Downy mildew resistance in *B. oleracea*

Abraham Gómez Gutiérrez

Thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

John Innes Centre
2022

Abstract

Downy mildew (*Hyaloperonospora brassicae*) is one of the most devastating and widely reported diseases on horticultural and agricultural *Brassica* plants worldwide. Yield losses can be up to 90-100% if environmental conditions are favourable for downy mildew, especially at the early stages of plant development. In recent years, downy mildew incidence has increased in many *Brassica* growing countries, most likely due to the steady rise of seasonal temperatures in some of these regions of the world caused by climate change, making the disease more widespread. To palliate this, cultural and chemical techniques for controlling downy mildew have been reported to be ineffective due to the systemic nature of the disease. As a result, *Brassica* cultivars with inherent genetic resistance to this pathogen can provide the best and most cost-effective alternative to avoid yield losses in the future.

To achieve this, in this thesis, I have created a collection of single spore *H. brassicae* isolates from 9 field isolates collected from distinct geographical locations in the UK and USA (Chapter 2). In Chapter 3, I assessed different sources of plant material at the cotyledon stage and identified new sources of resistance to *H. brassicae*. The host specificity of all *H. brassicae* isolates used in this thesis was very high, only being able to infect *B. oleracea* and a very few *B. napus* genotypes (Chapter 3). In Chapter 4, a specific F_2 cross-segregating for a single dominant downy mildew resistant gene at the seedling and adult plant stage was used for genetic mapping using a bulk

segregant analysis approach. The bulk segregant analysis was able to map the downy mildew resistance to an NLR cluster in the *B. oleracea* chromosome 2. The NLR cluster was observed to have a high level of copy number variation when compared to other genome assemblies. Therefore, to resolve the NLR cluster, a draft genome assembly of the resistant parental line was created using long-read sequencing. I used the assembly to create diagnostic KASP markers of the resistance and validate them using 466 F_2 individuals for marker-assisted selection.

These newly identified sources of downy mildew resistance could provide a solid foundation to improve the management of this disease in the fields of many *B. oleracea* crops, especially vegetable crucifers. To date, only 4 sources of resistance to downy mildew have been mapped in *B. oleracea*. The new source of resistance located to the *B. oleracea* Chromosome 2 and the diagnostic KASP markers created in this thesis can be used in marker-assisted breeding selection to develop new commercial varieties resistant to downy mildew. Finally, the data generated in this thesis of the NLR cluster and sequences of the candidate genes based on the long-read draft genome assembly could be used to clone the gene responsible for the resistance.

List of Contents

List of Figures

List of Tables

“Show a human a closed door, and no matter how many open doors she finds, she’ll be haunted by what might be behind it.”

— James S.A. Corey

Acknowledgements

This PhD thesis would not have been possible without the support of my family, friends, colleagues and of course, my two cats.

Firstly, I would like to express my gratitude to my supervisor, Chris Ridout, for the opportunity to do this PhD and for his trust in me, even when I was unsure of my abilities. Besides my supervisor, I would like to thank Kirsty Windle and specially Sujit Tha Shrestha at CN Seeds for teaching me how to work with downy mildew and how to pollinate *Brassica* plants early on my PhD journey.

I would also like to thank my fellow labmates and colleagues for all the stimulating talks, the sleepless nights we spent working to meet deadlines and all the fun we have had over the last four years. I want to thank Henk-jan Schoonbeek for taking the time to talk about my experiment ideas and Rachel Burns for helping me with the downy mildew disease assays and making them more enjoyable. I am very glad to have shared an office with two wonderful scientists, Catherine Jacott and Hicret Aslı YALÇIN, thank you for all the conversations and coffee breaks we had. Especially to Aslı for helping me with the bulk segregant analysis.

My sincere thanks also go to Burkhard Steuernagel and Ricardo Humberto Ramírez González, who gave me invaluable assistance when I was starting to analyse my sequencing data and I didn't know where to start.

I can't name here all the good friends I have made over the last four year, but you know who you are, so thank you for being there in the good and the bad. A mis amigos en España, a mi familia y especialmente a mi madre, mi padre y mi hermana, por su amor incondicional y por darme la oportunidad con su esfuerzo y apoyo de poder estar aqui haciendo este doctorado.

Lastly, I am very thankful to the BBSRC for funding my PhD project and to the John Innes Centre for hosting it.

Chapter 1

Introduction

Downy mildew (*Hyaloperonospora brassicae*) is a very common and devastating disease in many *Brassica* growing countries (Carlier et al., 2012; Farinhó et al., 2007), resulting in huge yield losses of up to 90-100% if climatic situations are favourable for downy mildew (Butler, 1918; Monot et al., 2010). In recent years, downy mildew incidence has increased in many brassica crops, especially at the seedling stage due to a steady rise in seasonal temperatures in many *Brassica* growing countries as a result of climate change. To palliate this, cultural and chemical techniques for controlling downy mildew have been reported to be ineffective due to the systemic nature of the disease (Vicente et al., 2012). Fungicides were likewise ineffective in preventing downy mildew disease, which lead to the emergence of new more virulent isolates (Brophy & Laing, 1992; Singh et al., 2013). As a result, *Brassica* cultivars with inherent genetic resistance to this pathogen can provide the best and most cost-effective alternative to avoid yield losses in the future. For this purpose identification of new sources of resistance and knowledge of the genetics of resistance are prerequisites to develop new, durable and effective downy mildew resistant varieties.

1.1 Brassicaceae family

1.1.1 Economic Importance

The Brassicaceae family, previously known as Cruciferae currently consist of 372 genera, and 4.060 species (*The Plant List. Version 1.*, 2010). This family includes some of the world's most economically important crops, particularly the members of the *Brassica* genus which include cabbage, oilseed rape, cauliflower, broccoli, kale, collards, mustards, turnips, Chinese cabbages, etc; *Raphanus* L. (radish), *Armoracia* G.Gaertn., B.Mey. & Scherb. (horseradish), *Eutrema* R. Br. (wasabi), *Eruca* Mill. (rocket), *Lepidium* Fabr. (garden cress) and *Nasturtium* W. T. Aiton (watercress) (Al-Shehbaz, 2012). In food markets, the main uses of these *Brassica* cultivars are oilseed, forage, condiment and leafy vegetables, by using seeds, stems, roots, buds, inflorescences and leaves. *Brassica* vegetables are a staple food in many countries around the world and the oil produced from brassica oilseed crops represents 14% of the world's vegetable oil, making it the third most important supply of edible oil after soybeans and palm oil (OECD, 2012). The principal vegetable market species is *B. oleracea*, also referred to as "cole crops", which includes vegetable and forage cultivars like kale, cabbage, broccoli, Brussels sprouts, kohlrabi and collard greens. *B. rapa* is also another source of vegetable and forage cultivars like turnip, Pak Choi and Chinese cabbage. *B. napus* cultivars (oilseed rape, OSR) are mainly used for oil production, but on a smaller scale, they are also used as a vegetable. *B. nigra*, *B. carinata* and *B. juncea* are included together in the "mustard group" and are mainly used as a condiment (OECD, 2012; Sadowski & Kole, 2016).

All the world's commercial *Brassica* products are subdivided into two categories (cabbages and other brassicas and cauliflowers and broccoli), in the

FAO database (FAOSTAT, 2020). In 2020 the total production for these two categories combined was 96.4 million tonnes from 3.8 million hectares. By region, Asia holds the highest brassica production with 74.9 Mt, followed by Europe with 12 Mt and America and Africa with 4.7 and 4.3 Mt respectively. Since 1961 the total worldwide vegetable *Brassica* production has increased five times. In terms of *Brassica* consumption, around 88% of the production goes for fresh human consumption, 1.2% is processed, 4% is used for animal feeding and the rest is considered waste (FAOSTAT, 2009).

Import and export FAO data estimated that in 2020 the total exported *Brassicas* in the world was approximately 4.119.069 tonnes, representing a value of approximately 3,557 million USD (Figure 1.1).

The annual growth worldwide for both brassica FAO groups production and export/imports has been gradually increasing from 1961 until 2020 (Figure 1.1, Figure 1.2). Therefore, to meet the growing demand and reduce environmental impacts, varieties with beneficial marketable traits will be needed by farmers.

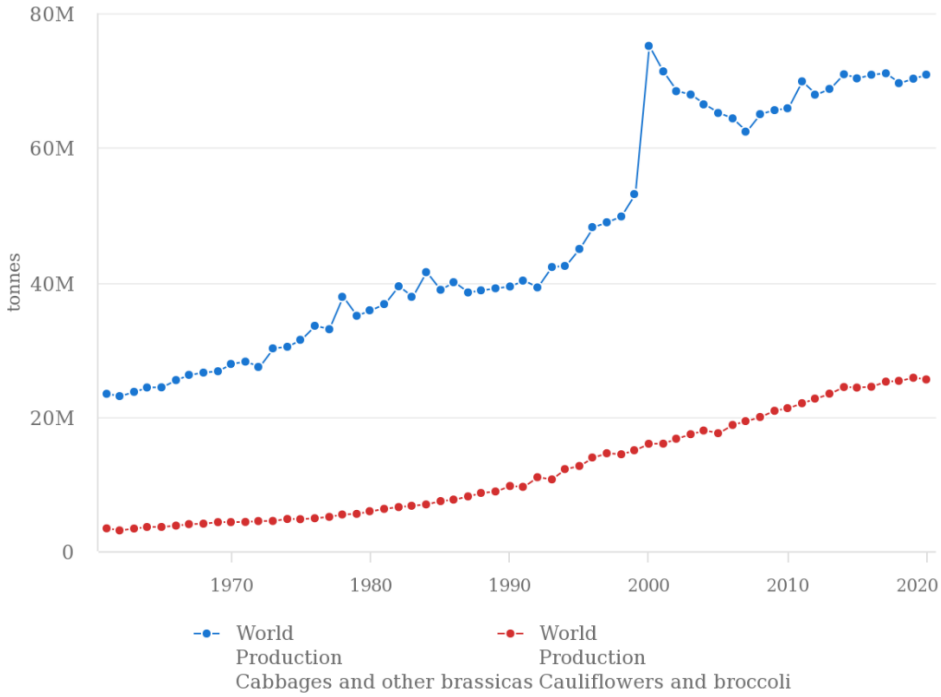


Figure 1.1. Worldwide annual *Brassica* production (Cabbages and other *Brassicas* & Cauliflowers and broccoli) from 1961 to 2020 (FAOSTAT, 2020). Data is shown as a million tonnes produced per year.

