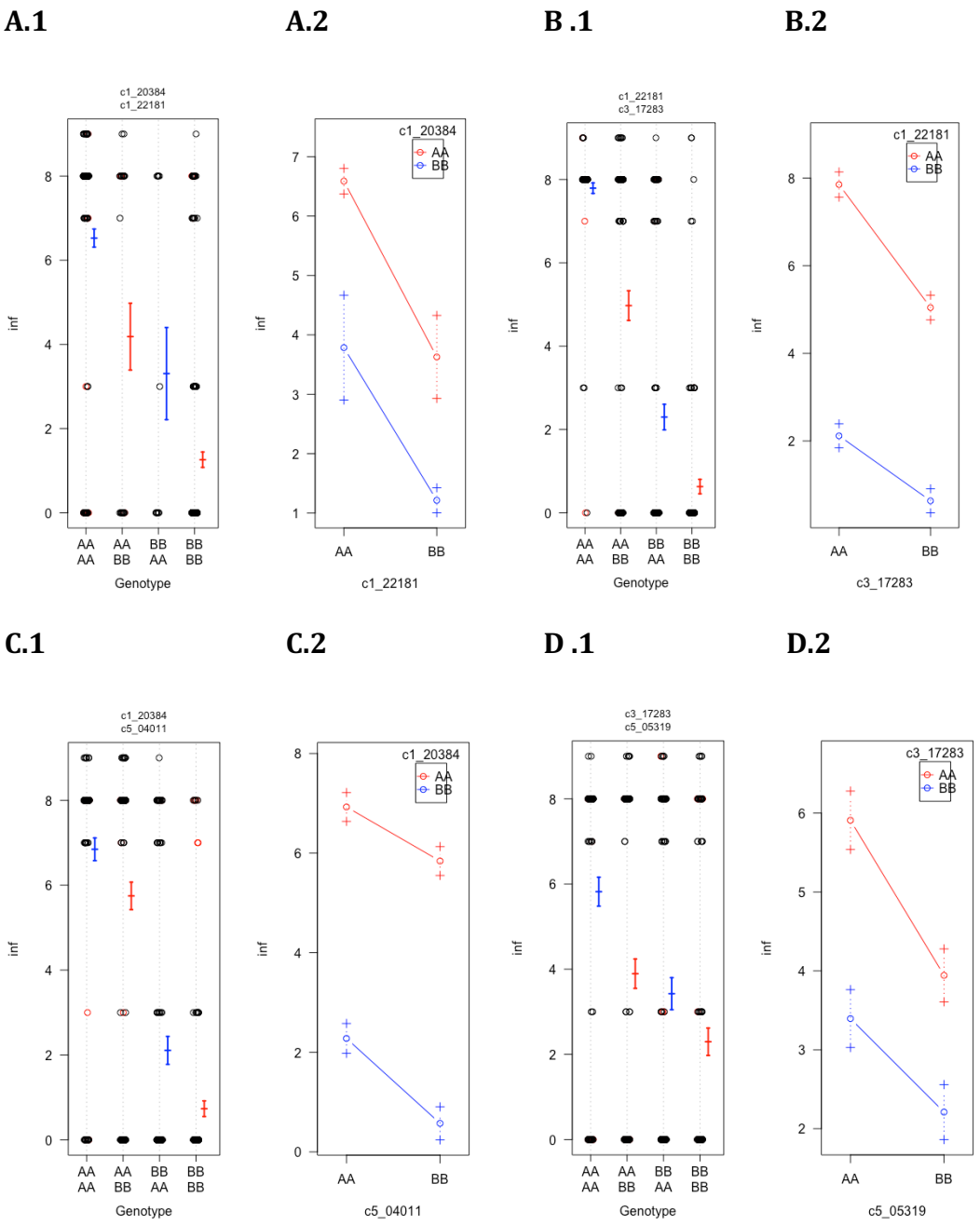


Figure 3.7. The effect of putative QTLs identified as contributing to white rust resistant phenotype in *Arabidopsis thaliana* Col-0 (AA) x Oy-0 (BB) mapping population following cotyledon inoculation with *Albugo candida* isolate AcExeter. **.1,** dot plots of phenotype as a function marker genotypes identified as interactive. Black dots correspond to observed genotypes and red dots correspond to missing, consequently imputed genotypes. **.2,** estimated phenotype averages for each two locus genotype group. **A**, putative QTLs chromosome 1 at physical positions 20384 and 22181 Mb; **B,** interactive QTLs on chromosomes 1 and 3 at positions 22181 and 17283 Mb, respectively; **C,** interactive QTLs on chromosomes 1 and 5 at positions 20384 and 04011 Mb, respectively; and **D,** interactive QTLs on chromosomes 3 and 5 at positions 17283 and 05319 Mb, respectively.

The results for the two putative QTLs on chromosome 1 show a mean susceptible phenotype when both loci are homozygous for Col-0, and a mean resistant phenotype when both loci are homozygous for Oy-0. A greater mean phenotype with lower variance was observed when marker C1_20384 is homozygous for Col-0 and C_22181 is homozygous for Oy-0, than when C1_20384 is homozygous for Oy-0 and C_22181 is homozygous for Col-0. This indicates a stronger contribution to resistance at C1_20384, yet Oy-0 alleles are required at both loci for a resistant phenotype.

The results for the interactive effect between the chromosomes 1 and 3 loci reveal greatest mean susceptible phenotype with the lowest degree of variance when both loci are homozygous for Col-0, and the greatest mean resistant phenotype with the lowest degree of variance when both loci are homozygous for Oy-0. This indicates that an Oy-0 allele at both loci is required for a fully resistant phenotype. In addition, the mean phenotype when the chromosome 1 locus is homozygous for Col-0 allele and the chromosome 3 locus is homozygous for the Oy-0 allele is significantly greater than when the chromosome 1 locus is homozygous for Oy-0 allele and the chromosome 3 locus is homozygous for the Col-0 allele. This indicates that the chromosome 1 locus is the greatest contributor to the resistant phenotype.

For the interaction between the chromosome 1 and chromosome 5 locus, the greatest mean susceptible phenotype is observed when both loci are homozygous for Col-0, and the lowest mean resistant phenotype is observed when both loci are homozygous for the Oy-0 allele. Again there is a significantly higher mean phenotype when the chromosome 5 locus is homozygous for the Oy-0 allele and the chromosome 1 locus is homozygous for the Col-0 allele than when the chromosome 5 locus is homozygous for the Col-0 allele and the chromosome 1 locus is homozygous for the Oy-0 allele. This indicates that the predominant contribution to resistance is from the chromosome 1 locus.

In the interaction between the loci on chromosome 3 and chromosome 5, all groups possess an intermediate mean phenotype, likely a consequence of the absence of an Oy-0 allele at the C1 locus. Nevertheless, the greatest mean susceptible phenotype is observed when both loci are homozygous for Col-0, and the lowest mean resistant phenotype is observed when both loci are

homozygous for the Oy-0 allele.

A multiple imputation method was used to compensate for missing genotype data, and therefore improve the predictive model for locations of the putative QTLs. By pulling out the imputed genotypes at each location, a four QTL model was generated allowing for interactions between the QTLs on chromosomes 1 and 3, 1 and 5, and 3 and 5. The overall fit of the model provided a LOD score of 79.04 relative to null model, with 54.44% of the phenotypic variance being accounted for. Each locus was dropped and reintroduced in succession, allowing comparison to be made between the full model and the model with the term omitted. The results provided strong evidence for a role of both loci on chromosome 1, as well as the additional loci on chromosomes 3 and 5. Evidence for interactions between the primary focus on chromosome 1 and the chromosome 3 locus was apparent. The interactions between chromosomes 1 and 5, and 3 and 5 fell below the threshold for significance (Table 3.3).

Table 3.3. Analysis of the effect of dropping each QTL independently from a five QTL model generated by identifying fixed locations of QTLs contributing to resistance to *Albugo candida* isolate AcExeter in *Arabidopsis thaliana* F_9 Col-0 x Oy-0 mapping population.

Chr position (cM)	Type 3 sum of squares	LOD	% variance	F value	P value
C1 70	206.88	6.25	2.92	29.17	0.00
C1 76	279.64	8.36	3.95	13.14	0.00
C3 58	624.63	17.79	8.82	29.36	0.00
C5 16	259.33	7.77	3.66	12.19	0.00
C1 76 & C3 58	49.78	1.54	0.70	7.02	0.01
C1 76 & C5 16	8.28	0.26	0.12	1.17	0.28
C3 58 & C5 16	16.92	0.53	0.24	2.34	0.12

3.3.4 QTL interval mapping of resistance in *Arabidopsis thaliana* accession Oy-0 to AcCarlisle

The same mapping analyses as described above were used to identify QTLs and possible interactions between loci following inoculation of F_9 Col-0 x Oy-0 inbreds with a second Columbia-virulent *A. candida* isolate AcCarlisle. In this case, the genome-wide LOD significance threshold was 2.5 with the maximum likelihood estimation and multiple imputation method and 2.4 with Haley-Knott regression for a 5% confidence interval. All three methods detected QTLs above the corresponding threshold on chromosomes 1 and 3, but not on chromosome 5 as detected with AcExeter (Figure 3.4D,E and F). The predicted position of a major effect QTL on chromosome 1 is 73.0 cM with maximum likelihood estimation and Haley-Knott regression (LOD scores of 43.80 and 41.67, respectively), and 74.0 cM with the multiple imputation scan (LOD of 40.67). The predicted position of a minor effect QTL on chromosome 3 is 49.2cM with all three methods (LOD score of 2.64) with maximum likelihood estimation and Haley-Knott regression; and 2.69 with multiple imputation scan. Interestingly, the position of the second minor effect QTL was approximately 16 cM above the one predicted on the same chromosome with AcExeter.

A multiple QTL scan was conducted on each model controlling the major effect locus on chromosome 1 to make any minor effects more apparent (Figures 5 and 6). No further significant QTLs could be detected. The same position of the minor effect QTL is predicted on chromosome 3 of 50 cM for the maximum likelihood genome scan and Haley-knot regression (LOD of 3.19 and 3.13, respectively) and 49.2 for the multiple imputation genome scan (LOD of 3.18).

A two dimensional genome scan was performed to begin searching for multiple linked or interacting QTLs underlying resistance to *A. candida* AcCarlisle (Table 3.4). When comparing the full two-QTL model with the best single fit QTL model, there is evidence of the major effect QTL on chromosome 1, and a secondary QTL on chromosomes 3 and 5 with allowance for epistasis. The effects of both loci were assessed through analyzing phenotype as a function of the genotype of the most tightly, and by assessing the phenotype averages of each of the four two locus genotype groups for each pair of contributing effects. (Figure 3.8).

The results reveal the greatest mean susceptible phenotype when both alleles are homozygous for Col-0, and the lowest mean resistant phenotype when both alleles are homozygous for Oy-0. There is a greater mean phenotype when the chromosome 1 locus is homozygous for the Col-0 allele and the chromosome 3 locus is homozygous for the Oy-0 allele than when the chromosome 1 locus is homozygous for the Oy-0 allele and the chromosome 3 locus is homozygous for the Col-0 allele. This indicates the chromosome locus is a greater contributor to a resistant phenotype.

3.3.5 Oy-0 x Col-0 RIL phenotyping with AcEm2 suggests shared resistance locus

The experimental procedure as described above was performed using AcEm2 on the Oy-0 x Col-0 RIL population. None of the 470 lines developed pustules, with all exhibiting a phenotype of 0 or 1. Single QTL genome scans using the three statistical approaches described and two dimensional two QTL genome scans were performed using the phenotype data. No significant QTLs were detected as expected.

Table 3.4. Predicted second gene effects for resistance to *Albugo candida* isolate AcCarlisle in *Arabidopsis thaliana* Col-0 x Oy-0 F_9 mapping population calculated using a multiple imputation genome scan. Significance thresholds for the LOD of the full model *(Sf),* LOD of the additive model *(Sa),* test for epistasis *(Si),* the comparison of the full model to the best single-QTL model *(Sfv1)*, and the comparison of the additive model to the best single-QTL model *(Sav1)* were calculated using 100 permutations.

	Position 1 full	Position 2 full	*Sf*	*Sfv1*	*Si*	Position 1 additive	Position 2 additive	*Sa*	*Sav1*
Significance threshold	-	-	4.45	3.58	3.08	-	-	3.52	1.85
C1:C3	74	8	44.4	3.7	0.887	74	50	43.5	2.82

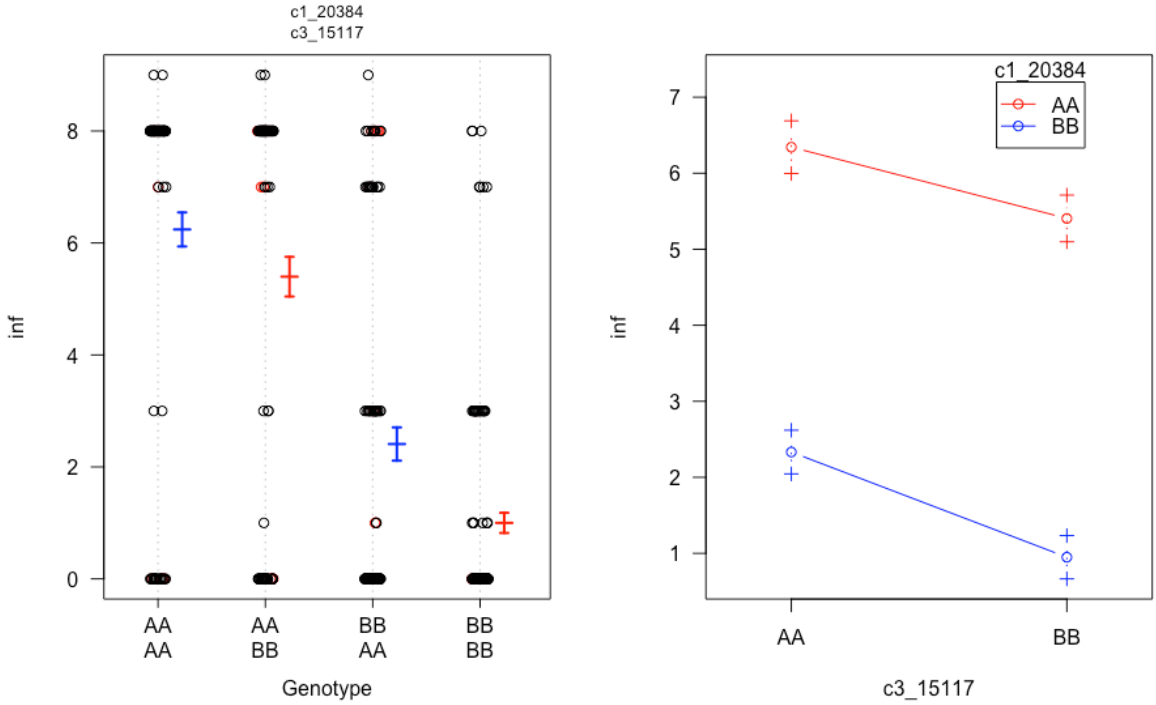


Figure 3.8. The effect of putative QTLs on chromosomes 1 and 3 identified as contributing to white rust resistant phenotype in *Arabidopsis thaliana* Col-0 (AA) x Oy-0 (BB) mapping population following cotyledon inoculation with *Albugo candida* isolate AcCarlisle. **Left,** dot plots of phenotype as a function marker genotypes identified as interactive. Black dots correspond to observed genotypes and red dots correspond to missing, consequently imputed genotypes. **Right,** estimated phenotype averages for each two locus genotype group.

The predicted locations of the putative QTLs were defined as a model in order estimate the effects of each locus. For this purpose, the multiple imputation method was used to compensate for missing genotype data. By pulling out the imputed genotypes at each location, a two QTL model was generated suggesting an additive effect between the QTL's on chromosomes 1 and 3. The overall fit of the model provided a LOD score of 41.18 relative to null model, with 33.9% of the phenotypic variance being accounted for. Each locus was dropped and reintroduced in succession, allowing comparison to be made between the full model and the model with the term omitted (Table 3.5). The results provided strong evidence for both loci.