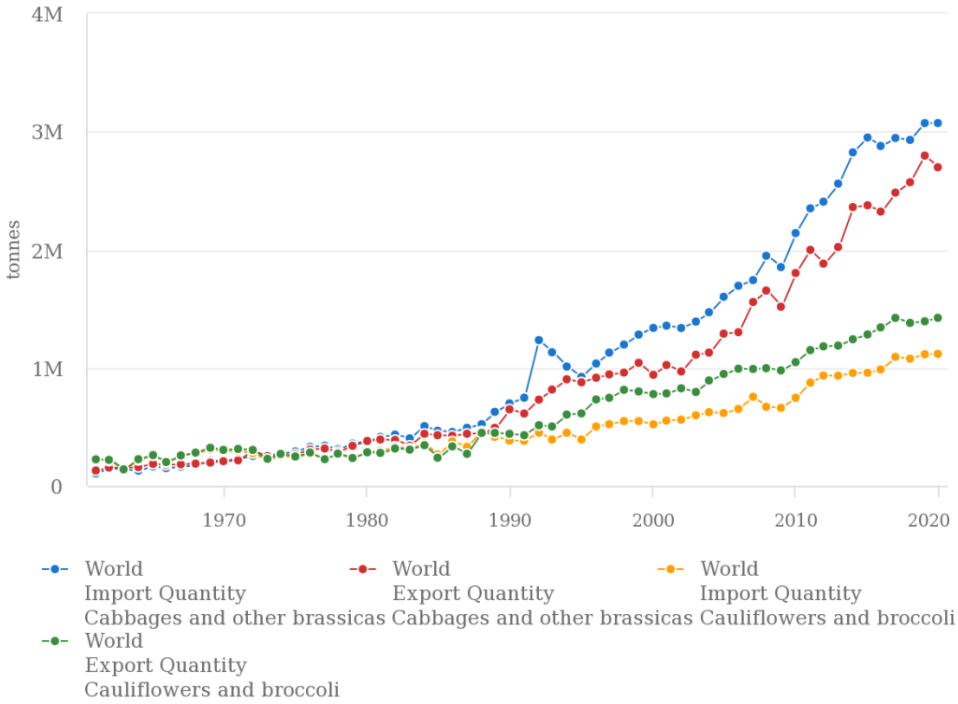


Figure 1.2. Worldwide annual imported and exported *Brassicas* (Cabbages and other *Brassicas* & Cauliflowers and broccoli) from 1961 to 2020 (FAOSTAT, 2020). Data is shown as million tonnes produced per year.

1.1.2 *Brassica* genetics and resources

As previously mentioned, the Brassicaceae family includes many important crop species. Of all the cultivated species included in the family, almost all belong to six species, *Brassica rapa*, *Brassica nigra*, *Brassica oleracea*, *Brassica juncea*, *Brassica carinata* and *Brassica napus*. The relationship between these six cultivated *Brassica* crops was investigated and clarified by inter-specific hybridization experiments, creating the well-known and highly accepted "triangle of U" model (Nagaharu, 1935); (Figure 1.3). Three of these species are diploid (*Brassica rapa* [n=10; **A**], *Brassica nigra* [n=8; **B**] and *Brassica oleracea* [n=9; **C**]), and the other three, intercrossed with each other to create the three polyploid species (*Brassica juncea* [n=18; **AB**], *Brassica napus* [n=19; **AC**] and *Brassica carinata* [n=17; **BC**]). Compared to the model organism *Arabidopsis thaliana*, *Brassica* species underwent an extra whole-genome triplication (WGT) event, which occurred around 9-15 (Beilstein et al., 2010; Wang et al., 2011) to 28 million years ago (Arias et al., 2014; Lukens et al., 2004).

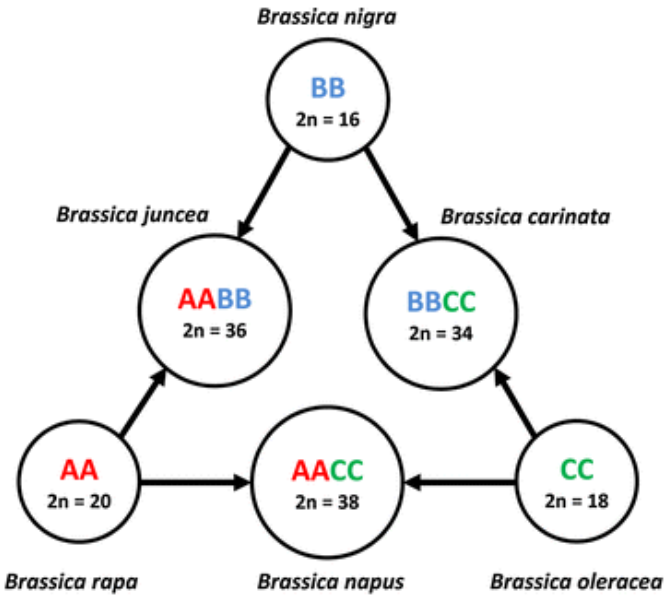


Figure 1.3. *Brassica* species within the triangle of U (Nagaharu, 1935)

Recent developments in genome sequencing and assembly techniques have enabled the scientific community to access the genetic resources of various crop species and wild relatives, in which many brassica crop species are included (Bevan et al., 2017). To date, several genome assemblies are available for the three diploid *Brassica* species, *B. oleracea* (Belser et al., 2018; Golicz et al., 2016; Liu et al., 2014; Lv et al., 2020; Parkin et al., 2014), *B. rapa* (Belser et al., 2018; Cai et al., 2017; Cai et al., 2021; Li et al., 2021; Li et al., 2020; Wang et al., 2011; L. Zhang et al., 2018) and *B. nigra* (Paritosh et al., 2020; Paritosh et al., 2021; Yang et al., 2016).

1.2 Downy Mildew in Brassicas

1.2.1 *Hyaloperonospora brassicae* classification

Downy Mildew (Downy Mildew) is a worldwide plant disease caused by the obligate biotrophic oomycete *Hyaloperonospora brassicae* (Gäumann) Göker, Voglmayr, Riethmüller, Wei and Oberwinkler (syn. *Hyaloperonospora parasitica* (Pers. ex Fr.) Constant. and *Peronospora parasitica*) (Göker *et al.,* 2009). In the past, Downy Mildews were considered to be part of the fungi kingdom due to their morphology, physiology and ecology, but molecular and biochemical studies demonstrated that they belong to the kingdom Chromista (Dick, 2001, 2002). The first to describe downy mildew on a brassica species (*Capsella bursa-pastoris*) was Persoon (1796), considering it as a fungus and identifying it as *Botrytis parasitica* Pers. Fries (1849) classified the pathogen in the *Peronospora* genus, which was introduced by Corda in his description of *Peronospora ramicis* in 1837 (Corda, 1837). After this, all isolates obtained from Brassicaceae plants were assigned to *Peronospora parasitica* (Pers. ex. Fr.). Gaumann (1918) started naming downy mildew isolates obtained from *Brassica* species as *Peronospora brassicae* Gaum, due to

the narrow host specialization observed at the species level. After this point, researchers have been constantly reviewing and changing the taxonomy of downy mildew members. Constantinescu and Fatehi (2002) divided the genus *Peronospora* into three genera, *Peronospora* s. str., *Hyaloperonospora* and *Perofascia*. At the present time, based on molecular, and morphological features and also, ITS 1, ITS 2 and 5.8 S rDNA sequence and phylogenetic analysis, *Hyaloperonospora* was confirmed to be a pathogen-specific to brassica crops. *H. brassicae* to Brassicaceae species, *H. arabidopsidis* to *Arabidopsis thaliana* and *H. parasitica* to *Capsella bursa-pastoris* (Göker et al., 2004; Göker et al., 2009; Göker et al., 2007; Göker et al., 2003; Young-Joon et al., 2003). Curiously, compared to *Peronospora*, *Hyaloperonospora* and *Perosfacia* seem to be highly restricted to the Brassicaeae family. With more than 100 species the *Hyaloperonospora* genera is the third largest genus for Downy Mildews infecting many brassica crops. On the other side, the downy mildew specific for the model organism *A. thaliana* (*H. arabidopsidis*) has become a plant pathogen model organism to investigate plant pathogen interactions (Baxter et al., 2010; Coates & Beynon, 2010). After the manuscript published by Constantinescu (1991), even though new monophyletic groups have been found in different phylogenetic analyses (Göker et al., 2009; Thines & Choi, 2016), very few new species have been discovered and published (Voglmayr et al., 2014). In the next chapters of this thesis, the term downy mildew will be used to refer to *Hyaloperonospora brassicae* (Gaum.) Goker.

1.2.2 Downy Mildew host specificity

As previously noted, the species concept for downy mildew has been a very contentious subject. The fundamental issues with downy mildew species delimitation are that one, they have a wide range of plant host species, two, they are morphologically highly similar and three, most of them have very

confined host ranges due to their biotrophic nature, therefore could represent genetically different species. Molecular phylogenetic studies on Downy Mildews have reported a high level of specialization for both, the *Peronospora* (Belbahri et al., 2005; Choi et al., 2010; Choi et al., 2008; Choi et al., 2007; Choi et al., 2015; Choi et al., 2009; Thines et al., 2009) and *Hyaloperonospora* (Goker et al., 2009; Voglmayr et al., 2014) genera. Disease downy mildew assessments on brassica plants also indicate that they are often very host-specific (Byford, 1967; Gaumann, 1918; Gaumann, 1923; Lebeda & Syrovátko, 1988; Sherriff & Lucas, 1990), and only a few exceptions exist where a downy mildew species appear to have a wide range of hosts (Kenneth, 1981; Runge et al., 2011). The phylogenetic analysis performed by Göker et al. (2009) revealed that downy mildew species could be easily interpreted biologically in terms of host specialization, supporting the concept of confined species that are highly specialized to a few or a single closely related host species (Constantinescu & Fatehi, 2002; Göker et al., 2004; Göker et al., 2003; Young-Joon et al., 2003). Therefore, even if they are morphologically indistinguishable, some authors proposed that downy mildew collections with distinctive DNA sequences and host ranges should be recognized as independent species (García-Blázquez et al., 2008; Göker, 2006; Göker et al., 2004; Göker et al., 2009; Göker et al., 2003). So, a combination of molecular and host traits could be a reliable method to distinguish downy mildew species (Göker et al., 2009).

1.2.3 Phylogenetic analyses

The species classification for plant pathogens is especially important for practical reasons, such as plant protection, and biological control as well as playing a very important role in comparative research involving model organisms and crops. Nonetheless, defining species boundaries remains a

very challenging taxonomic problem in numerous groups. As previously noted, this challenge is prominently observed in the obligate biotrophic downy mildew genera, which exhibit substantial variations in terms of genetic divergence and host specificity. Despite this genetic diversity, the downy mildew genera are, for the most part, characterized by relatively uniform morphological traits.

Although very few studies have been performed to solve this issue, to date, the most comprehensive molecular study of the downy mildew genus is the one performed by Göker et al. (2009), exploring the phylogenetic relationships within the genus *Hyaloperonospora*. This study used a wide array of samples from multiple host species and geographical locations and performed a molecular analysis of the nuclear ribosomal internal transcribed spacer (ITS) and large subunit (LSU) sequences. Utilizing *Perofascia* as an outgroup, *Hyaloperonospora* was strongly supported as a monophyletic group with a 100% bootstrap support (BS) value (Figure 1.4). Although sequences were obtained from multiple *Lepidium* host species, *Perofascia lepidii* emerges as a genetically uniform single species. Within Hyaloperonospora, six distinct clades were identified, each of them containing more than one species and supported by high BS values, in addition to several isolated species.

Clade 1 acts as a sister group to a clade encompassing the remaining five clades and all single species lineages, with both low (65%) and high (98%) BS values in maximum likelihood (ML) and maximum parsimony (MP) analyses, respectively (Figure 1.4). The central structure connecting these groups exhibits low resolution compared to the well-supported terminal nodes. This clade also includes six unique lineages that parasitize various *Cardamine* or *Rorippa* species from Europe and East Asia, as well as

European *Barbarea vulgaris* and *Arabis soyeri*. All host genera, except the latter two, belong to Cardamineae.

Clade 2 consists of samples from European *Draba verna* and East Asian *Draba nemorosa*, forming two separate clusters with significant genetic differences. *Hyaloperonospora arabidis-alpiniae* (parasitic on Arabis alpina) diverges before Clade 2, and *Hyaloperonospora niessleana* after Clade 2, both without support in the phylogenetic tree. *Hyaloperonospora isatidis*, *Hyaloperonospora thlaspeos-arvensis*, and the parasite of *Biscutella auriculata* are closely related to Clade 3, but also lack support.

Clade 3 contains parasites of a diverse array of hosts, such as *Reseda*, one of the few non-Brassicaceae host genera. The basal split within Clade 3 between the East Asian *Sisymbrium luteum* pathogen and the other specimens is moderate to highly supported (87 to 95% BS). Subclades within Clade 3 are challenging to delineate, and it is unclear if the parasite of *Arabis turrita* belongs to *Hyaloperonospora arabidopsidis* or if *Hyaloperonospora thlaspeos-perfoliati* includes specimens from *Noccaea caerulescens*. Isolated lineages with high monophyletic support (Figure 1.5), yet unresolved relationships, include *Hyaloperonospora cheiranthi*, *Hyaloperonospora hesperidis*, *Hyaloperonospora nesliae*, *Hyaloperonospora camelinae*, and *Hyaloperonospora parasitica*, along with pathogens of *Lepidium draba*, *Descurainia sophia*, *Iberis sempervirens*, and *Helianthemum* (Cistaceae). Clade 4 consists of *Hyaloperonospora galligena* and *Hyaloperonospora berteroae*, while Clade 5 includes two distinct lineages of *Sisymbrium* parasites.

Finally, Clade 6 has a notably large number of specimens and is split into two subclades: those primarily parasitizing *Cardaminae* (Cardamine and relatives like Nasturtium) and those primarily parasitizing *Brassica* and

relatives (tribe Brassiceae; *Diplotaxis*, *Eruca*, *Erucastrum*, *Raphanus*, *Sinapis*). This division is strongly supported by both ML and MP analyses. Within the first subclade, differentiation between samples from North America (specimen 21-01), East Asia (GenBank sequences), and Europe is evident. The second subclade also includes parasites of *Armoracia*, *Lobularia*, *Lunaria*, *Sisymbrium*, *Teesdalia*, and *Tribulus* (Zygophyllaceae). Parasites of *Diplotaxis* and *Sinapis* occur within at least two lineages. Both *Sinapis* and *Brassica* downy mildew appear to infect *Armoracia* as well. The classification of *Hyaloperonospora lunariae* is ambiguous, as it cannot be clearly distinguished from the downy mildew affecting *Erucastrum nasturtiifolium*.

In addition to the findings by Göker et al. (2009), a more recent study by Lee et al. (2017) included additional samples from Korea and made a phylogenetic analysis based on internal transcribed spacer (ITS) rDNA but also on cox2 mtDNA sequences.

The phylogenetic analyses performed by Lee et al. (2017) using samples collected in Koread (Figure 1.6 and Figure 1.7) revealed that accessions from different *Arabis* species fell into four distinct lineages: one representing *H. arabidis-alpinae*, another representing *H. arabidis-turritae*, a third on *Arabis soyeri* (previously described by Göker et al. (2009)), and a fourth clade with two Korean accessions on *Arabis glabra*, representing a new combination, *H. arabidis-glabrae*. An unknown *Hyaloperonospora* lineage from *Catolobus pendulus* (= *Arabis pendula*) clustered with *Hyaloperonospora arabidis-turritae*, *Hyaloperonospora mayli*, and *Hyaloperonospora thlaspeos-perfoliati* and may represent an unknown species.

Accessions from *Rorippa* species formed two major clades (Figure 1.6): one representing *Hyaloperonospora nasturtii-montani*, and another comprising

Hyaloperonospora nasturtii-islandicae and *Hyaloperonospora nasturtii-palustris*. The latter two species, parasitic on *Rorippa palustris*, formed a highly supported group (91/98%), but each represented a distinct *Hyaloperonospora* species.

The *Hyaloperonospora brassicae* complex fell into four related clades in the ITS tree, as suggested by Göker et al. (2009): one clade parasitic on *Brassica* species, another on *Sinapis alba*, a third on *Sinapis arvense*, and a fourth on *Raphanus* species. Intriguingly, accessions from spider flower, *Tarenaya hassleriana* (= *Cleome spinosa*), were placed within the first *Hyaloperonospora brassicae* clade, indicating a close genetic relationship based on ITS and cox2 sequences.

Korean accessions from Cardamine impatiens grouped within the *Hyaloperonospora dentariae* clade, albeit with weak support in the ITS dataset and strong support in the cox2 mtDNA dataset. They formed a well-supported subclade in the cox2 dataset, consistent with Voglmayr et al. (2014), who observed host-specific substructures within *Hyaloperonospora dentariae*. Notably, there was a divergence between European and Korean accessions from *Cardamine impatiens*. An unknown *Hyaloperonospora* lineage from *Cardamine scutata* was found to be closely related to *Hyaloperonospora* sp. ex *Arabis soyeri* and may represent an overlooked species. Korean accessions from *Thlaspi arvense* grouped with GenBank sequences of *Hyaloperonospora thlaspeos-arvensis*, marking the first report of *Hyaloperonospora dentariae* and *Hyaloperonospora thlaspeos-arvensis* in Korea.

Overall, the phylogenetic analysis of *Hyaloperonospora* reveals a complex picture with multiple clades and lineages, each exhibiting diverse host specificity and geographic distribution. These studies underscore the challenges

in delineating species boundaries and emphasize the need for continued research to better understand the evolutionary relationships within this genus.