Table 3.5. Analysis of the effect of dropping each QTL independently from a five QTL model generated by identifying fixed locations of QTLs contributing to resistance to *Albugo candida* isolate AcCarlisle in *Arabidopsis thaliana* F9 Col-0 x Oy-0 mapping population.

Position	Type 3 sum of squares	LOD	% variance	F value	P value (F stastic)
C1 70cM	2015.8	38.67	31.405	216.21	0.00
C3 76cM	112.9	2.84	1.914	13.18	0.00

3.3.5 Identification of candidate disease resistance-like genes located within QTLs that confer resistance to *A. candida* isolates AcExeter and AcCarlisle

The physical locations of markers flanking the putative QTLs on chromosomes 1, 3 and 5 were related to the TAIR10 reference genome and used to identify candidate NBS-LRR genes within each interval (Table 3.6). The major effect QTL on chromosome 1 locus from C1_20308 to C1_22181 spans 500 predicted genes that include 16 which are predicted to encode NBS-LRR proteins. The minor effect QTL on chromosome 3 from C3_15117 to C3_23412 spans 2219 predicted genes that include 11 predicted NBS-LRR genes. The minor effect QTL on chromosome 5 locus from C5_02900 to C5_06820 spans 1152 predicted genes that include seven predicted NBS-LRR genes. Given the broad mapping intervals,

these three loci will be referred to below by temporary names *WRR-OyC1, WRR-OyC3* and *WRR-OyC5,* respectively.

Table 3.6. Disease resistance-like genes located in putative QTLs that were predicted to contribute to segregation of resistance in *Arabidopsis thaliana* Col-0 x Oy-0 F_9 mapping population to *Albugo candida* isolates AcExeter and AcCarlisle.

Locus	Gene ID	Protein Description	Source
WRR-OyC1	At1g56470	Pseudogene, disease resistance-like protein	TAIR
	At1g56510	WRR4A white rust resistance protein (TNL)	TAIR
	At1g56520	Disease resistance-like protein (TNL)	TAIR
	At1g56540	Disease resistance-like protein (TNL)	TAIR
	At1g57840	pseudogene, putative disease resistance-like protein	TAIR
	At1g58390	rpp7-like protein (CNL)	UniProt
	At1g58400	rpp7-like protein (CNL)	TAIR
	At1g58410	rpp7-like protein (CNL)	UniProt
	At1g58602	rpp7 downy mildew resistance protein	UniProt
	At1g58807	rpp7-like protein (CNL)	UniProt
	At1g58848	rpp7-like protein (CNL)	UniProt
	At1g59124	rpp7-like protein (CNL)	TAIR
	At1g59218	rpp7-like protein (CNL)	UniProt
	At1g59620	rpp7-like disease resistance protein (CNL)	TAIR
	At1g59780	rpp7-like disease resistance protein (CNL)	UniProt
	At1g61100	Disease resistance-like protein (TNL)	TAIR
WRR-OyC3	At3g44400	Disease resistance-like protein (TNL)	TAIR
	At3g44480	Disease resistance-like protein (TNL)	TAIR
	At3g44630	Disease resistance-like protein (TNL)	TAIR
	At3g44670	Disease resistance-like protein (TNL)	TAIR
	At3g51560	Disease resistance-like protein (TNL)	TAIR
	At3g51570	Disease resistance-like protein (TNL)	TAIR
	At3g46530	rpp13 downy mildew resistance homolog (CNL)	UniProt
	At3g50950	rpp13-like protein 4	UniProt
	At3g44410	rpp1 downy mildew resistance homolog (TNL)	TAIR
	At3g46710	rpp13-like disease resistance protein (CNL)	UniProt
	At3g46730	rpp13-like disease resistance protein (CNL)	UniProt
WRR-OyC5	At5g11250	Disease resistance protein (TNL)	TAIR
	At5g17680	Disease resistance protein (TNL)	TAIR
	At5g17880	WRR5 white rust resistance protein (TNL)	TAIR
	At5g17890	WRR6 white rust resistance protein (TNL-LIM)	TAIR
	At5g17970	Disease resistance protein (TNL)	TAIR
	At5g18350	Disease resistance protein (TNL)	TAIR
	At5g18360	Disease resistance protein (TNL)	TAIR
	At5g18370	Disease resistance protein (TNL)	TAIR

3.4 DISCUSSION

The Columbia allele of WRR4 has been previously shown to provide broad spectrum resistance and mask the effect of other *R* genes (Borhan *et al.*, 2008). Here, Columbia-virulent isolates (AcExeter and AcCarlisle) have enabled the identification of additional sources of resistance with potential applicability in transgenic crop production, including intriguing evidence for a potentially broader spectrum resistance in Oy-0 that maps to the same region on chromosome 1 as *WRR4*-Columbia, but also spans a cluster of genes including a downy mildew resistance gene *RPP7* in Columbia.

Comparative sequence analysis of MAGIC parental lines susceptible to AcEm2 and resistant to AcExeter and AcCarlisle using the data from the 1001 Arabidopsis Genome Project revealed Ct-1 and Edi-0 to have an identical Col-0 allele of *WRR4*. Bur-0 and Ler-0 showed multiple non-synonymous mutations throughout the exons of *WRR4*, *WRR5* and *WRR6*, suggesting that these accessions possess alternative sources of resistance to Col-avirulent isolates. The mapping of broad spectrum resistance in Oy-0, Mt-0 and Rsch-4 to both Col-avirulent and Col-virulent isolates indicated that these accessions share a major gene resistance on the bottom arm of chromosome 1. This was supported by fine-scale mapping of resistance in Oy-0 to AcExeter, AcCarlisle and AcEm2 using an F_8 Oy-0 x Col-0 recombinant inbred population.

The observation of a full spectrum of interaction phenotypes amongst the Oy-0 x Col-0 inbreds following inoculations with AcExeter and AcCarlisle, indicates that the major effect resistance at the *WRR-OyC1* locus is masking resistance conferred by genes at other loci. Thus, complete resistance to Col-virulent isolates appears to be dependent on additive and interactive effects from elsewhere in the genome. Most noticeably, the *WRR-OyC3* locus on the bottom arm of chromosome 3 was detected above the threshold for significance for both isolates. This locus contains several genes that encode NB-LRR proteins including examples of specificities for downy mildew resistance, although none of these have so far been attributed to resistance to *A. candida.* However, a second minor effect locus *WRR-OyC5* was detected on the top arm of chromosome 5 using AcExeter, which contains a pair of NBS-LRR genes (*WRR5*/*WRR6*) that are both required for a 'yellowing patch' resistance

phenotype in Col-0 to *A. candida* race 4. Re-sequencing data from the 1001 Arabidopsis Genome Project was used to detect polymorphisms and showed both *WRR5* and *WRR6* are highly polymorphic in Oy-0 compared with Col-0.

Allelic variation in specificity of white rust resistance at the *WRR4-Col* locus has implications for the way crop resistance to *A. candida* could be developed. Col-virulent isolates are evidence of how single *R* genes can be overcome by functional mutations in the pathogen, and this will be discussed further in Chapter 4. Yet Oy-0 demonstrates that alternate and potentially more useful *R*-alleles exist in natural variation of *A. thaliana* with different specificities. New alleles can only be discovered through pathology with isolates capable of breaking known resistance. To extend this, a global diversity collection of 400 *A. thaliana* accessions was screened with AcEm2, AcExeter and AcCarlisle, and 50 accessions were identified with resistance to all three isolates (Appendix 4). Some of these share a similar Oy-0 haplotype of the *WRR4* gene, however most are polymorphic at this locus and may therefore provide a source for alternative broad spectrum resistance. Interestingly the set also revealed a potential host differential between AcExeter and AcCarlisle, indicating that they are different pathotypes.

Pyramiding of functional alleles from multiple loci of *R* genes could potentially enable anticipatory breeding of durable resistance for disease control using conventional marker-assisted breeding. Similarly, decisions of what R-alleles to include in a single 'stacked' construct for GM transformation of a crop variety will depend on decision-support from complementary pathology research (focus of the next Chapter). Knowledge about pathogen avirulence elicitors that match each *R*-gene specificity is essential for choosing the best *R*-alleles and therefore an optimum combination of component genes. For example, stacking *WRR-OyC1* resistance with *WRR4-Col* in the same construct for transformation would provide a good combination for use in transgenic varieties of all brassica crops (*B. rapa, B. juncea* and *B. oleracea*) if they each detect a different avirulence elicitor. However, this gene combination would be redundant, and most likely less durable, if both genes detect the same elicitor. Consequently, the next chapter initiates research to identify candidate avirulence genes in *A. candida.*

Chapter 4

Characterisation of *Albugo candida* diversity including efforts to identify candidates for the predicted WRR4-avirulence gene

4.1 INTRODUCTION

In order to achieve sustainable disease management, it is necessary to understand the effect of deploying resistant cultivars on the population structure of the pathogen. This will enable proactive changes to the resistance gene complement in the crop to counteract adaptive changes in virulence of the pathogen. Many pathogens including *Albugo candida* are not restricted to crops, and therefore wild relatives of the crops such as *Capsella bursa-pastoris* (Shepherds Purse) should also be considered as an overwintering host and potential reservoir of virulent pathotypes that could overcome a newly deployed *R* gene in a crop (Saharan *et al.*, 2014). Thus, an assessment of the genetic variability in the wider pathogen population is important to identify recombination between sub-populations, and hence determine whether wild relatives provide a source of inoculum for crops.

In theory, sustainable disease control could be achieved through monitoring pathogen virulence in a cropping system and using this information as decision-support, for example, to choose a combination of *R*-alleles that target different pathogen effectors for 'stacking' in a single construct for GM transformation of the crop. In practice, such a strategy is already being used in cereals with conventional breeding and release of cultivars with pyramiding of *R*-alleles to control fungal rusts (Chen *et al.*, 2008, Fukuoka *et al.*, 2015). By persistent monitoring of pathogen variability, cultivars can be withdrawn from commercial production and replaced with new cultivars possessing different recognition specificities to control the emergence of new virulent races. In other words, anticipating new races before they can cause significant crop losses is the key challenge for a combined strategy of conventional and GM control options. The alternative of uninformed and widespread deployment of *R* genes will ultimately select for virulence that is no longer controllable through host resistance (McDonald and Linde, 2002, Garcia-Arenal and McDonald, 2003). This situation is analogous to the growing ineffectiveness of chemical control against a range of pests, weeds and diseases.

Brassica hosts and physiological races of *A. candida* provide a useful experimental system to investigate adaptation of a pathogen to a family of closely related hosts. *A. candida* isolates are classified into physiological races

according to their pathogenicity on a defined set of lines that are representative of different host species (Saharan *et al.*, 2014, Srivastava *et al.*, 2004, Hill *et al.*, 1988, Pound and Williams, 1963). This provides an indication of discrete variation in host range, which has been useful for comparing results from researchers in different countries. However, race classification is not an absolute measure of host adaptation. A recent comparison of *A. candida* genomes indicates that sexual recombination has occurred amongst races (McMullan *et al.*, 2015). Thus, recombination and derivation of new pathotypes could provide a source for hybrid variants and new physiological races.

White rust commonly occurs in natural populations of *A. thaliana* (Holub *et al.,* 1995). The disease primarily appears in seedlings and in rosettes of adult plants. The causal pathogen was originally thought to be *A. candida*. However, it was later determined that it is a distinct species based on divergent *ITS1* and *COX2* sequences (Thines *et al.*, 2009). *Albugo laibachii* appears to be narrowly restricted to a single host species, but is broadly virulent in most accessions of *A. thaliana* including the favoured laboratory accession Columbia (Col-0). In contrast, *A. candida* is broadly avirulent in a diverse collection of *A. thaliana.* However, accessions have been found that are susceptible to *A. candida* race 4 (collected from *C. bursa-pastoris*) under laboratory conditions. For example, the Col-0 allele of *WRR4* has been shown to provide broad spectrum resistance to *A. candida* races 2, 4, 7 and 9 (Borhan *et al.*, 2008). This suggests that host genes are a major factor in determining the host range of *A. candida*, with host recognition and subsequent resistance being triggered by races with the corresponding AVR gene.

Although *A. candida* has been useful to investigate disease resistance in *A. thaliana* under controlled environment conditions, it is important to know whether *A. thaliana* is a potential reservoir for this pathogen under field conditions. Thus, a primary aim of this chapter was to determine whether *A. candida* could be readily detected as the cause of white rust in natural populations of *A. thaliana,* particularly in floral stem and leaf tissue (*i.e.,* a largely ignored niche for white rust in *A. thaliana* field biology) of plants growing in close proximity to *C. b. pastoris*, which is the most prolific source of inoculum.

Perennial species of Arabidopsis such as *A. lyrata* (syn. *Arabis lyrata*) are also a natural host for white rust in North America and northern Europe (Jacobson *et al.*, 1998). In the UK, an isolate of *A. candida* from *A. halleri* was collected in 2008 from a white rust epidemic in a glasshouse (controlled environment) at the University of Exeter (AcExeter), which provided the first evidence that *A. candida* can overcome WRR4-mediated resistance (Fairhead, 2012 Masters thesis). This was the first evidence of a previously race non-specific resistance breaking down. Thus, a second aim of this chapter was to determine whether additional Col-virulent isolates could be collected from Arabidopsis under field conditions.

Association genetics can provide a means for identifying virulence determinants in microbial pathogens (Bart *et al.*, 2012). Thus, a third aim of this chapter was to determine whether a collection of Col-0-virulent isolates could be used to search genome-wide effector sequences and identify candidates for *avrWRR4-Col* (*i.e.,* that are conserved amongst Col-avirulent *A. candida* of races 2, 4, 7 and 9 but altered by mutations in the same candidates that could cause loss-of-function amongst the Col-virulent isolates).

As described in the previous chapter, the Norwegian accession of *A. thaliana* Oy-0 carries resistance to AcExeter that maps to an interval containing *WRR4.* It is currently not known whether this *WRR-OyC1* resistance is conferred by an alternative allele of *WRR4* or a tightly linked gene in Oy-0 conferring resistance to both Col-virulent and Col-avirulent *A. candida*. Consequently, there are two possible explanations for a Col-virulence pathotype of *A. candida*. Either AcExeter has mutations in a highly conserved effector (*avrWRR4*) corresponding to recognition by the WRR4-Col protein (as predicted by Borhan *et al.,* 2010) that is also still recognised by an WRR-OyC1 protein, or AcExeter possesses a further avirulence effector that is specifically recognised by WRR-OyC1. Thus, a final aim of this chapter was to use the effector database to search and determine whether candidates exist that could match the *WRR-OyC1* resistance.

4.2 MATERIALS AND METHODS

4.2.1 Collection of *Albugo candida* isolates from diseased floral tissue

Sampling of *A. candida* from floral tissue of *A. thaliana* displaying visible white rust was performed from sites where plants were growing naturally in close proximity to *C. b-pastoris* that was heavily infested with white rust for the purpose of finding Col-0 virulent isolates. Several samples were collected by Professor Eric Holub in 2008 and 2010, and additional samples were collected at the start of this project from fields at Warwick Crop Centre. In addition, isolate Ac167 sampled from a Northern European accession of *Arabidopsis lyrata* with confirmed resistance to Col-0 avirulent isolates was added to the collection. *A. lyrata* (AL) 167 was cultivated on a flat roof site in Stratford-upon-Avon for the purpose of seed multiplication in a location presumed to be isolated from Albugo. In April 2014, plants developed white rust, which could either be the result of infection the previous autumn or infection brought in on the seed from Sweden (the original source of the accession). The isolates were bulked and maintained on of on juvenile leaves of *A. thaliana* Ws-*eds1* as described in Chapter 3.