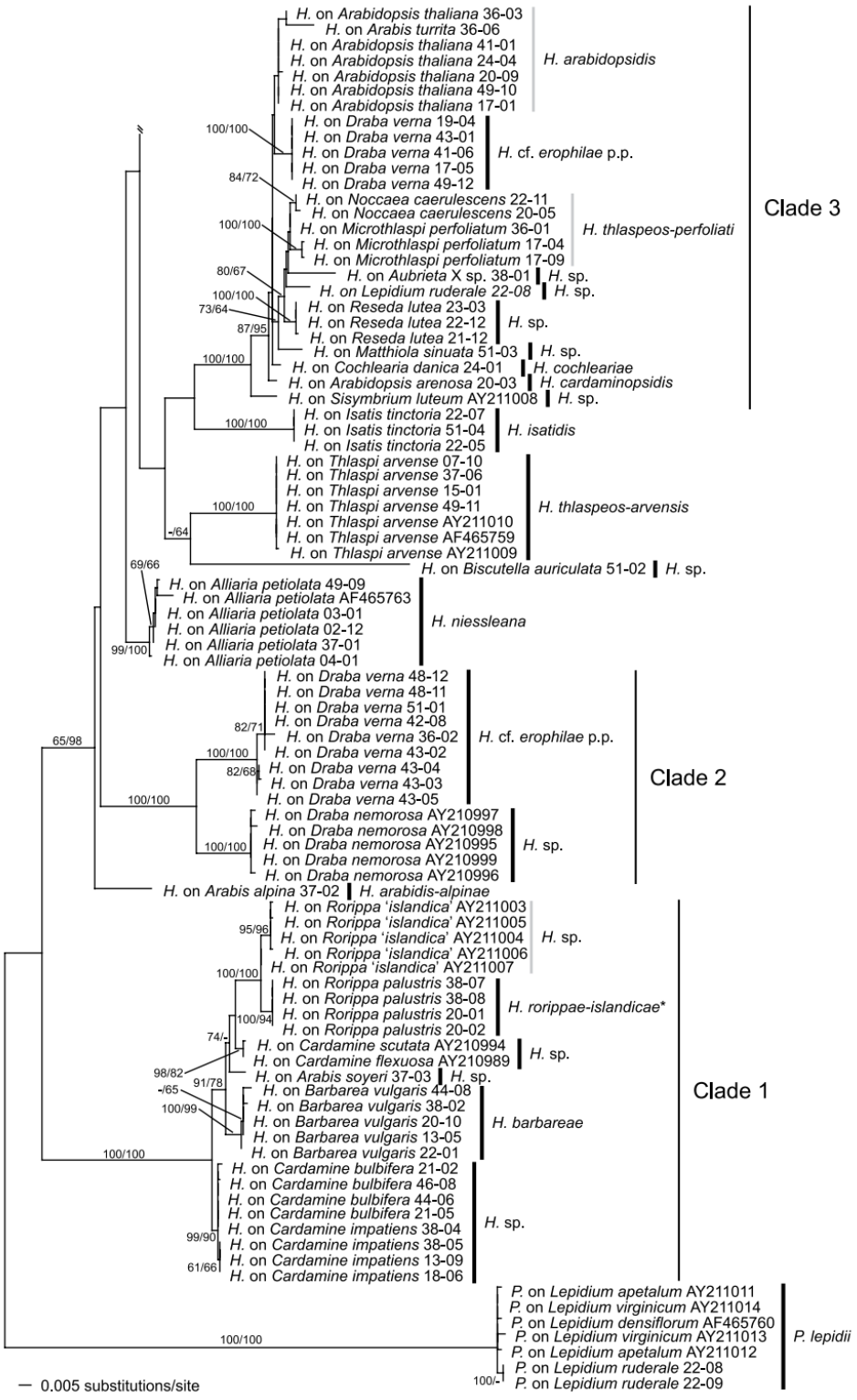


Figure 1.4. The lower portion of the phylogenetic tree, generated using RAxML with a GTRMIX nucleotide substitution model approximation, was rooted with Perofascia and is presented with branch lengths scaled to reflect the expected number of substitutions per site Bootstrap support (BS) values above 50% from both maximum likelihood (ML, left) and maximum parsimony (MP, right) bootstrapping are denoted above the branches. Specimens of *Hyaloperonospora* (H.) and

Perofascia (P.) are labelled with their DNA isolation numbers and host species, and GenBank accession numbers are provided where sequences were sourced from the database. Thick vertical bars, along with accompanying names, represent the proposed species names, with asterisks indicating binomials introduced in this study. In instances where species boundaries are uncertain, the bars are depicted in light grey. Thin vertical bars and corresponding numbers identify clades that appear to be above the species level, as further discussed in the text. The upper portion of this tree can be found in Figure 1.5 (Göker et al., 2009).

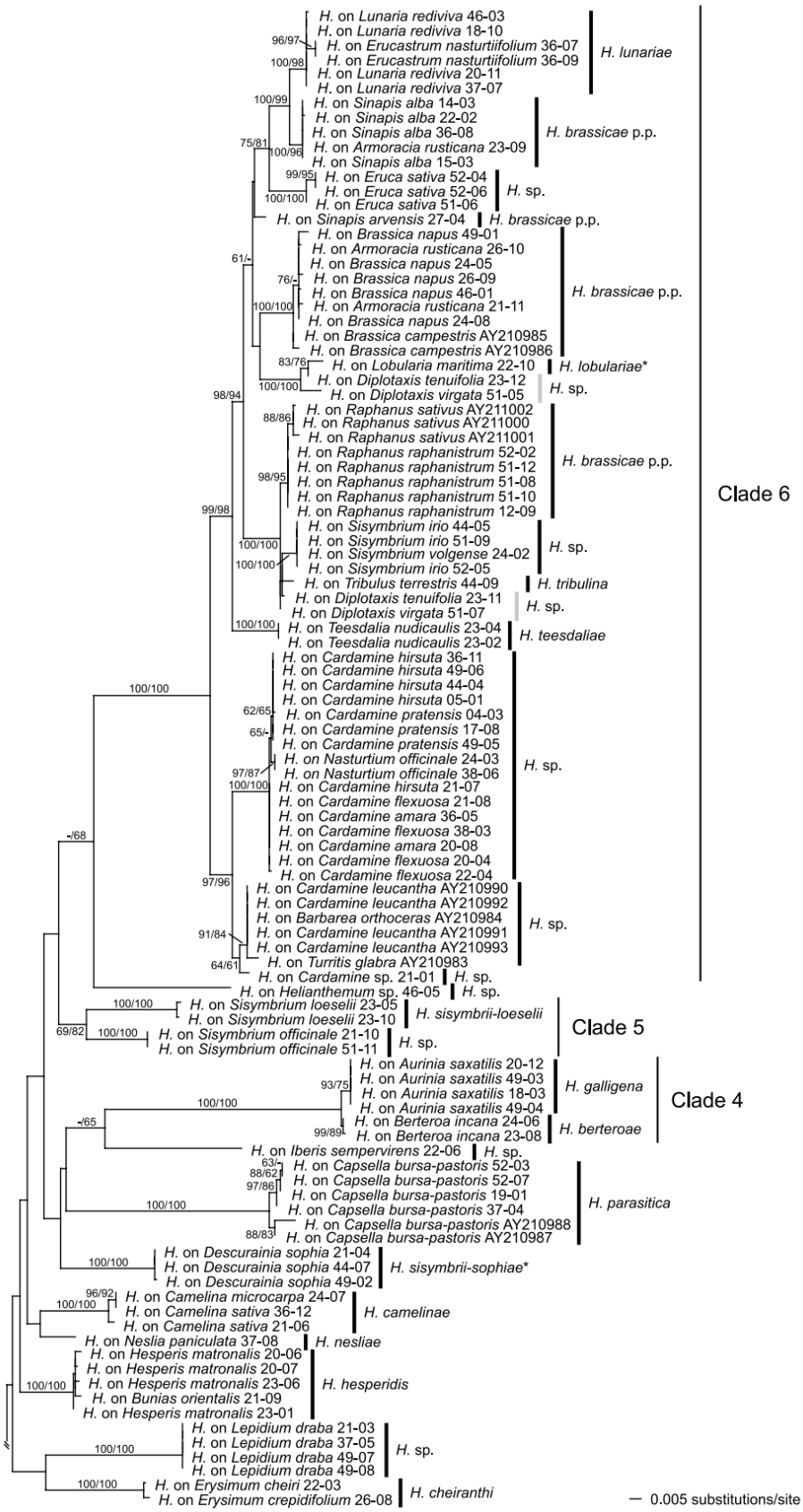


Figure 1.5. Upper half portion of the phylogenetic tree from Figure 1.4, including clades 4-6. For a description, see Figure 1.4 (Göker et al., 2009).

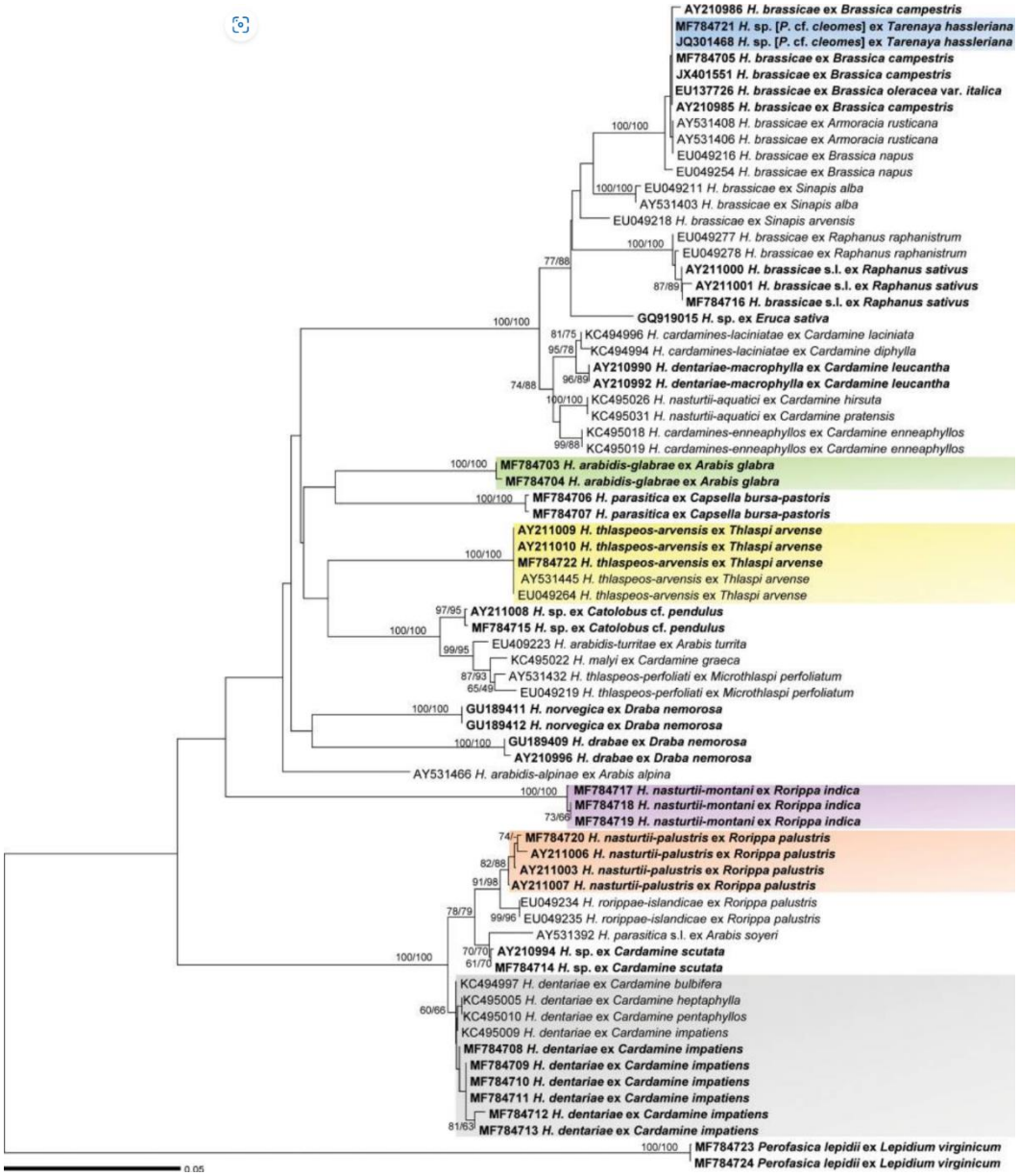


Figure 1.6. Minimum evolution trees, generated using the internal transcribed spacer (ITS) rDNA sequences, are presented with support values obtained from maximum likelihood inference. Bootstrap support values greater than 60% are displayed above/below the branches, indicating the level of confidence for each clade. Specimens collected in Korea are highlighted in bold text, and species that were previously unknown in Korea are emphasized with coloured boxes. The scale bar represents the number of nucleotide substitutions per site, providing a visual representation of the genetic distance between taxa within the tree (Lee et al., 2017).

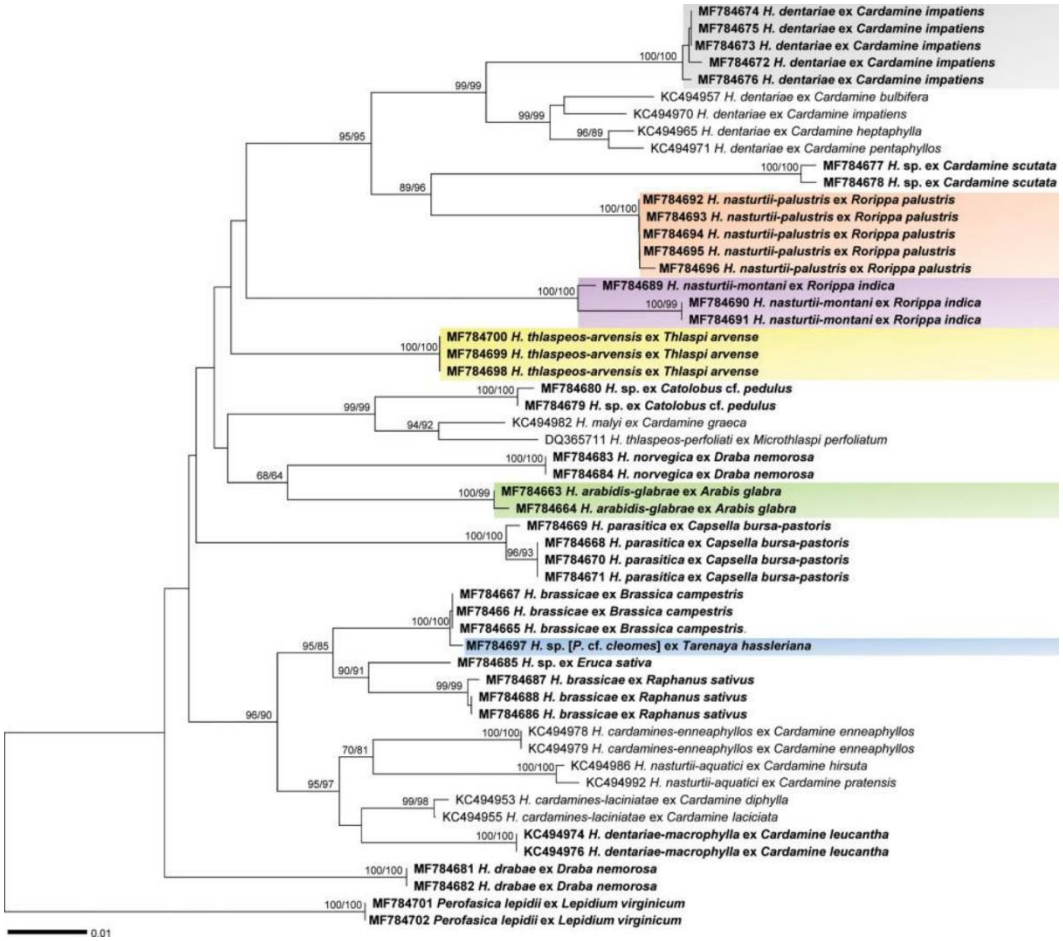


Figure 1.7. Minimum evolution trees, generated using the cox2 mitochondrial DNA (mtDNA) sequences, are presented with support values obtained from Bayesian inference. Bootstrap support values, derived from both minimum evolution and maximum likelihood analyses, are displayed above/below the branches for values greater than 60%, indicating the degree of confidence associated with each clade. Specimens that were collected in Korea are emphasized in bold text, and species that were previously unidentified in Korea are accentuated within coloured boxes. The scale bar represents the number of nucleotide substitutions per site, visually illustrating the genetic divergence between taxa within the tree (Lee et al., 2017).

1.2.4 *Hyaloperonospora brassicae* life cycle

Downy mildews are a group of obligate parasitic pathogens that are very specific to the host species they infect. To multiply and progress in their life cycle, downy mildews require specialized nutrients that are created or become available during pathogenesis (S. Kluczewski & J. Lucas, 1982; Singh et al., 1980). On Brassicaceae adult plants downy mildew is primarily a foliar disease but the infection may extend to other parts of the plant like the inflorescence of cauliflower and broccoli cultivars. Compared to adult plants, seedlings and young plants are more vulnerable to this disease (Koike et al., 2009), especially in plant nurseries where they could be very destructive (Channon, 1981). In the field, new downy mildew infections usually occur via sexually produced oospores stored over winter in crop residues in the soil (LeBeau, 1945), but after this initial infection successive colonization takes place via asexually produced spores formed on the cotyledons, hypocotyls and young leaves and spread by wind or water droplets to other uninfected parts of the plants (Chang et al., 1963).

1.2.4.1 Asexual reproduction

When environmental conditions are favourable (10-15°C and high relative humidity), conidiophores germinate, developing a short germ tube and a haustorium that penetrates the surface of the cotyledon/leaf/inflorescence at the anticlinal juncture of two epidermal cells or occasionally, through the stomata (Shiraishi et al., 1975). Following penetration, the hypha continues growing, branching out in all directions and changing in diameter and form according to the size of the space between the cells of the host plant tissues (Chou, 1970). After the mycelium has completed its vegetative growth, conidiophores emerge vertically, singly or in groups, through stomata on the

abaxial side of the cotyledons and leaves of the host plant, primarily during the night (Figure 1.8). The entire process takes about 4-6 hours from emergence to spore formation (Davison, 1968). The intercellular hyphae differentiate to form the conidiophores which are white/grey hyaline and measure 200-500 μm in length. Conidiophores can branch 3-8 times which bifurcate to form pairs of long, slender, pointed sterigmata, ending in a single conidium (Figure 1.8). Conidia are hyaline, elliptic or globose in shape and measure 24-27 by 12-23 μm and are delimited from the sterigmata by cross walls at maturity.

Disease intensity can be increased using concentrations of up to $3x10^4$ spores/mL, but increasing it further have little or no effect (Achar, 1992). In the fields, sporulation takes place mostly at night, disseminating the spores when the sporangia start to dry out in the morning, producing a violent twist that releases the spores into the air (Pinckard, 1942). In the fields, the average survival time for spores stored on plant debris is up to 10 days on moist soils and warm days but if the soil is dry it can survive for up to 115 days.

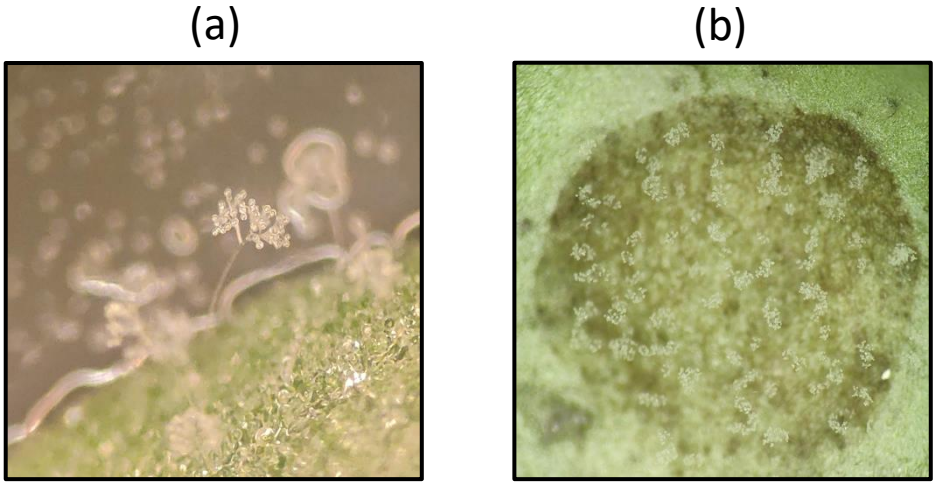


Figure 1.8. (a) Sporangia and hyaline spores of *H. brassicae* (30x magnification). (b) Early sporangia emergence at the point of inoculation on a *B. oleracea* cotyledon.

1.2.4.2 Sexual reproduction

Downy mildew oospores generated in senescent host tissues are a very important mechanism for survival during periods of adversity. To form oospores, downy mildews can reproduce sexually by passing nuclear material from an antheridium to an oogonium (Figure 1.9). Senescence of the host tissues caused by a deficiency of nitrogen or phosphorous favour oospore development (McMeekin, 1960). As a result, oospores can be found more frequently in senescent rather than green healthy tissues and in comparison to leaves oospores were found more frequently in cotyledons because they senesced earlier if the infection starts at early stages of the host development (McMeekin, 1960). Downy mildew oospores can survive in the soil for up to three years (Jang & Safeeulla, 1990). As a result, oospores are a crucial source of survival for downy mildew so after a growing season, oospores stored in plant debris could remain in the soil for many years (Edwards et al., 2011).

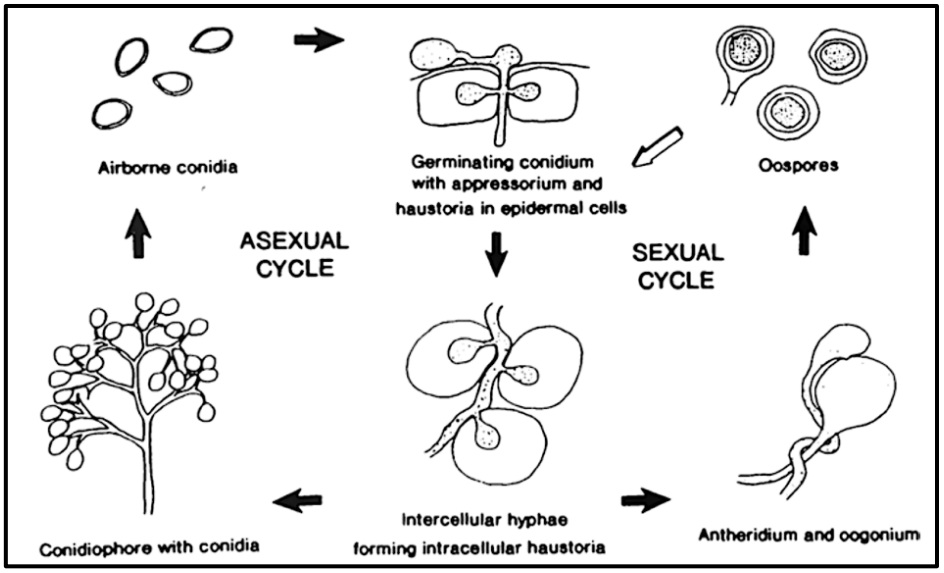


Figure 1.9. Downy mildew life cycle showing asexual and sexual cycles to form spores and oospores respectively on Brassicaceae species (Lucas et al., 1994).