4.2.2 Phenotyping for functional genomics

Phenotypes on Col-0, WS and four F_9 Col-0 x Ws-0 recombinant inbreds possessing single Col-0 WRR genes (Chapter 3, Figure 1) were examined and assessed using the experimental procedure described in Chapter 3.

4.2.3 Molecular genotyping

Infected tissue samples from *A. thaliana* Ws-*eds1* for each isolate were harvested ten days post inoculation. DNA was extracted using the DNeasy plant mini kit according protocol.

The selective amplification of the complete ITS region was performed on all isolates using the oomycete-specific forward primer DC6 (5-GAGGGACTTTTGGGTAATCA-3) and complementary reverse primer LR-0R (5-GCTTAAGTTCAGCGGGT-3) (Kaur *et al.*, 2011). PCR reactions were conducted using Phusion® high-fidelity DNA polymerase. Reaction volumes of 25 µl were created containing 5 µl of 5 x GC buffer, 0.5 µl of 10 mM dNTPs, 1.25 µl of 10 µm

forward and reverse primers, 1 µl of template, 15.75 µl nuclease-free water and 0.25 µl of DNA polymerase added last. All reactions were prepared on ice. PCR was performed using the standard touchdown program (Table 4.1). A 5 µl sample of PCR product was added to 2 µl of loading dye for assessment using electrophoresis.

Table 4.1. Standard touch down polymerase chain reaction PCR program used for the amplification of Internal Transcribed Spacers (ITS) to determine the species of pathogen causing white rust-like symptoms. D: denaturation, A: annealing, E: extension.

	Start	**Touchdown (8 Cycles)**			**(28 Cycles)**			**Finish**
	D	D	A	E	D	A	E	E
Temperature (°C)	98	98	65 – 57	72	98	56	72	72
Time (minutes)	3.00	0.30	0.30	1.30	0.30	0.30	1.30	7.00

Genomic DNA and PCR products were assessed through electrophoresis, using agarose gel of 1% volume to weight agarose/ Tris Borate buffer (TBE). Centrifuge driven PCR purification was performed using the Qiagen PCR purification kit according to protocol.

A 5 µl aliquot of PCR product and 5 µl of primer at a concentration of 0.5 mM were pipetted into 1.5 ml Eppendorf tubes and sequenced through the LightRun service provided by GACT Biotech. Sequence analysis was performed *in silico*. Geneious version 9.1.4 (Kearse *et al.*, 2012) was used for trimming regions with a greater than 0.5% chance of an error per base, quality control, reference based alignment and polymorphism detection for all sequenced PCR products.

4.2.4 Exome sequence capture

An exome sequence capture method called PathSeq, developed at the Sainsbury Laboratory, was used to characterise sequence variation of genes that encode effector-like proteins in a collection of *A. candida* isolates including two Col-virulent isolates (AcExeter and Ac167) and four *A. candida* race 9 isolates from UK production of *B. oleracea* (Table 4.2). Baits for the PathSeq assay were

designed to capture sequences of secreted proteins ("the effectorome") with respect to alleles of genes annotated in the AcNc2 – SRX884047 reference genome (McMullan *et al.*, 2015), and an additional set of multi-locus sequence typing (MLST) housekeeping genes for isolate characterisation. This sequence information was combined with whole genome sequences from a third Col-virulent isolate AcCarlisle (unpublished) and three standard isolates that are representative of races 2, 4 and 7. Ac2V - SRX884063, AcEM2 - SRX879636 and Ac7V (unpublished).

The combined dataset enabled comparative analyses across isolates, through the presence/absence and allelic variation amongst samples, variation in heterozygosity of individual genes, and between-isolate variation of individual genes. Enabling the determination of candidate *avr* genes through sequence similarity between isolates with the same host differential.

Table 4.2. *Albugo candida* isolates collected from different host species in the UK and Canada. Next generation sequencing data from each isolate was assembled for comparative genomics using short read sequencing (assembled contigs or PathSeq) or and consensus sequence from a the AcNc2 reference genome.

Isolate	Host	Location	Data Type
AcNc2	*Arabidopsis thaliana*	Norwich, UK	Reference genome
AcEm2	*Capsella bursa-pastoris*	East Malling, UK	Whole genome assembly
AcExeter	*A. halleri*	Exeter, UK	PathSeq contigs
AcCarlisle	*A. halleri*	Carlisle, UK	Whole genome assembly
Ac167	*A. lyrata*	Stratford, UK	PathSeq contigs
AcBol	*Brassica oleracea*	Wellesbourne, UK	PathSeq contigs
Ac101	*Brassica oleracea*	Cornwall, UK	PathSeq contigs
Ac102	*Brassica oleracea*	Cornwall, UK	PathSeq contigs
Ac116	*Brassica oleracea*	Cornwall, UK	PathSeq contigs
Ac2v	*Brassica juncea*	Saskatoon, Canada	Whole genome assembly
Ac7v	*Brassica rapa*	Saskatoon, Canada	Whole genome assembly

4.2.5 Data pre-processing and analysis

Bioinformatics support for PathSeq data analysis was provided by Dr. Laura Baxter. The reference genome was defined as the exome sequence of the 411

genes from the AcNc2 genome from which the baits were designed. PathSeq reads, assembled contigs and extracted genome sequence in the case of AcCarlisle were mapped to the reference genome using Bowtie2 aligner (Langmead and Salzberg, 2012). The aligned .sam files were converted into .bam files using SAMtools (version 0.1.19; (Li *et al.*, 2009)). BEDTools was used to obtain the depth and breadth of coverage of each target sequence (Quinlan and Hall, 2010). An additional filtering step characterised all alleles supported by a depth of coverage as less than 10. The *.tsv files were then processed with a custom Perl script (Baxter, unpublished) to generate breadth and depth metrics.

4.3 RESULTS

4.3.1 *Albugo candida* commonly occurs as a pathogen of floral tissue of *Arabidopsis thaliana*

The ten-class phenotype scale (Figure 3.2) was used to characterise the pathogen-host interactions of five isolates that were revived from *A. thaliana* inflorescence tissue, collected from plants that were growing in close proximity to *C. bursa-pastoris*, and another isolate collected from *A. lyrata* (Table 4.3). All six isolates were fully virulent in Ws-0, producing extensive sporulation and no visible host response. ITS sequence and NCBI BLAST against AcNc2 confirmed that all six isolates are *A. candida,* with 100% identity to the reference sequence. Full virulence was also observed in Col-0 with two of these isolates (AcCarlisle and Ac167), similar to AcExeter, thus confirming that natural mutants capable of breaking all layers of Col-0 resistance readily occur in the UK under natural conditions in annual and perennial species of Arabidopsis.

These isolates were tested further using a differential set of Col-0 x Ws-0 recombinant inbreds, with each inbred possessing individual Col-0 WRR genes. Four isolates (AcGear, AcHardewijk, AcTenterden and AcWCC2) each showed a fully non-compatible phentype on CW20, and yellowing phenotype on CW5. These phenotypes were consistent with the standard *A. candida* race 4 isolate AcEm2, suggestive of recognition by WRR4 and WRR5/WRR6. CW14 created a differential, indicating that variation occurs between the isolates in the effector

responsible for recognition by WRR7 (Table 4.3). The three Col-virulent isolates were fully virulent in the inbred different as expected.

Table 4.3. Phenotypic interactions between *Albugo candida* isolates and five *Arabidopsis thaliana* lines that contain differential combinations of White Rust Resistance (*WRR*) genes. The isolates were derived from UK field samples of diseased floral tissue in two annual hosts (*Capsella bursa-pastoris* and *A. thaliana*) and two perennial hosts (*A. lyrata* and *A. halleri*).

		Host differential line (*WRR* gene)				
Pathogen Isolate	Host origin	Col-0 (4,5,6,7)	Ws-0 (none)	CW20 (4)	CW5 (5,6)	CW14 (7)
AcEM2	*C.b. pastoris*	0	8	0	5	7
AcGear	*A. thaliana*	0	8	0	5	7
AcHardewijk	*A. thaliana*	0	8	0	5	0
AcTenterden	*A. thaliana*	0	8	0	5	8
AcWCC2	*A. thaliana*	0	8	0	5	5
AcCarlisle	*A. thaliana*	8	8	8	8	8
Ac167	*A. lyrata*	8	8	nt	nt	nt
AcExeter	*A. halleri*	8	8	8	8	8

nt= not tested

4.3.2 Candidate *avrWRR4-Col* and *avrWRR-OyC1* elicitors were identified by association genetics

A high degree of conservation was observed across seven MLST genes for all *A. candida* race 4 and race 9 isolates enriched through PathSeq.

Col-virulence of *A. candida* could be explained by loss-of-function in an avirulence elicitor detected by the WRR4 protein that is conserved in Col-avirulent isolates, as predicted by Borhan *et al.* (2008). To identify candidate host genes for this predicted *avrWRR4-Col* gene, a threshold of 99% sequence identity was set to identify alleles in the database that are conserved within the Col-avirulent group (AcNc2, AcEm2, Ac2v and Ac7v), and either absent or polymorphic within the Col-virulent group (AcExeter, AcCarlisle and Ac167). Five candidate effectors met these selection criteria, including two examples (CCG31 and CCG71) of presence/absence polymorphisms between the isolate groups (Table 4.4). Conservation of alleles across the Col-avirulent group with non-synonymous mutations compared to the virulent group was used to identify candidate effectors. (Figure 4.1).

Table 4.4 Candidates for an avirulence determinant in *Albugo candida* of recognition by *WRR4*-Col in *A. thaliana*, based on sharing a predicted conserved allele (at least 99% sequence similarity) between a representative group of Col-avirulent isolates including race2 (Ac2V from *Brassica juncea*) and race 4 (AcEm2 from *Capsella bursa-pastoris*), and comparative sequence variation or absence in three Col-virulent isolates (AcExeter, AcCarlisle and Ac167 collected from *Arabidopsis halleri, A. thaliana* and *A. lyrata*, respectively). Identity is shown as a percentage similarity with the allele sequence from the AcNc2 reference genome.

	A. candida isolate					
Effector	AcEm2 race 4	Ac2V race 2	Ac7V race 7	AcExeter	AcCarlisle	Ac167
CCGlike10	100	99.8	99.9	94.9	94.9	94.9
CCG31	100	99.6	100	ND	ND	ND
CCG71	100	99.9	100	ND	ND	ND
CCGlike42	99.9	99.9	99.5	97.9	89.4	97.8
CCG63	99.7	99.2	99.2	98.1	89.1	95.1

ND= sequence not detected (deleted or highly divergent).

Alternatively, Col-virulence of *A. candida* could be explained by an effector that is capable of suppressing an *avrWRR4-Col* avirulence elicitor. In this case, candidates for a suppressive effector would be conserved within the Col-virulent group (at least 99% similarity) and either absent or polymorphic (<99% similarity) in Col-avirulent isolates. The same candidates could also be target avirulence elicitors detected by the R protein encoded by the *WRR-OyC1* gene predicted in Chapter 3. Nine candidate effectors met these selection criteria, including one example (CCG66) of presence/absence polymorphisms between the isolate groups (Table 4.5). Conservation of alleles within the Col-virulent group was confirmed by non-synonymous mutations in nucleotide translations relative to the reference genome (Figure 4.2).

Table 4.5. Candidates for an avirulence determinant in *Albugo candida* of recognition by *WRR4*-Col in *A. thaliana*, based on sharing a predicted conserved allele (at least 99% sequence similarity) between a group of Col-virulent isolates (AcExeter, AcCarlisle and Ac167 collected from *Arabidopsis halleri, A. thaliana* and *A. lyrata*, respectively), and comparative sequence variation or absence in a representative group of Col-0 avirulent isolates including race2 (Ac2V from *Brassica juncea*) and race 4 (AcEm2 from *Capsella bursa-pastoris*), Identity is shown as a percentage similarity with the allele sequence from the AcNc2 reference genome.

	A. candida isolate					
Effector	AcEm2 race 4	Ac2V race 2	Ac7V race 7	AcExeter	AcCarlisle	Ac167
CCG66	ND	ND	ND	99.5	99.5	99.5
CCG27	99.5	ND	95.0	99.1	99.1	99.1
CCGlike41	100	99.7	99.7	99.3	99.3	99.3
CCGlike8	99.6	99.0	ND	99.2	99.2	99.2
CCGlike5	100	99.6	95.8	98.5	98.5	98.5
CCGlike25	99.9	97.6	97.7	97.6	97.6	97.6
CCGlike10	100	99.8	99.9	94.9	94.9	94.9
CCG80	ND	93.7	100.	92.9	92.9	92.9

ND= sequence not detected (deleted or highly divergent).

4.4 DISCUSSION

In the current study, ITS sequencing revealed that *A. candida* could be readily detected in floral tissue as the cause of white rust in natural populations of *A. thaliana,* and that this particular niche can harbour strains which are capable of breaking the broad-spectrum resistance conferred by *WRR4* in *A. thaliana* Columbia (Col-0). Association genetics enabled the identification of candidates for the matching avirulence elicitor *avrWRR4-Col* by using a database of genome-wide effector sequences to search for genes with a conserved allele amongst Col-avirulent isolates of races 2, 4, 7 and 9 but altered by predicted loss-of-function mutations amongst three Col-virulent isolates (table 4.4). Applying a reciprocal search of the database, effectors were identified that have a conserved allele amongst the Col-virulent isolates and mutational variation amongst the Col-avirulent group (table 4.5). These are potentially candidate *avr* elicitors that match a new broad spectrum resistance gene identified near *WRR4* in *A. thaliana* Oy-0 (designated *WRR-OyC1*).

Collectively, this evidence provides a basis from pathology to support anticipatory breeding of white rust control in brassica crops such as oilseed mustard, which is major staple food and biofuel crop in India. Primarily, if white rust resistance mediated by *WRR4* can be broken in wild host populations, then use of this gene in crop production as a sole means of controlling the disease would rapidly select for existing virulence specificity in the pathogen. Verifying which *avr* protein is recognised by WRR4 would provide means of monitoring pathogen populations within and in close proximity to cropping systems for WRR4 virulent pathotypes through conventional PCR and genotyping. Of the five *avrWRR4-Col* candidates identified through association genetics, only one (CCG71) exhibited a hypersensitive cell death response when co-expressed in a transient assay with the WRR4-Col protein in tobacco (Cevik, Sainsbury Lab, *pers. comm*.). The other candidates did not exhibit cell death in the same tobacco experiment and are therefore not expected to function as *avrWRR4-Col* elicitors.

Interestingly, the tobacco transient assay identified three additional effector-like proteins (CCG28, CCG30, CCG33) that elicited hypersensitive cell

death in a *WRR4-Col* dependent manner. None of these met the search criteria because they all had sequence variation (not conserved) within the Col-avirulent isolates across races 2, 4 and 7. For example, CCG28 and CCG30 have conserved alleles between AcNc2 and AcEm2 (race 4 isolates), yet neither were detected in the genomes of Ac2v or Ac7v. Similarly, CCG33 was conserved between AcNc2 and AcEm2, but showed 98.4% variation including multiple non-synonymous mutations throughout the coding sequence in Ac2v andAc7v.

These effectors may not elicit hypersensitive cell death either because they are not expressed naturally by the pathogen in a compatible host in a way that would expose them to detection, or because their detection is actively suppressed from the earliest stages of infection. Dominant suppressors of an avirulence elicitor have been described from genetic studies of fungal pathogens including the seminal work in the flax rust pathosystem by H. H. Flor in *Melampsora lini* reviewed recently by Ellis *et al.* (2007) and in rice blast by Ellingboe (2001). In *A. candida,* the three 'transient elicitors' may be masked by a matching HR-suppressor when expressed by the pathogen in the host-isolate combinations that have been tested thus far.