1.2.5 Downy Mildew symptoms on Brassicaceae species

Downy mildew is one of the most widely reported diseases on horticultural and agricultural *Brassica* plants (D. Silue et al., 1996). The disease primarily affects young seedlings, which can be damaged or killed in severe situations. Later stages of infection cause debilitation and a decline in the host plant's quality and performance (D. Silue et al., 1996). After initial infection, the first symptoms include chlorotic small pale yellow spots on the upper surface of the cotyledons and leaves, and white/grey sporangia carrying spores on the underside of the cotyledon and leaf. In the field, the most common visible symptoms on adult brassica plants are black, greyish or brown spots on the surface of leaves, curd heads and radish and turnip heads. The entire undersurface of *Brassica* seedlings can become completely covered by downy mildew sporangia. After infection, downy mildew can create multiple sunken black spots on cotyledons/leaves (Sherf & MacNab, 1986) and also on cauliflower and broccoli curds (Channon, 1981; Chorin, 1946). These dark black spots can also be seen at the point of inoculation on cotyledon and leafs inoculated in the laboratory (Figure 1.10-a). On highly susceptible genotypes, damage caused by downy mildew can make the entire leaf or cotyledon yellow, dry up and drop off the plant. On moderately resistant genotypes, initial spots caused by downy mildew on leaves remain confined to the point of inoculation not being able to extend to other parts of the leaves (Figure 1.10-b). Adult brassica plants have been reported to be more resistant to downy mildew infection compared to cotyledons from the same genotypes (Coelho & Monteiro, 2003a, 2003b; Coelho et al., 2009).

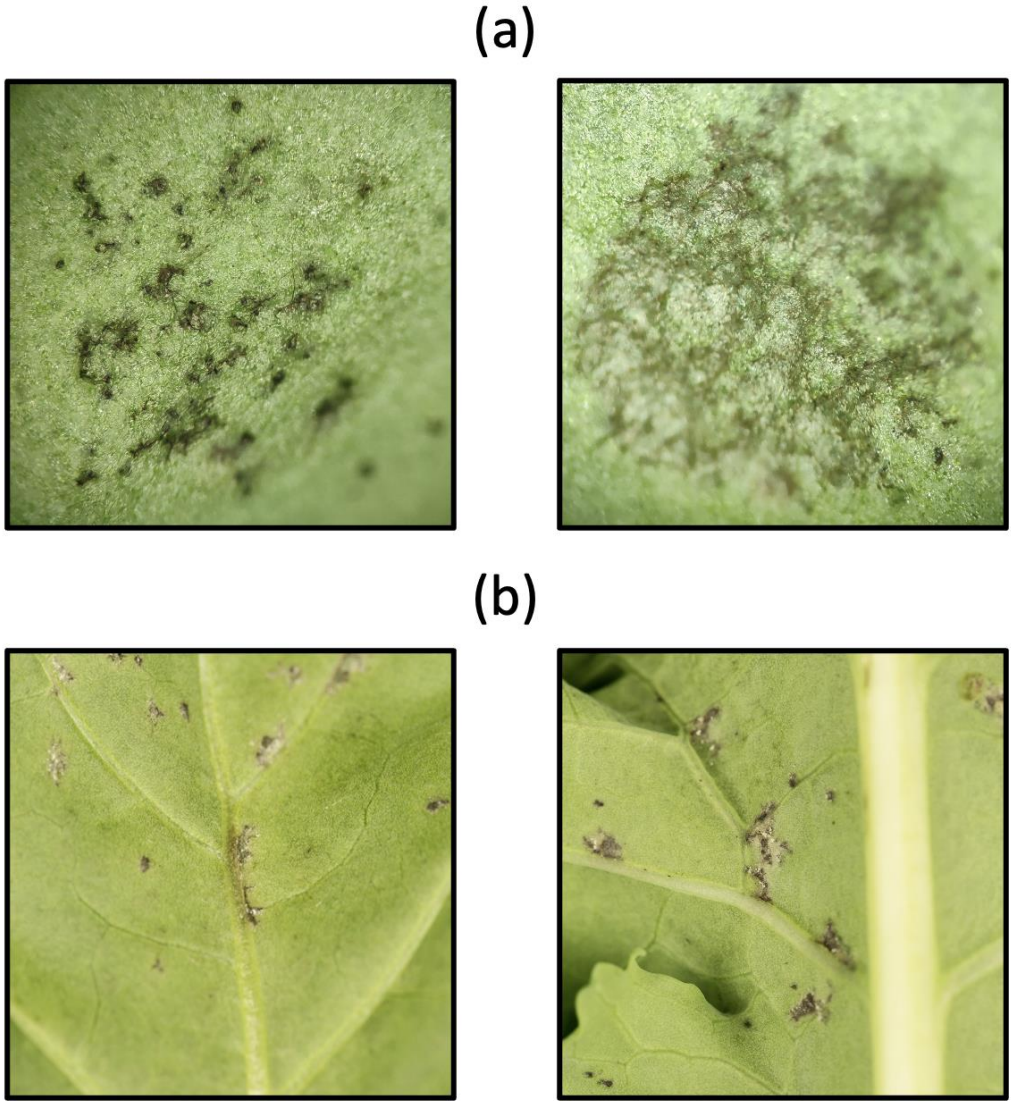


Figure 1.10. (a) Black spots at the Downy Mildew point of inoculation on the abaxial side of *B. oleracea* cotyledons. (b) localized Downy Mildew black spots and sporangia on the abaxial side of *B. oleracea* leaves.

1.2.6 Biotic and abiotic factors affecting downy mildew in the fields and disease management

In the case of downy mildew on Brassiceae species, the interaction between the host plant, the pathogen, and the environment is a complicated process that controls the rate of the disease progression. In addition, downy mildew infections in the field have a relatively short latency period, resulting in a fast disease progression making the epidemics rapid and unpredictable (Jeger & Spencer-Phillips, 2004).

Environmental temperature and relative humidity are the two most important abiotic factors that affect significantly downy mildew development before and after infection. Other environmental variables like wind and rainfall duration and intensity, and dew have an important effect on the downy mildew infection, survival, multiplication and spread as primary and secondary sources of inoculum. In downy mildews, primary inoculum usually starts with a modest amount of pathogen inoculum and then grows exponentially through repeated cycles on the plant host over the growing season. Therefore, the presence or absence of the aforementioned environmental abiotic variables will have a significant impact on the overall severity of the downy mildew infection in the fields, especially on the secondary spread of the pathogen. Cultural practices such as sowing time, nutritional conditions (macro and micronutrients), cultural operations, type of cultivar, crop rotations, continuous harvesting, plant debris burning, and so on have a significant impact on the downy mildew life cycle on Brassicaceae crops. Therefore, since downy mildew oospores remain in the plant debris of the infected host tissue, cultural practices for controlling downy mildew in the fields should focus on sanitation and management by removing and

destroying the infected plant debris to limit the supply of the primary inoculum in the next growing season. In addition to these measures, other cultural practices to make the environment less favourable for downy mildew infection should be applied. For example, increasing the aeration and avoiding dense seeding will reduce the relative humidity around plants (Butler, 1918; Sherf & MacNab, 1986).

Biotic factors like the age of the host tissue, availability of young tissues, host-like proximity to the primary source of inoculum, leaf wetness duration, low wax content on leaves, pathogen virulence, susceptibility of the host sown, etc are all factors that affect the downy mildew disease severity favourably. From all these biotic factors the most cost-effective solution to control downy mildew in the fields is to use naturally resistant genotypes. So the identification and mapping of new sources of resistance should be the priority for brassica breeders

Alternatively, brassica growers use fungicides or other chemical treatments to protect their crops from downy mildew (Monteiro et al., 2005) or to control it to some extent (Brophy & Laing, 1992). Fungicides have been used in a variety of ways, including foliar spraying on affected plants or seed treatment before sowing (Gisi, 2002). Jensen et al. (1999) reported that the sporulation intensity of downy mildew was significantly reduced on brassica seedlings by using CGA245704 (systemic acquired resistance chemical activator) as a seed treatment. Seed treatment with metalaxyl was also effective in controlling downy mildew in cauliflower (Verma et al., 1994). However, downy mildews have the capability to evolve and become resistant to fungicide treatments (Brophy & Laing, 1992; Tyler, 2007) including the previously mentioned metalaxyl (Crute et al., 1985). As with any other biological system, no single strategy can be viable, effective, ecologically safe or

economically feasible. Fungicides and cultural practices have only provided limited control to downy mildew in the fields (Barbetti et al., 2011). Also, fungicides and other chemicals cannot be used regularly in certain brassica crops, like crops with short periods between sowing and harvesting and or ecologically branded, like baby leaves. Therefore the use of resistant genotypes is the most cost-effective way to control downy mildew in the fields, and for this reason, this was the main research focus of this thesis.