As discussed in Chapter 3, it is still not known which gene on chromosome 1 in Oy-0 provides major effect resistance to Col-virulent isolates of *A. candida* such as AcExeter and AcCarlisle. However, it is logical to consider whether this source of white rust resistance detects a second *avr* elicitor or the same elicitor as *WRR4-Col.* CCG71 remains as the only potential candidate that is consistent with both genetic and co-expression evidence. Yet it appeared absent from the WRR4 virulent group which we independently confirmed through conventional genotyping as part of this study. Consequently, if CCG71 is responsible for Col-0 WRR4 recognition then this suggests that the Col-0 avirulent group possess a highly conserved unique effector. If *WRR4-OyC1* also detects the same effector as *WRR4-Col*, then stacking the both of these *R*-alleles in the same construct for transformation would be an ineffective strategy. Instead, the durability of the broader spectrum allele (*WRR4-OyC1*) would be more effectively improved by combining it in a background of race non-specific resistance such as conferred by the recessive *aca2* gene from *B. oleracea* described in Chapter 2.

Chapter 5

General discussion

The development of increasingly powerful DNA sequencing technologies such as Genotyping-by-Sequencing and exome sequencing (*e.g.,* RenSeq and PathSeq) are beginning to enable the targeted deployment of durable disease resistance, unachievable through conventional plant breeding. The cultivation of varieties possessing durable resistance has the capacity to sustainably intensify production in two ways. Firstly, production capacity under disease pressure is increased whilst reducing cost and environmental impact of production by reducing the requirement for excessive agrochemical application. Secondly, reduced and targeted agrochemical application on varieties possessing durable resistance will both reduce the capacity of pathogens to evolve specificities capable of overcoming crop resistance and withstanding chemical treatment, enhancing the longevity of both methods of control.

The introgression of dominant or semi-dominant *R* genes into commercial varieties has been achieved through conventional breeding, as selecting for functionality according to phenotype through each generation is relatively straight forward. Yet single *R* genes rapidly become redundant when deployed in monoculture cropping systems through selecting for pathogen variants with mutations or loss of the corresponding effector allele. Examples of such redundancy can be found in *P. infestans*, stem rust of wheat caused by *P. graminis*, blackleg of oilseed rape caused by *L. maculans* and downy mildew of lettuce caused by *B. lactucae* (Ballini *et al.*, 2013, Zhang *et al.*, 2009, Sivasithamparam *et al.*, 2005, Crute and Norwood, 1981).

The increased efficiency and reduced cost of isolating of *R* genes and their corresponding effectors opens the possibility of new breeding strategies for durable resistance. Such as the isolation of core pathogen effectors and the identification of the corresponding *R* gene from crop wild relatives through transient expression experiments.

The feasibility of mapping *R* genes is increasing as the development of next generation sequencing technologies progresses. The ability to enrich and assess variation across gene families identified from reference genomes through technologies such as RenSeq has rapidly enhanced the ability to isolate functional *R* genes. Combined with accurate phenotype data, this can now be performed across extensive diversity collections to identify candidates through

association genetics. The application through bulk segregate analysis has also been realised, where crosses between resistant and susceptible accessions enables the bulk extraction and analysis between resistant and susceptible groups. In the event of a single dominant *R* gene conferring resistance, which can be readily determined by assessing patterns of Mendelian inheritance, representation of both parental lines would be expected across all loci other than that conferring, or tightly linked to resistance.

Although a pioneering development, the successful application of such technology is dependent on the identification and targeting of the correct genes. With different copy numbers present throughout the *Brassicaceae*, and variation between accessions within each species, distinguishing between paralogs remains problematic. The development of pangenomes for bait design based on the whole genome sequence of multiple accessions will ultimately alleviate the problem, yet the ultimate and most compressive method of detection will be association genetics applied to whole genomes. Given the progression of technology this will become a reality across complex plant genomes in the foreseeable future. Currently, sequence capture is based on the assumption that what is baited for will confer resistance, yet as we have demonstrated with the *A. candida* - *B. oleracea* pathosystem this is not always the case. An unbiased approach to *R* gene detection may well reveal resistance mechanisms currently unconsidered. Such as the transfer of complete networks conferring non-host resistance from one organism to another.

Following the identification of *R* genes their function can now be confirmed through gene editing approaches such as clustered regulatory interspaced short palindromic repeat (CRISPR), and introgression into crop varieties through conventional breeding and marker assisted or transgenic approaches. These approaches will increase the feasibility of breeding recessive resistant traits in crop varieties, either through the selection of parental lines homozygous for the resistant gene, or through inducing a resistant allele through gene editing. In *B. oleracea* it would now be possible to induce the recessive race non-specific resistance found in EBH527 with the race specific resistance provided by *WRR4-Col*. Inducing different recognition specificities in the resulting variety would increase the durability of control as the pathogen would

require multiple mutations to subvert host resistance.

A race specific *R* gene will only provide resistance if the corresponding effector is present in the pathogen isolates local to the cropping system and is important for virulence. Consequently, targeted *R* gene deployment is necessary to achieve durable resistance. This has particular relevance to the *WRR-OyC1* allele. Should it be a broader spectrum of *WRR4-Col*, then it would likely provide resistance to races 2, 7 and 9 in *B. juncea, B. rapa* and *B. oleracea,* as described by Borhan et al (2008). If recognition is caused by an effector present in only race 4, then it would have little current application in crop production. Yet there is the potential of sexual reproduction between crop virulent and *WRR4-Col* virulent isolates (which are crop avirulent) through cohabitation of the same host, made possible through the suppression of resistance by the initial infection of a virulent isolate. Such an occurrence may have already occurred, but has not yet been detected through sampling. Consequently, the isolation of *WRR-OyC1* resistance for deployment in crop varieties would be a pre-emptive strategy for mitigating against potential new virulence specificities in crop virulent races.

As demonstrated by the synteny between of *ACA2* in EBH527 with the *WRR5-Col* and *WRR6-Col* locus in *A. thaliana*, and the Aca1 locus in *B. rapa*, conserved syntenic blocks across the *Brassicaceae* appear to confer resistance to different races of *A. candida*. It can be speculated the allelic variation within these blocks has driven pathogen divergence, and the emergence of races with different host ranges. This would in part explain variation within races apparent through different host specifies within a given host species.

It is intriguing that in the centre of the QTL conferring resistance to AcAus is a single QTL homologous to a CC-NB-LRR in the vicinity of *RPP7*, and within the *WRR-OyC1* locus for resistance AcExeter and AcCarlisle. *R* genes often appear in clusters, with different alleles conferring recognition of different avirulence specificities. It is therefore possible that other CC-NB-LRR genes exist in the *A12-AcAus* locus. The increased read length achievable through sequencing technologies such as Pacific Biosciences (PacBio) sequencing would allow effective cDNA sequencing of accession A12 to investigate this further. Whether or not this is the case, it does imply that a CC-NBS-LRR is implicated in resistance to *A. candida* race 4, and a homolog of gene identified in the A12 QTL

exists in the *WRR-OyC1* locus.

It would be possible to investigate whether a CC-NBS-LRR is implicated in resistance to AcExeter and AcCarlisle through investigating the effects of non-functional NDR1 and EDS1 on the phenotype. Since TIR-NBL-LRR gene signal through EDS1, and CC-NBS-LRR genes signal through NDR1, F2 progeny would between Oy-0 and Ws-*eds1* would provide a resource to investigate this further, with differing segregation ratios expected between the phenotypes of Col-0 virulent and Col-0 avirulent isolates expected if Oy-0 recognition is dependent on CC-NBS-LRR recognition (McHale *et al.*, 2006). If the same ratios were observed it would add to the evidence that a different allele of *WRR4* with a different specificity is responsible for recognition of Col-0 virulent isolates, justifying co-expression experiments with the Oy-0 *WRR4* allele and the eleven effectors we have identified as conserved between the Col-0 virulent isolates yet polymorphic within the Col-0 avirulent isolates. Should it be that different *R* genes that are recognising different effectors are responsible for *WRR4-Col* mediated resistance to AcEm2 and *WRR-OyC1* resistance to AcExeter and AcCarlisle, the isolation of the *WRR-OyC1* gene would enable stacking of potentially two NB-LRR type genes. The requirement of the pathogen to mutate multiple effectors would again reduce or delay the occurrence of isolates capable of breaking crop resistance.

In this study we have built on knowledge developed in a model organism and applied it to related crop species. The translation of oomycete resistance from *A. thaliana* into brassica production has the capacity to sustainably intensify the production of oilseed and vegetable brassicas. A small but significant system change in the quest for local, regional and global food security.

Bibliography

Aarts, M. G., Te Lintel Hekkert, B., Holub, E. B., Beynon, J. L., Stiekema, W. J. & Pereira, A. (1998a). Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in Arabidopsis thaliana. *Molecular Plant-Microbe Interactions : MPMI,* **11,** 251-8.

Aarts, N., Metz, M., Holub, E., Staskawicz, B. J., Daniels, M. J. & Parker, J. E. (1998b). Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. *Proceedings of the National Academy of Sciences,* **95,** 10306-10311.

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J. W., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Most, M., Mullen, S., Nichols, R., Nolte, A. W., Parisod, C., Pfennig, K., Rice, A. M., Ritchie, M. G., Seifert, B., Smadja, C. M., Stelkens, R., Szymura, J. M., Vainola, R., Wolf, J. B. W. & Zinner, D. (2013). Hybridization and speciation. *Journal of evolutionary biology,* **26,** 229-246.

Adhikari, T. B., Liu, J. Q., Mathur, S., Wu, C. X. & Rimmer, S. R. (2003). Genetic and molecular analyses in crosses of race 2 and race 7 of *Albugo candida*. *Phytopathology,* **93,** 959-965.

Ahdb (2015). AHDB recommended lists for cereals and oilseeds 2016/2017.

Alford, D. V. (2008). Pest and disease management handbook, John Wiley & Sons.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C. & Ecker, J. R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science,* **301,** 653-7.

Andolfo, G., Jupe, F., Witek, K., Etherington, G. J., Ercolano, M. R. & Jones, J. D. (2014). Defining the full tomato NB-LRR resistance gene repertoire using genomic and cDNA RenSeq. *BMC plant biology,* **14,** 120.

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. *Reference Source*.

Axtell, M. J. & Staskawicz, B. J. (2003). Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell,* **112,** 369-377.

Ballini, E., Lauter, N. & Wise, R. (2013). Prospects for advancing defense to cereal rusts through genetical genomics.

Bansal, V., Tewari, J., Stringam, G. & Thiagarajah, M. (2005). Histological and inheritance studies of partial resistance in the *Brassica napus*–*Albugo candida* host – pathogen interaction. *Plant Breeding,* **124,** 27-32.

Bart, R., Cohn, M., Kassen, A., Mccallum, E. J., Shybut, M., Petriello, A., Krasileva, K., Dahlbeck, D., Medina, C., Alicai, T., Kumar, L., Moreira, L. M., Rodrigues Neto, J., Verdier, V., Santana, M. A., Kositcharoenkul, N., Vanderschuren, H., Gruissem, W., Bernal, A. & Staskawicz, B. J. (2012). High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences USA,* **109,** E1972-9.

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R. & Mathews, S. (2010). Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences,* **107,** 18724-18728.

Bennetzen, J. L., Qin, M.-M., Ingels, S. & Ellingboe, A. H. (1988). Allele-specific and Mutator-associated instability at the Rpl disease-resistance locus of maize.

Biffen, R. H. (1907). Studies in the inheritance of disease-resistance. *The Journal of Agricultural Science,* **2,** 109-128.

Biffen, R. H. (1912). Studies in the inheritance of disease resistance. II. *The Journal of Agricultural Science,* **4,** 421-429.

Biga, M. (1955). Review of the species of the genus Albugo based on the morphology of the conidia. *Sydowia,* **9,** 339-358.

Blanc, G. & Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *The Plant Cell,* **16,** 1679-1691.

Block, A. & Alfano, J. R. (2011). Plant targets for Pseudomonas syringae type III effectors: virulence targets or guarded decoys? *Current Opinion in Microbiology,* **14,** 39-46.

Boisson, B., Giglione, C. & Meinnel, T. (2003). Unexpected protein families including cell defense components feature in the N-myristoylome of a higher eukaryote. *The Journal of Biological Chemistry,* **278,** 43418-29.

Borhan, M. H., Brose, E., Beynon, J. L. & Holub, E. B. (2001). White rust (*Albugo candida*) resistance loci on three Arabidopsis chromosomes are closely linked to downy mildew (*Peronospora parasitica*) resistance loci. *Molecular Plant Pathology,* **2,** 87-95.

Borhan, M. H., Gunn, N., Cooper, A., Gulden, S., Tor, M., Rimmer, S. R. & Holub, E. B. (2008). WRR4 encodes a TIR-NB-LRR protein that confers broad-

spectrum white rust resistance in *Arabidopsis thaliana* to four physiological races of *Albugo candida. Molecular Plant-Microbe Interactions : MPMI,* **21,** 757-68.

Borhan, M. H., Holub, E. B., Beynon, J. L., Rozwadowski, K. & Rimmer, S. R. (2004). The Arabidopsis TIR-NB-LRR gene RAC1 confers resistance to *Albugo candida* (white rust) and is dependent on EDS1 but not PAD4. *Molecular Plant-Microbe Interactions,* **17,** 711-719.

Borhan, M. H., Holub, E. B., Kindrachuk, C., Omidi, M., Bozorgmanesh-Frad, G. & Rimmer, S. R. (2010). WRR4, a broad-spectrum TIR-NB-LRR gene from *Arabidopsis thaliana* that confers white rust resistance in transgenic oilseed Brassica crops. *Molecular Plant Pathology,* **11,** 283-91.

Bowers, J. E., Chapman, B. A., Rong, J. & Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature,* **422,** 433-438.

Bowler, P. J. (1989). Evolution: the history of an idea, Univ of California Press.

Boyes, D. C., Nam, J. & Dangl, J. L. (1998). The *Arabidopsis thaliana* RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. *Proceedings of the National Academy of Sciences,* **95,** 15849-15854.

Broman, K. W. & Sen, S. (2009). A Guide to QTL Mapping with R/qtl, Springer.

Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics,* **19,** 889-90.

Caillaud, M. C., Wirthmueller, L., Sklenar, J., Findlay, K., Piquerez, S. J. M., Jones, A. M. E., Robatzek, S., Jones, J. D. G. & Faulkner, C. (2014). The Plasmodesmal Protein PDLP1 Localises to Haustoria-Associated Membranes during Downy Mildew Infection and Regulates Callose Deposition. *Plos Pathogens,* **10**.

Cevik, V., Holub, E. B. (2014) Unpublished data, cited with permission.

Chattopadhyay, C., Agrawal, R., Kumar, A., Meena, R., Faujdar, K., Chakravarthy, N., Kumar, A., Goyal, P., Meena, P. & Shekhar, C. (2011). Epidemiology and development of forecasting models for *White rust* of *Brassica juncea* in India. *Archives of Phytopathology and Plant Protection,* **44,** 751-763.

Chen, Y., Singh, S., Rashid, K., Dribnenki, P. & Green, A. (2008). Pyramiding of alleles with different rust resistance specificities in *Linum usitatissimum L. Molecular Breeding,* **21,** 419-430.

Cheng, F., Wu, J. & Wang, X. (2014). Genome triplication drove the diversification of Brassica plants. *Horticulture Research,* **1,** 14024.

Cheung, W., Gugel, R. & Landry, B. (1998). Identification of RFLP markers linked to the white rust resistance gene (Acr) in mustard (*Brassica juncea* (L.) Czern. and Coss.). *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada,* **41,** 626-628.

Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. (2006). Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell,* **124,** 803-814.

Choi, D. & Priest, M. J. (1995). A key to the genus Albugo. *Mycotaxon (USA)*.

Choi, Y. J., Shin, H. D. & Thines, M. (2009). The host range of Albugo candida extends from Brassicaceae through Cleomaceae to Capparaceae. *Mycological Progress,* **8,** 329-335.

Cooper, A. J., Cevik, V., Holub, E. B. (2012) Unpublished data, cited with permission.

Cooper, A. J., Latunde-Dada, A. O., Woods-Tor, A., Lynn, J., Lucas, J. A., Crute, I. R. & Holub, E. B. (2008). Basic compatibility of *Albugo candida* in *Arabidopsis thaliana* and *Brassica juncea* causes broad-spectrum suppression of innate immunity. *Molecular plant-microbe interactions : MPMI,* **21,** 745-56.

Crute, I. & Norwood, J. M. (1981). The identification and characteristics of field resistance to lettuce downy mildew (Bremia lactucae Regel). *Euphytica,* **30,** 707-717.

Crute, I. R., Holub, E. B., Tor, M., Brose, E. & Beynon, J. L. 1993. The identification and mapping of loci in *Arabidopsis thaliana* for recognition of the fungal pathogens: *Peronospora parasitica* (downy mildew) and *Albugo candida* (white blister). *Advances in Molecular Genetics of Plant-Microbe Interactions, Vol. 2.* Springer.

Crute, I. R. & Pink, D. (1996). Genetics and utilization of pathogen resistance in plants. *The Plant Cell,* **8,** 1747.

Cusack, B. P. & Wolfe, K. H. (2007). When gene marriages don't work out: divorce by subfunctionalization. *Trends in genetics,* **23,** 270-272.

Dangl, J. L., Horvath, D. M. & Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science,* **341,** 746-51.

Defra (2003). Variability in Fungal Pathogens. *Horticulture Research International*.

Delourme, R., Chevre, A., Brun, H., Rouxel, T., Balesdent, M., Dias, J., Salisbury, P., Renard, M. & Rimmer, S. (2006). Major gene and polygenic resistance to Leptosphaeria maculans in oilseed rape (Brassica napus). *European Journal of Plant Pathology,* **114,** 41-52.

Delwiche, P. A. (1976). Genetic studies in Brassica nigra (L.) Koch, University of Wisconsin--Madison.

Dixon, M. S., Jones, D. A., Keddie, J. S., Thomas, C. M., Harrison, K. & Jones, J. D. (1996). The tomato Cf-2 disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell,* **84,** 451-459.

Downey, R. & Rimmer, S. (1993). Agronomic Improvement in. *Advances in Agronomy,* **50,** 1.

Ellingboe, A. H. (2001). Plant—pathogen interactions: Genetic and comparative analyses. *European Journal of Plant Pathology,* **107,** 79-84.

Ellis, J. G., Dodds, P. N. & Lawrence, G. J. (2007). Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annual Review of Phytopathology,* **45,** 289-306.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S. & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One,* **6,** e19379.

Eynck, C., Koopmann, B., Grunewaldt-Stoecker, G., Karlovsky, P. & Von Tiedemann, A. (2007). Differential interactions of *Verticillium longisporum* and V. dahliae with *Brassica napus* detected with molecular and histological techniques. *European Journal of Plant Pathology,* **118,** 259-274.

Fairhead, S. E. (2012). Developing marker-assisted breeding of white resistance for use in oilseed and vegetable brassicas (Masters thesis). University of Warwick

Fan, Z., Rimmer, S. & Stefansson, B. (1983). Inheritance of resistance to *Albugo candida* in rape (*Brassica napus L.*). *Canadian Journal of Genetics and Cytology,* **25,** 420-424.

Fao (2016). Food security, Nutrition and Peace. *Proceedings of the United Nations Security Council Meeting: New York, 29 March 2016*

Fitt, B. D., Brun, H., Barbetti, M. & Rimmer, S. (2006). World-wide importance of phoma stem canker (*Leptosphaeria maculans and L. biglobosa*) on oilseed rape (*Brassica napus*). *Sustainable strategies for managing Brassica napus (oilseed rape) resistance to Leptosphaeria maculans (phoma stem canker).* Springer.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology,* **9,** 275-296.

Fukuoka, S., Saka, N., Mizukami, Y., Koga, H., Yamanouchi, U., Yoshioka, Y., Hayashi, N., Ebana, K., Mizobuchi, R. & Yano, M. (2015). Gene pyramiding enhances durable blast disease resistance in rice. *Scientific Reports,* **5,** 7773.

Gan, X., Stegle, O., Behr, J., Steffen, J. G., Drewe, P., Hildebrand, K. L., Lyngsoe, R., Schultheiss, S. J., Osborne, E. J., Sreedharan, V. T., Kahles, A., Bohnert, R.,

Jean, G., Derwent, P., Kersey, P., Belfield, E. J., Harberd, N. P., Kemen, E., Toomajian, C., Kover, P. X., Clark, R. M., Ratsch, G. & Mott, R. (2011). Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature,* **477,** 419-23.

Garcia-Arenal, F. & Mcdonald, B. A. (2003). An analysis of the durability of resistance to plant viruses. *Phytopathology,* **93,** 941-52.

Goritschnig, S., Steinbrenner, A. D., Grunwald, D. J. & Staskawicz, B. J. (2016). Structurally distinct *Arabidopsis thaliana* NLR immune receptors recognize tandem WY domains of an oomycete effector. *The New Phytologist*.

Gupta, K. & Saharan, G. (2002). Identification of pathotypes of *Albugo candida* with stable characteristic symptoms on Indian mustard. *Journal of Mycology and Plant Pathology (India)*.

Haley, C. S. & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity,* **69,** 315-324.

Haudry, A., Platts, A. E., Vello, E., Hoen, D. R., Leclercq, M., Williamson, R. J., Forczek, E., Joly-Lopez, Z., Steffen, J. G. & Hazzouri, K. M. (2013). An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nature Genetics,* **45,** 891-898.

Hedrick, P. W. (2013). Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology,* **22,** 4606-4618.

Hill, C., Crute, I., Sherriff, C. & Williams, P. (1988). Specificity of *Albugo candida* and *Peronospora parasitica* pathotypes toward rapid-cycling crucifers. *Cruciferae Newsletter,* **13,** 112-113.

Hirai, M. (2006). Genetic analysis of clubroot resistance in Brassica crops. *Breeding Science,* **56,** 223-229.

Holub, E. B. (2007). Natural variation in innate immunity of a pioneer species. *Current Opinion in Plant Biology,* **10,** 415-24.

Holub, E. B., Brose, E., Tor, M., Clay, C., Crute, I. R. & Beynon, J. L. (1995). Phenotypic and genotypic variation in the interaction between *Arabidopsis thaliana* and *Albugo candida*. *Molecular plant-microbe interactions : MPMI,* **8,** 916-28.

Jacobson, D. J., Lefebvre, S. M., Ojerio, R. S., Berwald, N. & Heikkinen, E. (1998). Persistent, systemic, asymptomatic infections of *Albugo candida*, an oomycete parasite, detected in three wild crucifer species. *Canadian Journal of Botany,* **76,** 739-750.

Johnston, J. S., Pepper, A. E., Hall, A. E., Chen, Z. J., Hodnett, G., Drabek, J., Lopez, R. & Price, H. J. (2005). Evolution of genome size in *Brassicaceae*. *Annals of botany,* **95,** 229-235.

Jones, D. A., Thomas, C. M., Hammond-Kosack, K. E., Balint-Kurti, P. J. & Jones, J. D. (1994). Isolation of the tomato Cf-9 gene for resistance to Cladosporium fulvum by transposon tagging. *Science- New York Then Washington.* 789-789.

Jones, J. D. G. & Dangl, J. L. (2006). The plant immune system. *Nature,* **444,** 323-329.

Joshi, R. K. & Nayak, S. (2010). Gene pyramiding-A broad spectrum technique for developing durable stress resistance in crops. *Biotechnology and Molecular Biology Reviews,* **5,** 51-60.

Jupe, F., Chen, X., Verweij, W., Witek, K., Jones, J. D. G. & Hein, I. (2014). Genomic DNA Library Preparation for Resistance Gene Enrichment and Sequencing (RenSeq) in Plants. *Plant-Pathogen Interactions: Methods and Protocols, 2nd Edition,* **1127,** 291-303.

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I. & Jones, J. D. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *The Plant Journal : For Cell and Molecular Biology,* **76,** 530-44.

Kagale, S., Robinson, S. J., Nixon, J., Xiao, R., Huebert, T., Condie, J., Kessler, D., Clarke, W. E., Edger, P. P. & Links, M. G. (2014). Polyploid evolution of the *Brassicaceae* during the Cenozoic era. *The Plant Cell,* **26,** 2777-2791.

Kaur, P., Jost, R., Sivasithamparam, K. & Barbetti, M. J. (2011). Proteome analysis of the *Albugo candida-Brassica juncea* pathosystem reveals that the timing of the expression of defence-related genes is a crucial determinant of pathogenesis. *Journal of Experimental Botany,* **62,** 1285-98.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S. & Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* **28,** 1647-1649.

Keen, N. (1990). Gene-for-gene complementarity in plant-pathogen interactions. *Annual Review of Genetics,* **24,** 447-463.

Kim, H. G., Kwon, S. J., Jang, Y. J., Nam, M. H., Chung, J. H., Na, Y. C., Guo, H. W. & Park, O. K. (2013). GDSL LIPASE1 Modulates Plant Immunity through Feedback Regulation of Ethylene Signaling. *Plant Physiology,* **163,** 1776-1791.

Klosterman, S. J., Atallah, Z. K., Vallad, G. E. & Subbarao, K. V. (2009). Diversity, pathogenicity, and management of Verticillium species. *Annual Review of Phytopathology,* **47,** 39-62.

Kole, C., Williams, P. H., Rimmer, S. R. & Osborn, T. C. (2002). Linkage mapping of genes controlling resistance to white rust (*Albugo candida*) in *Brassica rapa* (syn. campestris) and comparative mapping to Brassica napus and *Arabidopsis thaliana*. *Genome / National Research Council Canada = Genome / Conseil national de recherches Canada,* **45,** 22-7.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant, C. & Mott, R. (2009). A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS genetics,* **5,** e1000551.

Krueger, F. (2015). Trim Galore. *A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files*.

Kwon, S. J., Jin, H. C., Lee, S., Nam, M. H., Chung, J. H., Kwon, S. I., Ryu, C. M. & Park, O. K. (2009). GDSL lipase-like 1 regulates systemic resistance associated with ethylene signaling in Arabidopsis. *The Plant Journal : For Cell and Molecular Biology,* **58,** 235-45.

Lander, E. S. & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics,* **121,** 185-199.

Langenbach, C., Schultheiss, H., Rosendahl, M., Tresch, N., Conrath, U. & Goellner, K. (2016). Interspecies gene transfer provides soybean resistance to a fungal pathogen. *Plant Biotechnology Journal,* **14,** 699-708.

Langmead, B. & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods,* **9,** 357-9.

Lava, S. S., Heller, A. & Spring, O. (2013). Oospores of Pustula helianthicola in sunflower seeds and their role in the epidemiology of white blister rust. *IMA Fungus,* **4,** 251-258.

Leckie, D., Astley, D., Crute, I., Ellis, P., Pink, D., Boukema, I., Monteiro, A. & Dias, S. The location and exploitation of genes for pest and disease resistance in European gene bank collections of horticultural brassicas. ISHS Brassica Symposium-IX Crucifer Genetics Workshop 407, 1994. 95-102.

Lee, T., Yang, S., Kim, E., Ko, Y., Hwang, S., Shin, J., Shim, J. E., Shim, H., Kim, H., Kim, C. & Lee, I. (2015). AraNet v2: an improved database of co-functional gene networks for the study of *Arabidopsis thaliana* and 27 other nonmodel plant species. *Nucleic Acids Research,* **43,** D996-1002.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics,* **27,** 2987-93.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* **25,** 2078-9.

Lincoln, S. E. & Lander, E. S. (1992). Systematic detection of errors in genetic linkage data. *Genomics,* **14,** 604-610.

Links, M. G., Holub, E., Jiang, R. H. Y., Sharpe, A. G., Hegedus, D., Beynon, E., Sillito, D., Clarke, W. E., Uzuhashi, S. & Borhan, M. H. (2011). De novo sequence assembly of *Albugo candida* reveals a small genome relative to other biotrophic oomycetes. *Bmc Genomics,* **12**.

Liu, J., Parks, P. & Rimmer, S. (1996). Development of monogenic lines for resistance to Albugo candida from a Canadian *Brassica napus* cultivar. *Phytopathology,* **86,** 1000.

Liu, Q. & Rimmer, S. R. (1990). Effects of host genotype, inoculum concentration, and incubation-temperature on white rust development in oilseed rape. *Canadian Journal of Plant Pathology-Revue Canadienne De Phytopathologie,* **12**.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, X., Wang, J., Lu, K., Fang, Z., Bancroft, I., Yang, T. J., Hu, Q., Wang, X., Yue, Z., Li, H., Yang, L., Wu, J., Zhou, Q., Wang, W., King, G. J., Pires, J. C., Lu, C., Wu, Z., Sampath, P., Wang, Z., Guo, H., Pan, S., Yang, L., Min, J., Zhang, D., Jin, D., Li, W., Belcram, H., Tu, J., Guan, M., Qi, C., Du, D., Li, J., Jiang, L., Batley, J., Sharpe, A. G., Park, B. S., Ruperao, P., Cheng, F., Waminal, N. E., Huang, Y., Dong, C., Wang, L., Li, J., Hu, Z., Zhuang, M., Huang, Y., Huang, J., Shi, J., Mei, D., Liu, J., Lee, T. H., Wang, J., Jin, H., Li, Z., Li, X., Zhang, J., Xiao, L., Zhou, Y., Liu, Z., Liu, X., Qin, R., Tang, X., Liu, W., Wang, Y., Zhang, Y., Lee, J., Kim, H. H., Denoeud, F., Xu, X., Liang, X., Hua, W., Wang, X., Wang, J., Chalhoub, B. & Paterson, A. H. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications,* **5,** 3930.

Lorieux, M. (2012). MapDisto: fast and efficient computation of genetic linkage maps. *Molecular Breeding,* **30,** 1231-1235.

Lysak, M. A., Cheung, K., Kitschke, M. & Bureš, P. (2007). Ancestral chromosomal blocks are triplicated in *Brassiceae* species with varying chromosome number and genome size. *Plant Physiology,* **145,** 402-410.

Majer, D., Lewis, B. & Mithen, R. (1998). Genetic variation among field isolates of Pyrenopeziza brassicae. *Plant Pathology,* **47,** 22-28.

Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Frary, A., Ganal, M. W., Spivey, R., Wu, T., Earle, E. D. & Tanksley, S. D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science,* **262,** 1432-6.

Mcdonald, B. A. & Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology,* **40,** 349-79.

Mchale, L., Tan, X., Koehl, P. & Michelmore, R. W. (2006). Plant NBS-LRR proteins: adaptable guards. *Genome Biology,* **7,** 212.

Mcmullan, M., Gardiner, A., Bailey, K., Kemen, E., Ward, B. J., Cevik, V., Robert-Seilaniantz, A., Schultz-Larsen, T., Balmuth, A., Holub, E., Van Oosterhout, C. & Jones, J. D. (2015). Evidence for suppression of immunity as a driver for genomic introgressions and host range expansion in races of Albugo candida, a generalist parasite. *eLife,* **4**.

Mei, J., Qian, L., Disi, J., Yang, X., Li, Q., Li, J., Frauen, M., Cai, D. & Qian, W. (2011). Identification of resistant sources against Sclerotinia sclerotiorum in Brassica species with emphasis on *B. oleracea*. *Euphytica,* **177,** 393-399.

Mendel, G. (1865). Experiments in plant hybridization. *Verhandlungen des naturforschenden Vereins Brünn.) Available online: www. mendelweb. org/Mendel. html (accessed on 1 January 2013)*.

Meyerowitz, E. M. & Somerville, C. R. (1994). Arabidopsis, Cold Spring Harbor Laboratory Press.

Monaghan, J. & Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Current Opinion in Plant Biology,* **15,** 349-357.