1.3 Sources of resistance to downy mildew in brassicas

Downy mildew is a very common and devastating disease in many *Brassica* growing countries (Carlier et al., 2012; Farinhó et al., 2007), resulting in huge yield losses of up to 90% if climatic situations are favourable for downy mildew (Butler, 1918; Monot et al., 2010). Several cultural and chemical techniques for controlling downy mildew have been reported to be ineffective due to the systemic nature of the disease (Vicente et al., 2012), fueling the research for new ways to reduce the yield losses caused by downy mildew (Eshraghi et al., 2007). Fungicides were likewise ineffective in preventing downy mildew disease, which lead to the emergence of new more virulent isolates (Brophy & Laing, 1992; Singh et al., 2013). As a result, *Brassica* cultivars with inherent genetic resistance to this pathogen can provide the best and most cost-effective alternative to avoid yield losses in the future. For this purpose, the identification of new sources of resistance and knowledge of the genetics of resistance are prerequisites to developing downy mildew resistant varieties. Downy mildew resistance has been identified to be controlled by different loci (Coelho et al., 1998; Coelho & Monteiro, 2003a; Dickson & Petzoldt, 1993) at both seedling (Monteiro & Williams, 1989;

Natti et al., 1967; Thomas & Jourdain, 1992) and adult plant stage (C. E. Thomas & E. Jourdain, 1990). For *B. oleracea*, some accessions were reported to be resistant at the seedling stage but susceptible at the adult plant stage (Coelho & Monteiro, 2003a, 2003b; Monteiro et al., 2005). But for other accessions the resistance observed at seedling stage for downy mildew was also observed at the adult plant stage (Jensen et al., 1999; Wang et al., 2000). At the seedling and adult-plant stages, multiple resistance sources were observed for all *B. oleracea* cultivars (Branca et al., 2005; Carlsson et al., 2004; Coelho & Monteiro, 2003b; Dickson & Petzoldt, 1996; Dickson & Petzoldt, 1993; Farnham et al., 2002; Hoser-Krauze et al., 1984; Hoser-Krauze et al., 1995; Jensen et al., 1999; Kontaxis et al., 1979; Mahajan et al., 1991; Mohammed et al., 2019; Monteiro et al., 2005; Monteiro & Williams, 1989; Natti et al., 1967; Pandey et al., 2001; Sharma et al., 1995; Silué et al., 1995; Singh et al., 1987; Singh et al., 2013; C. E. Thomas & E. L. Jourdain, 1990; Vicente et al., 2012) and other cultivated *Brassica* species (Ge et al., 2008; Lucas et al., 1988; Nashaat & Awasthi, 1995; Nashaat et al., 2004; Nashaat et al., 1997; Séguin-Swartz, 2005). The genetics of downy resistance has been reported to be complex with various modes of inheritance for both seedling and adult plant resistance (Table 1.1). At the seedling stage, a single dominant gene (Farnham et al., 2002; Jensen et al., 1999; Natti et al., 1967; Vicente et al., 2012), a recessive gene (Carlsson et al., 2004; Hoser-Krauze et al., 1984), multiple genes (Carvalho & Monteiro, 1996; Hoser-Krauze et al., 1995; B. D. Jensen et al., 1999; Monteiro et al., 2005) have been reported. At the adult plant stage, a single dominant gene (Barnes, 1968; Coelho & Monteiro, 2003b; Mahajan et al., 1995; Monteiro et al., 2005; Natti & Atkin, 1960 ; Saha et al., 2020; Verma & Singh, 2018) and a recessive gene (Mahajan et al., 1995) have been reported.

Table 1.1. Inheritance of resistance in *B. oleracea* to the pathogen *H. brassicae*.

Species	**Resistance gene**	**Reference**
B. oleracea	One dominant gene	Barnes (1968)
B. oleracea	One dominant gene	Natti and Atkin (1960)
B. oleracea	One dominant gene	Mahajan et al. (1995)
B. oleracea	One dominant gene	Farnham et al. (2002)
B. oleracea	One dominant gene	Natti et al. (1967)
B. oleracea	One dominant gene	Coelho and Monteiro (2003b)
B. oleracea	One dominant gene	Monteiro et al. (2005)
B. oleracea	One dominant gene	Saha et al. (2020)
B. oleracea	One dominant gene	Verma and Singh (2018)
B. oleracea	Two dominant genes	Natti et al. (1967); Vicente et al. (2012)
B. oleracea	Two dominant genes	Carvalho and Monteiro (1996)
B. oleracea	One recessive gene	Carlsson et al. (2004)
B. oleracea	One recessive gene	Hoser-Krauze et al. (1984)
B. oleracea	One recessive gene	Mahajan et al. (1995)
B. oleracea	One recessive gene	Carlsson et al. (2004)
B. oleracea	Multiple dominant genes	Hoser-Krauze et al. (1995)
B. oleracea	Multiple dominant genes	Moss et al. (1988)
B. oleracea	Partial resistance	B. D. Jensen et al. (1999)
B. oleracea	Quantitative genes	Dickson and Petzoldt (1996)

1.3.1 Identified QTLs conferring downy mildew resistance on *Brassica* crops

Positional cloning has been a popular approach for isolating resistance genes in model organisms and crop species. To date, twenty sources of resistance to downy mildew have been mapped in *A. thaliana* (Crute et al., 1997). These identified sources of resistance genes (Resistance to *Peronospora parasitica*, RPP) are NB-LRR which encode for receptor-like proteins with a conserved nucleotide binding motif and a leucine-rich repeat domain, and most of them are found in clusters across the genome instead of being randomly dispersed across the chromosomes (Bittner-Eddy et al., 2000; Botella et al., 1998; McDowell et al., 1998; Parker et al., 1997). Compared to *A. thaliana*, only a few sources of resistance have been identified and mapped in *B. oleracea* and other Brassicas.

In *B. oleracea*, most sources of resistance were found to be monogenically inherited. A single dominant gene located at the end of chromosome 8 shortest arm was identified (*Pp523*) in broccoli (Carlier et al., 2012; Coelho & Monteiro, 2003b; Farinho et al., 2004), and markers flanking the locus were identified (Farinhó et al., 2007). These markers were revealed to be in synteny with the top arm end of the *A. thaliana* chromosome 1 (Farinhó et al., 2007). Another source of resistance against downy mildew (*BoDowny Mildew1*) was mapped to the *B. oleracea* chromosome 5 (Gao et al., 2003; Giovannelli et al., 2002) and was found in close proximity to the glucosinolate pathway gene *BoGsl-elong* (Gao et al., 2007). Another single dominant resistance gene to downy mildew was mapped (*Ppa3*) in cauliflower (Singh et al., 2012). Saha et al. (2020) discovered another monogenic inherited dominant gene (*Ppa207*) located at the end of chromosome 2 longest arm.

In *B. rapa*, multiple sources of resistance were also mapped and candidate genes were reported (Kim et al., 2011; Li et al., 2011; Yu et al., 2016; Yu et al., 2009; Yu et al., 2011; B. Zhang et al., 2018; Zhi et al., 2016). Four closely linked QTLs (yBrDowny Mildew8, sBrDowny Mildew8, rBrDowny Mildew8, and hBrDowny Mildew8) were found to be associated with downy mildew resistance and were mapped to the long arm of the *B. rapa* chromosome 8 (Yu et al., 2016). Another two QTLs (BraDowny Mildew and BrDW) were located in the short arm of the *B. rapa* chromosome 8 (Li et al., 2011; Yu et al., 2009; Yu et al., 2011). Two other QTLs (rBrDowny Mildew6 and hBrDowny Mildew4) were mapped to the chromosome 6 and 4 top arm of the *B. rapa* genome respectively (Yu et al., 2016). A single resistance QTL (Br-Downy Mildew04) was identified to be located on the longest arm of the *B. rapa* chromosome 4 (B. Zhang et al., 2018). Another single QTL (BrRPHP1) was mapped to the shortest chromosome 1 arm on the *B. rapa* genome (Kim et al., 2011).

Table 1.2. Collection of genes/QTLs linked to Downy mildew resistance in *Brassica oleracea* (Shaw et al., 2021).

Mapping population	Locus/QTL	Chr	Linked Marker	References
F_2	-	C05	UBC359620 OPM16750	Giovannelli et al. (2002)
F_2	*Pp523*	C08	OPK17_980 ATCTA_133/134	Farinho et al. (2004)
F_2	*Pp523*	C08	SCR15 SCAFB1/BfuI	Farinhó et al. (2007)
F_2	*BoDowny Mildew1*	C05	Downy Mildew1-F Downy Mildew1-R	Gao et al. (2007)
	Pp523	C08	CB10139 CB10028	Carlier et al. (2012)
F_2	*Ppa3*	-	OPC141186 OPE141881 ISSR-231103	Singh et al. (2012)
F_2	-	-	OPC141186 OPE141881 ISSR-231103	Singh et al. (2015)
RIL	*Ppa207*	C02	BoGMS0486 BoGMS0900	Saha et al. (2020)

Table 1.3. Collection of genes/QTLs linked to Downy mildew resistance in *Brassica rapa* (Shaw et al., 2021).

Mapping population	**Locus/QTL**	**Chr**	**Linked Marker**	**References**
DH	BraDowny Mildew	A08	K14-1030 PGM Ol12G04	Yu et al. (2009)
F_2	*BrRPHP1*	A01	OPA08650 and BrPEK15B	Kim et al. (2011)
DH	*BrDW*	A08	Brb062-Indel230 Brb094-DraI787 Brb094-AatII666 Brb043-Bg1II715 Brh019-SNP137 bru1209	Li et al. (2011)
DH	*BrDW*	A08	SCK14-825 kbrb058m10-1 kbrb006c05-2	Yu et al. (2011)
DH	*sBrDowny Mildew8* *yBrDowny Mildew8* *rBrDowny Mildew8* *hBrDowny Mildew8* *rBrDowny*	A08 A06 A04	A08-709 A08-028 A08-018	Yu et al. (2016)

	Mildew6			
	hBrDowny			
	Mildew4			
Inbred lines	*QTL*	A01	A0124655323	Zhi et al. (2016)
DH, F_2	*Br-Downy*	A04	A04_5235282	B. Zhang et al. (2018)
	Mildew04		A04_5398232	

1.3.2 Molecular plant-microbe interactions

Plants are constantly exposed to a variety of microbial pathogens, including bacteria, fungi, oomycetes, and viruses that can cause disease and affect agricultural productivity (Jones & Dangl, 2006). One of these diseases, downy mildew, caused by the oomycete *Hyaloperonospora brassicae*, poses a significant threat to *Brassica oleracea* crops, including cabbage, broccoli and cauliflower. To defend against these pathogens, plants have evolved a two-level immune system consisting of molecular interactions between the plant surface-localized and intracellular protein receptors that recognize and respond to secreted pathogen molecules, also known as effectors, during pathogen invasion (Bigeard et al., 2015).