Munnik, T. (2001). Phosphatidic acid: an emerging plant lipid second messenger. *Trends in Plant Science.,* **6,** 227-233.

Nowicki, M., Nowakowska, M., Niezgoda, A. & Kozik, E. (2012). Alternaria black spot of crucifers: symptoms, importance of disease, and perspectives of resistance breeding. *Vegetable Crops Research Bulletin,* **76,** 5-19.

Oh, I. S., Park, A. R., Bae, M. S., Kwon, S. J., Kim, Y. S., Lee, J. E., Kang, N. Y., Lee, S., Cheong, H. & Park, O. K. (2005). Secretome analysis reveals an Arabidopsis lipase involved in defense against *Alternaria brassicicola*. *Plant Cell,* **17,** 2832-47.

Panjabi-Massand, P., Yadava, S. K., Sharma, P., Kaur, A., Kumar, A., Arumugam, N., Sodhi, Y. S., Mukhopadhyay, A., Gupta, V., Pradhan, A. K. & Pental, D. (2010). Molecular mapping reveals two independent loci conferring resistance to *Albugo candida* in the east European germplasm of oilseed mustard *Brassica juncea*. *Theoretical and Applied Genetics,* **121,** 137-45.

Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., Town, C. D., Nixon, J., Krishnakumar, V., Bidwell, S. L., Denoeud, F., Belcram, H., Links, M. G., Just, J., Clarke, C., Bender, T., Huebert, T., Mason, A. S., Pires, C. J., Barker, G., Moore, J., Walley, P. G., Manoli, S., Batley, J., Edwards, D., Nelson, M. N., Wang, X., Paterson, A. H., King, G., Bancroft, I., Chalhoub, B. & Sharpe, A. G. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biology,* **15,** R77.

Peele, H. M., Guan, N., Fogelqvist, J. & Dixelius, C. (2014). Loss and retention of resistance genes in five species of the Brassicaceae family. *BMC plant biology,* **14,** 298.

Penninckx, I., Eggermont, K., Terras, F., Thomma, B., De Samblanx, G. W., Buchala, A., Métraux, J.-P., Manners, J. M. & Broekaert, W. F. (1996). Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. *The Plant Cell,* **8,** 2309-2323.

Petrie, G. (1994). New races of *Albugo candida* (white rust) in Saskatchewan and Alberta. *Canadian Journal of Plant Pathology,* **16,** 251-252.

Pilet, M., Delourme, R., Foisset, N. & Renard, M. (1998). Identification of QTL involved in field resistance to light leaf spot (*Pyrenopeziza brassicae*) and blackleg resistance (*Leptosphaeria maculans*) in winter rapeseed (*Brassica napus L.*). *Theoretical and Applied Genetics,* **97,** 398-406.

Pound, G. & Williams, P. (1963). Biological Races of Albugo candida.

Prabhu, K. V., Somers, D., Rakow, G. & Gugel, R. (1998). Molecular markers linked to white rust resistance in mustard Brassica juncea. *Theoretical and Applied Genetics,* **97,** 865-870.

Provart, N. J., Alonso, J., Assmann, S. M., Bergmann, D., Brady, S. M., Brkljacic, J., Browse, J., Chapple, C., Colot, V. & Cutler, S. (2016). 50 years of Arabidopsis research: highlights and future directions. *New Phytologist,* **209,** 921-944.

Pryor, T. (1987). The origin and structure of fungal disease resistance genes in plants. *Trends in Genetics,* **3,** 157-161.

Quinlan, A. R. & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics,* **26,** 841-842.

Rimmer, S., Mathur, S. & Wu, C. (2000). Virulence of isolates of *Albugo candida* from western Canada to Brassica species. *Canadian Journal of Plant Pathology,* **22,** 229-235.

Ronald, P. C., Salmeron, J. M., Carland, F. M. & Staskawicz, B. J. (1992). The cloned avirulence gene avrPto induces disease resistance in tomato cultivars containing the Pto resistance gene. *Journal of Bacteriology,* **174,** 1604-11.

Saharan, G. & Verma, P. (1992). White rusts: a review of economically important species, International Development Research Centre.

Saharan, G. S., Verma, P. R., Meena, P. D. & Kumar, A. (2014). White Rust of Crucifers: Biology, Ecology and Management, Springer.

Sato, T., Okamoto, J., Degawa, Y., Matsunari, S., Takahashi, K. & Tomioka, K. (2009). White rust of Ipomoea caused by *Albugo ipomoeae-panduratae* and *A.*

ipomoeae-hardwickii and their host specificity. *Journal of General Plant Pathology,* **75,** 46-51.

Schmidt, R. & Bancroft, I. (2011). Genetics and Genomics of the *Brassicaceae*, Springer.

Sen, Ś. & Churchill, G. A. (2001). A statistical framework for quantitative trait mapping. *Genetics,* **159,** 371-387.

Seo, M. & Koshiba, T. (2002). Complex regulation of ABA biosynthesis in plants. *Trends in Plant Science,* **7,** 41-8.

Shah, J. (2005). Lipids, lipases, and lipid-modifying enzymes in plant disease resistance. *Annual Review of Phytopathology,* **43,** 229-260.

Shepherd, K. & Mayo, G. (1972). Genes conferring specific plant disease resistance. *Science,* **175,** 375-380.

Simon, M., Loudet, O., Durand, S., Berard, A., Brunel, D., Sennesal, F. X., Durand-Tardif, M., Pelletier, G. & Camilleri, C. (2008). Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics,* **178,** 2253-64.

Singh, U. S., Doughty, K. J., Nashaat, N. I., Bennett, R. N. & Kolte, S. J. (1999). Induction of Systemic Resistance to *Albugo candida* in *Brassica juncea* by Pre- or Coinoculation with an Incompatible Isolate. *Phytopathology,* **89,** 1226-32.

Sivasithamparam, K., Barbetti, M. J. & Li, H. (2005). Recurring challenges from a necrotrophic fungal plant pathogen: a case study with Leptosphaeria maculans (causal agent of blackleg disease in Brassicas) in Western Australia. *Annals of Botany,* **96,** 363-377.

Song, W.-Y., Wang, G.-L., Chen, L.-L. & Kim, H.-S. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science,* **270,** 1804.

Speulman, E., Bouchez, D., Holub, E. B. & Beynon, J. L. (1998). Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis thaliana*. *The Plant Journal : For Cell and Molecular Biology,* **14,** 467-74.

Srivastava, P. S., Narula, A., Srivastava, S. & Bhojwani, S. S. (2004). Plant biotechnology and molecular markers, Springer.

Staskawicz, B. J., Dahlbeck, D. & Keen, N. T. (1984). Cloned avirulence gene of *pseudomonas-syringae* pv glycinea determines race-specific incompatibility on glycine-max (l) merr. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* **81,** 6024-6028.

Stukenbrock, E. H., Christiansen, F. B., Hansen, T. T., Dutheil, J. Y. & Schierup, M. H. (2012). Fusion of two divergent fungal individuals led to the

recent emergence of a unique widespread pathogen species. *Proceedings of the National Academy of Sciences USA,* **109,** 10954-9.

Thines, M., Choi, Y. J., Kemen, E., Ploch, S., Holub, E. B., Shin, H. D. & Jones, J. D. (2009). A new species of Albugo parasitic to *Arabidopsis thaliana* reveals new evolutionary patterns in white blister rusts (*Albuginaceae*). *Persoonia,* **22,** 123-8.

Thines, M., Zipper, R., Schäuffele, D. & Spring, O. (2006). Characteristics of Pustula tragopogonis (syn. *Albugo tragopogonis*) newly occurring on cultivated sunflower in Germany. *Journal of Phytopathology,* **154,** 88-92.

Thomma, B. P., Eggermont, K., Tierens, K. F.-J. & Broekaert, W. F. (1999). Requirement of Functional Ethylene-Insensitive 2Gene for Efficient Resistance of Arabidopsis to Infection by *Botrytis cinerea*. *Plant Physiology,* **121,** 1093-1101.

U, N. (1935). Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Journal of Japanese Botany,* **7,** 389-452.

Uricaru, R., Rizk, G., Lacroix, V., Quillery, E., Plantard, O., Chikhi, R., Lemaitre, C. & Peterlongo, P. (2015). Reference-free detection of isolated SNPs. *Nucleic Acids Research,* **43,** e11.

Van Der Hoorn, R. A., Van Der Ploeg, A., De Wit, P. J. & Joosten, M. H. (2001). The C-terminal dilysine motif for targeting to the endoplasmic reticulum is not required for Cf-9 function. *Molecular Plant-Microbe Interactions,* **14,** 412-415.

Van Der Linden, C. G., Wouters, D. C., Mihalka, V., Kochieva, E. Z., Smulders, M. J. & Vosman, B. (2004). Efficient targeting of plant disease resistance loci using NBS profiling. *Theoretical and Applied Genetics,* **109,** 384-393.

Varshney, A., Mohapatra, T. & Sharma, R. (2004). Development and validation of CAPS and AFLP markers for white rust resistance gene in *Brassica juncea*. *Theoretical and Applied Genetics,* **109,** 153-159.

Verma, P., Harding, H., Petrie, G. & Williams, P. (1975). Infection and temporal development of mycelium of *Albugo candida* in cotyledons of four Brassica species. *Canadian Journal of Botany,* **53,** 1016-1020.

Verma, P., Saharan, G., Bartaria, A. & Shivpuri, A. (1999). Biological races of *Albugo candida* on *Brassica juncea* and *B. rapa* var. toria in India. *Journal of Mycology and Plant Pathology (India)*.

Vicente, J. G., Gunn, N. D., Bailey, L., Pink, D. a. C. & Holub, E. B. (2012). Genetics of resistance to downy mildew in *Brassica oleracea* and breeding towards durable disease control for UK vegetable production. *Plant Pathology,* **61,** 600-609.

Vicente, J. G., Taylor, J., Sharpe, A., Parkin, I., Lydiate, D. & King, G. (2002). Inheritance of race-specific resistance to *Xanthomonas campestris* pv. campestris in Brassica genomes. *Phytopathology,* **92,** 1134-1141.

Walker, J. & Priest, M. (2007). A new species of Albugo on Pterostylis (Orchidaceae) from Australia: confirmation of the genus Albugo on a monocotyledonous host. *Australasian Plant Pathology,* **36,** 181-185.

Walley, P. G., Teakle, G. R., Moore, J. D., Allender, C. J., Pink, D. A., Buchanan-Wollaston, V. & Barker, G. C. (2012). Developing genetic resources for pre-breeding in *Brassica oleracea* L.: an overview of the UK perspective. *Journal of Plant Biotechnology,* **39,** 62-68.

Wang, X., Richards, J., Gross, T., Druka, A., Kleinhofs, A., Steffenson, B., Acevedo, M. & Brueggeman, R. (2013). The rpg4-mediated resistance to wheat stem rust (Puccinia graminis) in barley (Hordeum vulgare) requires Rpg5, a second NBS-LRR gene, and an actin depolymerization factor. *Molecular Plant-Microbe Interactions : MPMI,* **26,** 407-18.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I. & Cheng, F. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics,* **43,** 1035-1039.

Warwick, S., Francis, A. & Al-Shehbaz, I. (2006). *Brassicaceae*: species checklist and database on CD-Rom. *Plant Systematics and Evolution,* **259,** 249-258.

Williams, P. (1985). White rust (*Albugo candida* (Pers. ex. Hook.) Kuntze.). Crucifer genetics cooperative (CRGC) resource book. *University Wisconsin, USA*, 1-7.

Witek, K., Jupe, F., Witek, A. I., Baker, D., Clark, M. D. & Jones, J. D. (2016). Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology,* **34,** 656-60.

Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics,* **2,** 333-341.

Wu, J., Cai, G., Tu, J., Li, L., Liu, S., Luo, X., Zhou, L., Fan, C. & Zhou, Y. (2013). Identification of QTLs for resistance to Sclerotinia stem rot and BnaC. IGMT5. a as a candidate gene of the major resistant QTL SRC6 in *Brassica napus*. *PloS one,* **8,** e67740.

Yang, Y., Costa, A., Leonhardt, N., Siegel, R. S. & Schroeder, J. I. (2008). Isolation of a strong Arabidopsis guard cell promoter and its potential as a research tool. *Plant Methods,* **4,** 1.

Yu, J., Tehrim, S., Zhang, F., Tong, C., Huang, J., Cheng, X., Dong, C., Zhou, Y., Qin, R., Hua, W. & Liu, S. (2014). Genome-wide comparative analysis of NBS-encoding genes between Brassica species and *Arabidopsis thaliana*. *BMC Genomics,* **15,** 3.

Zhang, N. W., Pelgrom, K., Niks, R. E., Visser, R. G. & Jeuken, M. J. (2009). Three combined quantitative trait loci from nonhost *Lactuca saligna* are sufficient to provide complete resistance of lettuce against *Bremia lactucae*. *Molecular Plant-Microbe Interactions,* **22,** 1160-1168.

Zhao, J. & Meng, J. (2003). Genetic analysis of loci associated with partial resistance to Sclerotinia sclerotiorum in rapeseed (Brassica napus L.). *Theoretical and Applied Genetics,* **106,** 759-764.

Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology,* **53,** 247-73.

Appendix 1

Brassica oleracea accessions from the wild species and diversity fixed foundation set phenotyped 12 days' post inoculation with *Albugo candida* AcBoWells

Founder accession no.	Crop type	12 day phenotype
A12DHd	Kale, Chinese white	6
ARS18	Kale	6
BI87053	Broccoli	7
C04001	*B. oleracea x B. alboglabra* wild species	7
C04001	*B. oleracea x B. alboglabra* wild species	7
C04001	*B. oleracea x B. alboglabra* wild species	6
C04001	*B. alboglabra*	7
C04003	*B. oleracea x B. atlantica*	6
C04003	*B. atlantica*	6.3
C04006	*B. bourgaei x B. oleracea*	6
C04006	*B. bourgaei x B. oleracea*	5.1
C04006	*B. bourgaei*	6.2
C04008	*B. cretica x B. oleracea*	6
C04008	*B. cretica x B. oleracea*	6
C04008	*B. cretica x B. oleracea*	6
C04008	*B. cretica*	6.1
C04009	*B. cretica*	6
C04010	*B. cretica*	6
C04010	*B. cretica*	6
C04011	*B. cretica x B. oleracea*	6
C04011	*B. cretica x B. oleracea*	6
C04011	*B. cretica*	6
C04012	*B. cretica x B. oleracea*	7
C04012	*B. cretica x B. oleracea*	6
C04012	*B. cretica x B. oleracea*	6
C04013	*B. oleracea x B. cretica*	6.1
C04013	*B. oleracea x B. cretica*	6
C04013	*B. oleracea x B. cretica*	6
C04013	*B. cretica*	7
C04014	*B. cretica x B. oleracea*	6
C04014	*B. cretica x B. oleracea*	6

C04014	*B. cretica*	6
C04015	*B. hilarionis*	6.1
C04016	*B. hilarionis x B. oleracea*	6
C04016	*B. hilarionis x B. oleracea*	6
C04016	*B. hilarionis*	6
C04017	*B. hilarionis*	6
C04018	*B. incana*	7
C04019	*B. incana*	6
C04020	*B. incana*	6
C04023	*B. oleracea x B. incana*	6
C04023	*B. incana*	6
C04024	*B. insularis*	5.1
C04028	*B. macrocarpa*	5
C04028	*B. macrocarpa*	5
C04029	*B. macrocarpa*	6.1
C04030	*B. macrocarpa*	6
C04032	*B. macrocarpa*	6
C04033	*B. macrocarpa*	7.3
C04035	*B. macrocarpa*	7
C04036	*B. macrocarpa*	7
C04037	*B. macrocarpa*	7
C04038	*B. macrocarpa*	7
C04039	*B. macrocarpa*	7
C04040	*B. macrocarpa*	7
C04041	*B. macrocarpa*	7
C04042	*B. macrocarpa*	7
C04043	*B. macrocarpa*	7
C04044	*B. maurorum*	2
C04045	*B. oleracea x B. montana*	6.1
C04045	*B. montana x B. oleracea*	6
C04045	*B. montana*	2
C04047	*B. oleracea*	7
C04048	*B. oleracea*	7
C04050	*B. oleracea*	6
C04051B	*B. oleracea* wild species *x B. oleracea*	6
C04051B	*B. oleracea* wild species *x B. oleracea*	6
C04051B	*B. oleracea*	7
C04052	*B. oleracea* wild species *x B. oleracea*	6.3
C04052	*B. oleracea* wild species *x B. oleracea*	7.1
C04052	*B. oleracea* wild species *x B. oleracea*	6.2
C04052	*B. oleracea*	7
C04053	*B. oleracea*	7
C04054	*B. oleracea x B. oleracea* wild cabbage	6

C04054	*B. oleracea x B. oleracea* wild cabbage	6
C04054	*B. oleracea*	7
C04055	*B. oleracea* wild species *x B. oleracea*	6
C04055	*B. oleracea* wild species *x B. oleracea*	6
C04056	*B. oleracea*	7
C04057	*B. oleracea*	7
C04060	*B. oleracea x B. oleracea* wild species	7
C04060	*B. oleracea x B. oleracea* wild species	6
C04060	*B. oleracea*	7
C04061	*B. oleracea*	6
C04062	*B. oleracea* wild cabbage *x B. oleracea*	7
C04062	*B. oleracea* wild cabbage *x B. oleracea*	6
C04062	*B. oleracea* wild cabbage *x B. oleracea*	7
C04062	*B. oleracea*	2
C04063	*B. oleracea x B. oleracea* wild cabbage	7.1
C04063	*B. oleracea*	6
C04066	*B. oleracea*	6
C04067	*B. oleracea*	6
C04067	*B. oleracea*	2
C04068	*B. oleracea*	7
C04069	*B. oleracea*	7
C04070	*B. montana x B. oleracea*	6.1
C04070	*B. montana x B. oleracea*	6
C04070	*B. montana x B. oleracea*	6
C04073	*B. rupestris*	7.1
C04077	*B. incana x B. oleracea*	6
C04077	*B. incana x B. oleracea*	6
C04077	*B. incana x B. oleracea*	6
C04077	*B. incana*	6
C04079	*B. incana x B. oleracea*	6
C04079	*B. incana x B. oleracea*	6
C04079	*B. incana* (listed as villosa in seed book)	6
C04080	*B. oleracea x B. incana*	6
C04080	*B. incana*	6.1
C04081	*B. incana*	6
C04082	*B. incana*	7
C04083	*B. villosa*	7.1
C04084	*B. villosa bivoniana*	7.1
C04085	*B. villosa bivoniana*	7
C04086	*B. villosa bivoniana*	7
C04087	*B. villosa bivoniana*	6.1
C04088	*B. villosa bivoniana*	6

C04092	*B. villosa drepanensis*	6.2
C04093	*B. villosa tinei*	2
C04093	*B. villosa tinei*	7
C04097	*B. oleracea capitata*	7
C04098	*B. oleracea capitata*	6
C04098	*B. oleracea capitata*	6
CA25	Cauliflower, autumn	6
Cal18b	Broccoli	7
Cor12b	Broccoli	7
Early Big Broccoli	Broccoli	7
GDDH33	Broccoli	6
HRIGRU000302	Brussels sprout	7
HRIGRU000434	Brussels sprout	7
HRIGRU002175	Brussels sprout	7
HRIGRU002291	Cabbage	6.1
HRIGRU002400	Broccoli	7
HRIGRU002401	Broccoli	2
HRIGRU002405	Broccoli	6
HRIGRU002484	Kohl rabi	7
HRIGRU002787	Brussels sprout	7
HRIGRU002891	Cauliflower, winter	5
HRIGRU003543	Broccoli, sprouting	7
HRIGRU003546	Broccoli	6
HRIGRU003591	Kale, borecole	7
HRIGRU003592	Kale, borecole	7.1
HRIGRU003595	Kale, borecole	6.2
HRIGRU004239	Cauliflower, autumn	5
HRIGRU004293	Kale	7
HRIGRU004492	Cauliflower, winter	7
HRIGRU004607	Brussels sprout	7.1
HRIGRU004701	Calabrese	2
HRIGRU004705	Broccoli	7
HRIGRU004707	Broccoli	7
HRIGRU004709	Calabrese	7
HRIGRU004710	Calabrese	7
HRIGRU004771	Cabbage, summer	7
HRIGRU004785	Brussels sprout	7
HRIGRU004818	Cauliflower, autumn	6.1
HRIGRU004845	Caulilfower, green	6
HRIGRU004846	Caulilfower, green autumn	6
HRIGRU004854	Cauliflower, green	7
HRIGRU004858	Cauliflower, romanesco	7.1
HRIGRU004860	Caulilfower, green	7

HRIGRU004864	Cauliflower, romanesco	7
HRIGRU004865	Caulilfower, green	5.1
HRIGRU004872	Broccoli	7
HRIGRU004885	Broccoli	6.3
HRIGRU004991	Cauliflower, summer	6.3
HRIGRU005085	Kale	6
HRIGRU005085	Kale	7
HRIGRU005086	Brussels sprout	7
HRIGRU005108	Kale, Chinese white	2
HRIGRU005259	Caulilfower, green	6.1
HRIGRU005281	Calabrese	7
HRIGRU005282	Calabrese	7
HRIGRU005295	Broccoli	2
HRIGRU005297	Broccoli	7
HRIGRU005312	Caulilfower, green	7
HRIGRU005364	Caulilfower, green	7
HRIGRU005389	Kohl rabi	7
HRIGRU005419	Broccoli	3
HRIGRU005429	Calabrese	7
HRIGRU005430	Cauliflower, autumn	5.1
HRIGRU005577	Cabbage, hybrid autumn	7
HRIGRU005611	Kohl rabi	7
HRIGRU005652	Cabbage	7
HRIGRU006212	Brussels sprout	7
HRIGRU006226	Kale	6
HRIGRU006318	Broccoli, black	2
HRIGRU006556	Cabbage, heading	7
HRIGRU006628	Cauliflower	6.1
HRIGRU006630	Calabrese	7
HRIGRU006797	Cauliflower, autumn	2
HRIGRU007458	Cauliflower	6
HRIGRU007474	Cauliflower, autumn	2
HRIGRU007514	Broccoli	2
HRIGRU007517	Broccoli	6
HRIGRU007518	Broccoli	6
HRIGRU007520	Cauliflower, romanesco	6
HRIGRU007543	Kale, Chinese	6
HRIGRU007543	Kale, Chinese	7
HRIGRU007544	Kale, Chinese white	7
HRIGRU007799	Tronchuda cabbage	7
HRIGRU007826	Cabbage, pickling	7
HRIGRU008202	Kale	2
HRIGRU008266	Cauliflower	2

HRIGRU008267	Kohl rabi	6
HRIGRU008362	Cabbage, savoy	7
HRIGRU008558	Cauliflower, autumn	5
HRIGRU008567	Cauliflower, romanesco	7
HRIGRU008571	Cauliflower, romanesco	7
HRIGRU008658	Calabrese	5
HRIGRU008723	Kohl rabi	6
HRIGRU008732	Cabbage, savoy	7
HRIGRU009467	Tronchuda kale	4
HRIGRU009489	Tronchuda kale	7
HRIGRU009553	Tronchuda kale	2
HRIGRU009579	Tronchuda cabbage	7.1
HRIGRU009598	Tronchuda cabbage	7.1
HRIGRU009617	Cabbage	7
HRIGRU009845	Kale, Chinese white hybrid	2
HRIGRU009846	Kale	7
HRIGRU009979	Cabbage	7
HRIGRU010772	Calabrese	7
HRIGRU011446	Tronchuda cabbage	7.3
HRIGRU011555	Cabbage, loose head	2
HRIGRU011729	Cauliflower, autumn	7
HRIGRU011800	Broccoli	7.1
HRIGRU011802	Broccoli	7.3
HRIGRU011803	Broccoli?	6
HRIGRU013023	Kale	2
MarDH34	Broccoli	7
Senna	Kale, Chinese white	7
Sho5a	Broccoli	7
Sir5a	Cauliflower	7
TO1000DH3	Rapid cyling alboglabra	6

Appendix 2

Internal primers used for whole gene sequencing of candidate genes within a locus conferring resistance to *Albugo candida* AcBoWells

MARKER	PRIMER SEQUENCE 5' TO 3
BO-02	GAATCTTCGTGGAGTAGTTAGAAAATCTAT TTAGTATTTTCTTGTACTCGGAGATTAGTA ATTGAAAACCGGTATCGTTGAGAATTCCAA
BO-03	TGCATAAAACTAAAACTCCCCAACC
BO-04	TAAAATTCTGTCCTCCTTTCAGTTTACATT AGATTGTGTACCCAAAAGAAAGCAG GGACCCAAAGATTGAACAAGTGAAT TTGAAGTTAATGGTGGTAGAGCAGA TCACCATTCTCATTTGGAAGGTTTG
BO-05	TACTTGACTCTGGAAACTTCGTGAT CTTTGACCATCCAACCAATGTGTTA TAACCTTAGAGCTGTCCAGAGAGTA

Appendix 3

Three replications of A12DH x EBH527 F5 recombinant inbred mapping population scored 10 days' post inoculation with *Albugo candida* AcBoWells. Quantitative phenotype score taken in accordance with figure 2.8

Accession	Rep 1	Rep 2	Rep 3	Mean Phenotype
RIL_1	1	1	1	1
RIL_2	1	6	7	4.7
RIL_3	6	6	6	6
RIL_4	NA	6	6	6
RIL_5	5.3	4	5.3	4.9
RIL_6	1	7	7	5
RIL_7	6	6	6	6
RIL_8	7	7	7	7
RIL_9	6	6	6	6
RIL_10	1	2	1	1.3
RIL_11	7	6	6	6.3
RIL_12	NA	2	2	2
RIL_13	6	7	7	6.7
RIL_14	1	NA	NA	1
RIL_15	1	4	2	2.3
RIL_16	3	6	7	5.3
RIL_17	7	5	7	6.3
RIL_18	7	7	7	7
RIL_19	6	NA	NA	6
RIL_20	2	2	2	2
RIL_21	7	6	7	6.7
RIL_22	7	4	5.3	5.4
RIL_23	NA	3	1	2
RIL_24	6	6	7	6.3
RIL_25	2	4	2	2.7
RIL_26	7	7	7	7
RIL_27	7	7	7	7

RIL_28	6	NA	1	3.5
RIL_29	1	NA	NA	1
RIL_30	7	6	7	6.7
RIL_31	6	NA	NA	6
RIL_32	6	6	7	6.3
RIL_33	3	3	3	3
RIL_34	1	2	2	1.7
RIL_35	6	6	6	6
RIL_36	7	7	7	7
RIL_37	NA	7	7	7
RIL_38	7	NA	NA	7
RIL_39	1	3	3	2.3
RIL_40	6	6	3.3	5.1
RIL_41	7	NA	NA	7
RIL_42	7	3	5	5
RIL_43	7	6	6	6.3
RIL_44	6	6	7	6.3
RIL_45	1	NA	1	1
RIL_46	6	6	7	6.3
RIL_47	1	1	1	1
RIL_48	1	4	1	2
RIL_49	2	3	1	2
RIL_50	5	3	1	3
RIL_51	NA	6	7	6.5
RIL_52	7	7	7	7
RIL_53	7	7	7	7
RIL_54	6	6	6	6
RIL_55	NA	6	7	6.5
RIL_56	6	NA	NA	6
RIL_57	7	4	5	5.3
RIL_58	NA	1	1	1
RIL_59	1	1	1	1
RIL_60	6	7	7	6.7
RIL_61	NA	6	6	6
RIL_62	7	6	7	6.7
RIL_63	6	6	7	6.3
RIL_64	2	3	5	3.3
RIL_65	7	7	7	7
RIL_66	7	6	7	6.7
RIL_67	2	4	5	3.6
RIL_68	7	6	6	6.3
RIL_69	3	6	NA	4.5
RIL_70	2	2	1	1.7

Appendix 4

Gene models retrieved from the *Brassica oleracea* TO1000 reference genome (Parkin *et al,.* 2014) in a QTL conferring resistance to *Albugo candida* AcAus in accession A12DH.

GENE ID	DESCRIPTION
BO4G037350	Rop guanine nucleotide exchange factor%2C putative
BO4G037330	Receptor-like protein kinase
BO4G037340	Eukaryotic aspartyl protease family protein
BO4G038410	TTF-type zinc finger protein with HAT dimerization domain-containing protein
BO4G038400	hypothetical protein
BO4G038390	serpin
BO4G038380	conserved hypothetical protein
BO4G038350	Copper amine oxidase family protein
BO4G038370	BREVIS RADIX-like protein
BO4G038360	Nucleotide-diphospho-sugar transferase family protein
BO4G038450	hypothetical protein
BO4G038440	Plant protein of unknown function (DUF641)
BO4G038430	O-sialoglycoprotein endopeptidase%2C putative
BO4G038420	tRNA 2'-phosphotransferase
BO4G038470	Transketolase
BO4G038460	3-phosphoshikimate 1-carboxyvinyltransferase
BO4G038500	Secretory carrier membrane protein family protein
BO4G038510	Aminotransferase-like%2C plant mobile domain family protein
BO4G038530	Ulp1 protease family protein
BO4G038520	Chaperone protein dnaJ
BO4G038550	Cystatin/monellin superfamily protein
BO4G038540	Myosin heavy chain-related protein
BO4G038560	G-box-binding factor
BO4G038490	O-sialoglycoprotein endopeptidase%2C putative
BO4G038480	tRNA 2'-phosphotransferase
BO4G038600	conserved hypothetical protein
BO4G038660	Chitinase family protein

BO4G038670	Disease resistance protein
BO4G038640	RING finger-like protein
BO4G038650	Ribose-phosphate pyrophosphokinase
BO4G038620	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
BO4G038610	conserved hypothetical protein
BO4G038630	conserved hypothetical protein
BO4G038570	WWE protein-protein interaction domain family protein
BO4G038590	Plastid-lipid associated protein
BO4G038580	Cysteine/histidine-rich C1 domain-containing protein
BO4G038700	Ubiquitin family protein
BO4G038720	EID1-like F-box protein
BO4G038730	Baculoviral IAP repeat-containing protein
BO4G038710	Protein phosphatase 2c%2C putative
BO4G038740	Seed maturation protein
BO4G038750	conserved hypothetical protein
BO4G038760	AT hook motif DNA-binding family protein
BO4G038680	Peroxidase superfamily protein
BO4G038690	Glycine cleavage system H protein
BO4G038800	DAG protein
BO4G038840	Ribonuclease H-like superfamily protein .
BO4G038830	conserved hypothetical protein
BO4G038820	VQ motif-containing protein
BO4G038810	TTF-type zinc finger protein with HAT dimerization domain-containing protein
BO4G038780	Glutathione S-transferase T3
BO4G038790	Ribosomal protein-like protein
BO4G038770	conserved hypothetical protein
BO4G038850	Separin
BO4G038880	RING finger-like protein
BO4G038860	Separase
BO4G038870	Zinc finger (CCCH-type) family protein
BO4G038900	conserved hypothetical protein
BO4G038920	O-fucosyltransferase-like protein
BO4G038910	WD40 repeat-containing protein SMU1
BO4G038950	Receptor-kinase%2C putative
BO4G038940	Putative plant snare
BO4G038930	hypothetical protein
BO4G038890	conserved hypothetical protein
BO4G039000	Pentatricopeptide repeat-containing protein
BO4G039010	Glycine cleavage system H protein
BO4G039020	Nck-associated protein
BO4G038970	Trypsin family protein
BO4G038960	Histone-lysine N-methyltransferase%2C putative

BO4G038980	Phosphate-responsive 1 family protein
BO4G038990	DCD (Development and Cell Death) domain protein
BO4G039050	Potassium transporter
BO4G039060	Bifunctional purine biosynthesis protein PurH
BO4G039030	Exostosin family protein
BO4G039040	En/Spm-like transposon protein
BO4G039070	Pentatricopeptide repeat-containing protein
BO4G039080	UDP-N-acetylglucosamine pyrophosphorylase
BO4G039150	Leucine-rich repeat receptor-like protein kinase family protein
BO4G039160	CLAVATA3/ESR (CLE)-related protein
BO4G039170	RING/U-box superfamily protein
BO4G039180	root hair specific 4 .
BO4G039090	Ring finger protein%2C putative
BO4G039100	SLL3 ORF2 protein%2C putative
BO4G039140	Translation initiation factor eIF-2B subunit epsilon%2C putative
BO4G039130	Phosphatidylinositol N-acetylglucosaminyltransferase subunit C
BO4G039120	Ring finger protein%2C putative
BO4G039110	hypothetical protein
BO4G039190	Bromodomain-containing factor
BO4G039200	CTP synthase family protein
BO4G039220	Chaperone protein dnaJ-related-like protein
BO4G039210	hydroxyproline-rich glycoprotein family protein
BO4G039240	conserved hypothetical protein
BO4G039230	UDP-glucose 4-epimerase%2C putative
BO4G039260	WRKY DNA-binding protein
BO4G039250	Zinc ion binding protein
BO4G039290	FAD-binding Berberine family protein
BO4G039270	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
BO4G039280	FAD-binding Berberine family protein
BO4G039300	Maternal effect embryo arrest
BO4G039390	Pistil extensin-like protein
BO4G039380	Class III homeodomain-leucine zipper
BO4G039370	Nuclear transcription factor Y subunit A-7
BO4G039360	RNA polymerase I specific transcription initiation factor RRN3 family protein
BO4G039350	Tubulin%2C beta chain
BO4G039340	Ribosomal protein-like protein
BO4G039330	fatty acid hydroxylase
BO4G039320	hypothetical protein
BO4G039310	fatty acid hydroxylase

BO4G039400	Cytochrome P450%2C putative
BO4G039420	Multidrug resistance protein ABC transporter family
BO4G039410	Glycolipid transfer protein (GLTP) family protein
BO4G039480	N(6)-adenine-specific DNA methyltransferase 2%2C putative
BO4G039470	Mitochondrial transcription termination factor family protein
BO4G039440	Protein kinase superfamily protein
BO4G039430	hypothetical protein
BO4G039460	Geranylgeranyl diphosphate synthase
BO4G039450	Plastid transcriptionally active
BO4G039490	conserved hypothetical protein
BO4G039510	Pyruvate dehydrogenase E1 component subunit beta
BO4G039520	hypothetical protein
BO4G039530	Homeodomain leucine zipper family IV protein
BO4G039500	jasmonate-zim-domain protein
BO4G039590	Alkylated DNA repair protein alkB 8 putative
BO4G039550	Similarity to non-LTR retroelement reverse transcriptase
BO4G039560	ATP-dependent Clp protease proteolytic subunit
BO4G039570	ARM repeat superfamily protein
BO4G039580	hypothetical protein
BO4G039540	auxin response factor
BO4G039600	Galactose oxidase/kelch repeat superfamily protein
BO4G039670	En/Spm-like transposon protein
BO4G039650	Galactose oxidase/kelch repeat superfamily protein
BO4G039660	Galactose oxidase/kelch repeat superfamily protein
BO4G039630	Retrotransposon protein%2C putative%2C Ty1-copia subclass
BO4G039640	conserved hypothetical protein
BO4G039610	hypothetical protein
BO4G039620	Galactose oxidase/kelch repeat superfamily protein

Appendix 5

Arabidopsis thaliana accession phenotyped 10 days post inoculation *Albugo candida* Col-0 avirulent isolate AcEm2, and two Col-0 virulent isolates AcCarlisle and AcExeter. Quantitative phenotype score taken in accordance with figure 3.2

Accession	AcEm1	AcCarlisle	AcExeter
Aco-1	8	NA	8
Alst-1	0	0	0
Amb-1	NA	8	8
Amb-1	NA	8	8
An-1	0	0	8
AN-1	3	8	0
An-2	0	0	0
Ang-0	5	0	8
Ba-1	0	8	8
Bak-0	0	8	NA
Bak-2	8	8	8
Bak-7	8	0	8
Ban-2	NA	NA	8
BAY-0	0	8	8
Bay-0	3	8	8
Bc-1	0	0	8
Bea-1	0	0	0
Bek-1	0	NA	0
Bet-1	8	8	8
Bg-2	3	0	8
Bid-1	0	0	0
Big-1	0	0	NA
BIL-5	0	8	8
Bil-5	0	8	NA
BIL-7	8	8	8
Bla-1	0	0	0

Bla-1	0	3	0
Blh-2	0	0	0
Bog-2	0	NA	8
BOR-1	0	8	8
BOR-4	0	8	8
BR-0	3	8	8
Br-0	NA	NA	8
Brm-1	0	NA	NA
Bs-1	3	8	8
Bsch-0	0	0	8
Bu-0	0	0	0
Bur-0	0	8	0
BUR-0	0	8	8
Bur-0	0	8	8
C24	0	0	0
C24 (CO-1)	0	0	0
Ca-0	3	0	8
Cal-2	0	8	8
Cam-1	NA	NA	8
Cdm-0	3	8	8
Che-1	0	0	0
Chi-1	3	NA	NA
Chr-1	0	8	NA
CIBC-17	8	8	8
CIBC-5	3	0	8
CIBC17	0	8	8
CIBC5	8	8	8
Cnt-3	0	8	8
Co-2	0	0	8
Coc-1	0	0	NA
COL-0	0	8	8
Col-0	0	8	8
Cra-1	3	NA	NA
CS22491	3	8	8
CT-0	0	8	8
Ct-1	0	0	8
Cul-1	0	8	NA
CVI-0	0	NA	NA
Cvi-0	8	8	8
Del-10	8	8	8
Dog-4	8	8	8
Duc-1	0	8	NA
Dun-1	0	0	0

EDEN-1	8	8	8
EDEN-2	8	8	8
EDI-0	3	8	8
Edi-0	3	NA	8
Edi-1	0	8	8
Edi-2	5	8	8
EI-2	0	8	8
Est-1	0	0	8
EST-1	0	8	8
Est-1	0	8	8
Ey-1.5.11	8	8	8
Ey-15.2	8	0	8
Ey-2	8	8	8
Ey-20	8	8	8
Ey-4	8	8	0
Fab-1	NA	NA	8
FAB-2	8	8	8
FAB-4	3	8	8
Far-1	3	NA	8
Fav-1	0	0	0
FEI-0	0	0	0
Fei-0	0	0	0
Fei-0	3	0	0
Fof-1	NA	0	NA
For-1	0	NA	0
Frd-1	0	0	0
Ga-0	0	0	8
GA-0	0	8	0
God-1	8	0	8
Goettingen-7	8	8	8
GOT-22	8	8	8
GOT-7	8	8	0
GU-0	3	8	8
Gü-0	0	8	8
Gy-0	0	0	8
GY-0	3	8	0
H55	0	8	8
Haes-1	3	8	8
Haes-6	0	8	8
Haw-1	0	0	NA
Hey-1	8	8	8
Hh-0	0	3	0
Hil-1	0	8	8

HKT2-4	0	8	8
HR-10	0	0	0
HR-10	0	3	8
HR-5	0	8	0
HR5	0	NA	NA
Hs-0	8	0	9
ICE-1 (Bolin-1)	8	8	8
ICE-102 (Glado-1)	0	8	8
ICE-104 (Lago-1)	NA	8	8
ICE-106 (Mammo-1)	0	NA	8
ICE-107 (Mammo-2)	3	0	8
ICE-112 (Moran-1)	0	0	8
ICE-119 (Timpo-1)	3	0	8
ICE-120 (Valsi-1)	NA	8	8
ICE-127 (Kly-1)	NA	0	8
ICE-130 (Kly-4)	3	0	8
ICE-134 (Koz-2)	3	0	8
ICE-138 (Leb-3)	3	8	8
ICE-150 (Sij-1)	3	0	8
ICE-152 (Sij-2)	0	8	8
ICE-153 (Sij-4)	0	8	8
ICE-163 (Altenb-2)	0	0	0
ICE-169 (Bozen-1-1)	NA	0	8
ICE-173 (Bozen 1-2)	3	0	8
ICE-181 (Mitterberg-1)	3	8	8
ICE-212 (Castelfed 4.1)	8	8	8
ICE-213 (Castelfed 4.2)	8	8	8
ICE-216 (Rovero-1)	8	8	8
ICE-226 (Vezzano 2.1)	8	8	8
ICE-228 (Vezzano 2.2)	8	8	8
ICE-29 (Slavi-1)	8	8	8
ICE-33 (Jablo-1)	8	8	8
ICE-36 (Dobra-1)	3	8	8
ICE-50 (Toufl-1)	0	8	8
ICE-60 (Stepn-2)	3	8	8
ICE-61 (Stepn-1)	0	8	8

ICE-63 (Copac-1)	3	NA	8
ICE-7 (Lecho-1)	0	0	8
ICE-70 (Borsk-1)	NA	8	8
ICE-71 (Shigu-1)	3	0	8
ICE-72 (Shigu-2)	0	8	8
ICE-73 (Kidr-1)	NA	8	8
ICE-75 (Krazo-2)	3	8	8
ICE-79 (Voeran-1)	8	8	8
ICE-92 (Angit-1)	0	0	0
ICE-93 (Apost-1)	8	8	8
ICE-97 (Ciste-1)	0	8	8
ICE-98 (Ciste-2)	0	0	0
Igt-1	0	8	8
In-0	8	0	8
Inv-1	3	8	8
KAS-1	0	NA	0
Kas-1	NA	8	8
Kastel-1	0	0	0
Kil-1	NA	8	8
Kn-0	8	8	8
KNO-10	0	0	0
KNO-18	0	0	0
KNO-18	0	0	0
Koch-1	3	8	8
KONDARA	0	8	8
Kondara	0	8	8
Ksk-1	0	8	8
Ksk-2	NA	8	8
Kyl-1	0	NA	NA
KZ-1	9	8	8
KZ-9	3	0	8
KZ1	3	8	8
Lan-1	0	NA	8
Laz-1	0	0	8
Lee-1	0	0	8
Leg-1	8	NA	8
LER-1	0	8	8
Ler-1	0	8	8
Lerik1-3	0	8	8
Lha-1	3	8	8
LL-0	0	0	8
LL-0	0	8	8
LOV-1	0	0	0

Lov-1	0	0	8
Löv-1	NA	8	NA
LOV-5	0	0	0
Lov-5	0	NA	8
LP2-2	0	8	8
Lp2-2	0	8	8
LP2-6	0	8	8
Lwe-0	0	0	0
LZ-0	0	0	0
Mer-6	8	8	8
Mit-1	0	8	0
MR-0	8	8	0
Mrk-0	8	0	8
MRK-0	8	8	8
MS-0	8	8	8
MT-0	0	0	0
Mt-0	0	NA	NA
MZ-0	8	8	8
Nas-0	0	8	8
ND-1	3	8	8
Nd-1	3	8	8
Nemrut-1	0	8	8
NFA-10	0	8	8
NFA-10	NA	8	8
NFA-8	0	8	0
Nie1-2	0	8	8
NOK-3	8	8	8
Nor-1	NA	0	0
Not-1	0	NA	8
Nun-1	0	NA	0
OMO-2-1	8	NA	8
OMO-2-3	8	8	8
Omo2-1	8	8	NA
Omo2-3	8	0	8
Or-0	8	0	8
Oy-0	0	0	0
OY-0	0	NA	NA
Pa-1	0	0	8
Pdw-1	3	8	8
Ped-0	0	8	8
Pee-1	0	8	8
PHW-26	0	0	8
PHW-28	0	0	8

PHW-33	8	8	8
PHW-35	0	0	0
PHW-36	0	0	8
PHW-37	0	8	8
Pla-0	NA	0	0
Ply-1	NA	0	NA
Pn-0	0	8	8
PNA-10	0	0	0
Pna-10	0	0	0
PNA-17	0	8	8
Pna-17	NA	8	8
Pog-0	0	0	8
Poo-1	3	8	8
Pra-6	0	8	8
PRO-0	0	8	0
Pu2-23	0	8	8
PU2-23	8	8	8
PU2-7	3	8	8
Pu2-7	3	8	8
Qui-0	0	8	8
Ra-0	0	0	9
RA-0	0	8	8
Rd-0	0	8	8
REN-1	3	8	8
REN-11	0	8	0
Rennes-1	3	0	8
Rennes-11	0	0	8
Rip-1	NA	8	8
Rmx-A02	0	0	0
RMX-A02	0	8	0
RMX-A180	0	0	0
Rmx-A180	0	0	3
Rom-1	0	0	8
Rot-1	0	8	8
Rou-0	9	0	0
Roy-1	0	8	0
RRS-10	0	0	0
RRS-10	0	0	0
RRS-7	0	8	8
RRS-7	8	8	8
Rsch-4	0	0	0
Rü3.1-27	0	8	8
Rut-1	0	0	NA

S96	8	0	8
Sap-0	3	8	8
Sau-0	0	8	8
Sco-1	NA	0	NA
SE-0	3	8	8
Se-0	NA	NA	0
Set-1	0	8	8
Sev-1	0	8	0
Sf-1	9	0	8
Sha	0	0	NA
Sha	0	0	NA
Sha-0	0	0	8
SHAKDARA	0	0	0
Si-0	0	0	8
Sis-1	0	8	8
Siz-1	NA	8	NA
Sma-1	NA	8	8
Sna-1	NA	0	NA
Som-1	0	8	8
Sorbo	0	0	8
SORBO	3	8	8
Sou-1	3	8	8
SPR-1-2	0	8	8
SPR-1-6	8	8	8
SQ-1	0	0	0
Sq-1	0	0	0
SQ-8	0	8	8
St-0	0	8	8
Sta-0	0	8	NA
Star-8	3	8	8
Su-1	0	8	NA
Ta-0	3	0	8
Tamm-2	0	0	0
TAMM-2	0	8	0
TAMM-27	0	8	0
Ting-1	8	8	8
Tiv-1	3	0	8
Ts-1	3	0	8
TS-1	3	8	8
TS-5	3	8	8
TSU-1	0	8	0
Tsu-1	0	8	8
Tsu-1	0	NA	8

Tü-SB30-2	3	8	8
Tü-Scha-9	3	8	8
Tü-U-M1	NA	7	8
Tü-V12	0	0	8
Tü-Wa1-2	3	0	8
Ty-0	3	8	NA
Uk-1	0	8	8
Uk-3	0	0	0
ULL-2-3	0	8	8
ULL-2-5	8	8	8
Ull2-5	8	8	8
Unt-1	3	8	8
Uod-1	0	8	8
UOD-1	3	8	8
Uod-7	0	0	NA
UOD-7	0	8	8
VAN-0	0	8	8
Van-0	0	8	8
VAR-2-1	7	8	8
VAR-2-6	0	8	8
Vash-1	0	0	0
Vie-0	0	8	8
WA-1	9	8	8
Wal-HäsB-4	8	8	8
WEI-0	0	0	0
Wen-1	3	8	8
Wen-2	0	0	0
Wig-1	0	0	NA
Wim	3	NA	0
Wis-1	0	0	8
WS-0	8	8	8
Ws-0	8	NA	8
WS-2	NA	8	8
Wt-5	0	0	8
WT-5	0	8	8
Xan-1	0	8	0
Yeg-1	0	8	8
YO-0	0	0	0
Yo-0	0	NA	8
ZDR-1	0	8	8
Zdr-1	0	8	8
ZDR-6	0	8	8
Zdr-6	0	NA	NA