The first layer of defence is mediated by membrane-anchored pattern recognition receptors (PRRs), subdivided into receptor kinases (RLKs) and receptor-like proteins (RLPs). Both groups share common structural features, such as extracellular domains for ligand recognition, but they differ in their intracellular domains, which play a role in signal transduction (Couto & Zipfel, 2016). This type of receptor has a variable ectodomain that can recognize molecular signatures coming from the pathogen as pathogen-associated molecular patterns (PAMPs) or danger-associated molecular patterns (DAMPs) and initiate defence responses against invading pathogens (Boller & Felix, 2009).

RLKs are a large family of transmembrane proteins that play an essential role in plants. These receptors typically consist of an extracellular ligand-binding domain, a transmembrane domain, and an intracellular kinase domain (Shiu & Bleecker, 2001). Based on their extracellular domains, receptor kinases can be further classified into various families, including leucine-rich

repeat receptor kinases (LRR-RKs) and lysin motif receptor kinases (LysM-RKs) (De Smet et al., 2009), which are responsible for recognizing and binding to PAMPs (Macho & Zipfel, 2014). Upon PAMP recognition, the intracellular kinase domain becomes activated, initiating a cascade of phosphorylation events that ultimately lead to the activation of downstream defence signalling pathways (Couto & Zipfel, 2016).

RLPs are another group of transmembrane proteins that also contribute to the recognition of PAMPs. Like RLKs, RLPs possess extracellular ligand-binding domains and single-pass transmembrane domains (Wang et al., 2010). However, RLPs lack the intracellular kinase domain found in RLKs. Instead, they rely on interactions with other proteins, such as kinases or adaptor proteins, to transduce signals upon ligand binding (Gust & Felix, 2014).

Despite this defence, many pathogens have evolved to be able to suppress this first layer of defence (PTI) by delivering effectors into the plant cells. As a result, plants have evolved a second layer of defence mediated by intracellular nucleotide-binding and leucine-rich repeat receptors (NLRs), that are capable to recognize specific pathogen effectors and coordinate a quick signalling response known as effector-triggered immunity (ETI), often resulting in a hypersensitive response (HR) (Chisholm et al., 2006; Jones et al., 2016).

NLRs represent a class of intracellular immune receptors that play a critical role in plant defence against diverse pathogens (Cui et al., 2015). NLRs are characterized by a nucleotide-binding domain (NBD) and a leucine-rich repeat (LRR) region, which are involved in ATP-binding and hydrolysis, and pathogen recognition, respectively (Takken & Goverse, 2012). NLRs can be

divided into several types based on their structural features, particularly the presence of distinct N-terminal domains. The three main types of plant NLRs are, coiled-coil (CC)-NLRs, Toll/interleukin-1 receptor (TIR)-NLRs and RPW8-NLRs (Meyers et al., 2003).

Both TIR-NLRs and CC-NLRs are involved in recognizing pathogen-derived effectors and activating downstream defence responses (Wu et al., 2017).

In addition to these main types, some atypical NLRs have been identified, such as NLRs with integrated domains (ID-NLRs) or unconventional N-terminal domains. ID-NLRs contain additional domains within their NLR structure, which are thought to directly recognize pathogen effectors (Cesari, 2018).

Further characterization of these NLRs and their potential roles in downy mildew resistance in *Brassica oleracea* crops could lead to the development of improved breeding strategies for enhanced disease resistance.

1.4 Aim and objectives.

As previously detailed, downy mildew is a very important foliar disease that can cause significant economic losses in all Brassicaceae crops worldwide. The pathogen is most dangerous to young plants and seedlings, but it can also be very destructive to adult plants, especially very susceptible genotypes, where infection may become systematic affecting the inflorescence and other parts of the plant reducing the overall quality of the crops and rendering the crop unmarketable. Fungicides can be used alone or in combination with cultural practices, but they have only provided limited control to downy mildew in the fields. Therefore, the use of resistant genotypes is the most cost-effective way to control downy mildew.

For these reasons, the main objective of this PhD was to find and map new sources of resistance to downy mildew in *Brassica oleracea*.

The objectives of this thesis:

1. To collect and maintain various downy mildew field isolates from distinct geographical locations in the UK and USA and to derive all field isolates into single spore isolates and preserve them long term (Chapter 2).

2. Use the downy mildew collection to find new sources for Downy Mildew resistance in *Brassica oleracea* from different sources of plant material and create biparental mapping populations (Chapter 3).

3. To map the downy mildew resistance observed from a selected genotype by bulk segregant analysis and create markers for marker-assisted selection (MAS) (Chapter 4).

Chapter 2

Culture collection establishment of *Hyaloperonospora brassicae* from *Brassica oleracea* crop fields in the UK and USA

2.1 Introduction

Downy mildews are a unique group of obligate plant pathogens classified in the *Mastigomycotina* of the *Oomycete* order *Peronosporales*. Some indicators suggest that due to their obligate parasitism, downy mildews have very limited host ranges and hence could represent separate species. Due to this reason, based on molecular analysis, Constantinescu and Fatehi (2002) segregated several isolates from the genus *Peronospora* into two new genera, i.e. *Perofascia* and *Hyaloperonospora*. *Hyaloperonospora* was the pathogen on cruciferous hosts, with *H. arabidopsidis* specific to *Arabidopsis thaliana*, *H. parasitica* to *Capsella bursa-pastoris* and *H. brassicae* specific to Brassicaceae. Some of the most extensively researched *H. brassicae* isolates have come from the

United Kingdom, France, Italy, Portugal, India, and South Africa.

Due to their obligate pathogenic nature, downy mildews cannot be cultivated in artificial conditions and must rely only on the host plant for nutrition and survival (Lucas et al., 1994). For this reason, to study *H. brassicae*, a steady supply of pathogen material by successive transfers on cotyledons and leaf material from established sources is necessary. Downy Mildew callus culture has been described using a variety of approaches (Achar, 1995; Ingram, 1969). In addition, callous tissue allowed Downy Mildew to survive for 14-21 days on the callous, and it was observed that it decreased cross-contamination compared to other methods. However, calluses infected with Downy Mildew could not be transferred directly, and cell wall structure differences compared with whole-plant cell walls could affect colonization and growth (Ingram, 1969). For these reasons, callous was not the ideal tissue for testing host germplasm for resistance purposes since it did not always reflect the reaction of the entire plant.

It is critical to keep *H. brassicae* in a controlled environment on appropriate brassica hosts or, preferably, *in vitro* under favourable environments for propagation. A simple method for maintaining *H. brassicae in vitro* on a suitable living host in a controlled environmental room was developed (S. M. Kluczewski & J. A. Lucas, 1982; Kluczewski & Lucas, 1983; Sheriff & Lucas, 1989a, 1989b). Cotyledons of the host plant were detached nine days after sowing and placed face down on a damp filter paper in a translucent plastic box (Bonnet & Blancard, 1987; S. M. Kluczewski & J. A. Lucas, 1982; Kluczewski & Lucas, 1983; Sheriff & Lucas, 1989a, 1989b). These cotyledons were inoculated with Downy Mildew in a sealed container and covered to maintain a constant relative high humidity. These containers were kept in a C.E.R at 14-20°C and a day-night cycle with a 12-16h photoperiod (Bonnet

& Blancard, 1987; Channon, 1981; S. M. Kluczewski & J. A. Lucas, 1982; Kluczewski & Lucas, 1983; Nashaat & Awasthi, 1995; Sheriff & Lucas, 1987; Sheriff & Lucas, 1989a, 1989b; Sherriff & Lucas, 1990; D. Silue et al., 1996). In approximately 5-7 days, these cotyledons will be sporulated and ready to be subcultured onto fresh cotyledons. Some control measures could be taken to minimise cross-contamination for experiments where this can be problematic, including autoclaving soil, surface sterilisation of seed and subculturing at intervals of 7-9 days in previously sterilised environments for every isolate.

Highly variability has been reported on *H. brassicae* isolates collected from the fields due to the presence of different races or *forma specialis* of *H. brassicae* (Natti et al., 1967; Sherriff & Lucas, 1990; Tham et al., 1994). Therefore, for long-term consistency between experiments, it is necessary to create a uniform genetic collection of isolates, also defined as monospore lines or single-spore isolates.

The viability of spore sources could complicate inoculum availability. Spores' ability to survive is determined by the environment in which they are stored or maintained. For example, *H. parasitica* spore on cabbage survives longer in cold, dry conditions (Krober, 1981). In general, a constant relative humidity is more important than the temperature for spore viability. Spores can survive for ten days on detached leaves in the field during warm days. If the spore is buried in dry soil, it can survive for 110 days. If the soil is moist, the survival span is drastically reduced to a maximum of 22 days. A marked decline in survival has also been reported following summer storage in dry and moist soils. When the spores are stored in air-dried soil at a constant temperature of 5 °C, the viability of the spores is extended for up to 130 days (Krober, 1970). Spores stored at −25 °C and

moderately dry on leaf discs sustain a relatively high germination rate after one year or more (Krober, 1970; Krober, 1981). Nevertheless, oospore formation is prolific in infected tissues of all cultivated brassica crops, and they represent a crucial source of survival for downy mildew over periods of unfavourable conditions (Jang & Safeeulla, 1990; Kolte, 1985; LeBeau, 1945; McMeekin, 1960). In the soil, oospores have been observed to stay viable and infective for two to three years.

Michelmore and Ingram (1980) described several long-term methods to store obligatory pathogens. Generally, the pathogen is maintained long-term by storing a few sporulating cotyledons sealed in tiny vials at -70 °C (Michelmore & Ingram, 1980; Nashaat & Rawlinson, 1994). Such treatment did not generally result in any significant loss of viability, even after one year of storage. (Paul et al., 1998)successfully used the cryoprotective agent's glycerine polyethene glycol (PEG) 400 or 1000 allowed for the long-term preservation of *H. brassicae* spore in a freezer at -21 °C for one year.

The objectives of the present chapter are:

a) To collect and preserve different *Hyaloperonospora brassicae* isolates from distinct geographical locations in the UK and USA.

b) To obtain single spore isolates and preserve them long term.

2.2 Material and Methods

2.2.1 Hyaloperonospora brassicae culturing in the laboratory

Brassica oleracea seeds, specifically from the commercial kale variety Dwarf Blue (CN Seeds), were used for maintaining cultures and multiplying the sporangial inoculum of all *H. brassicae* isolates. Before using this line for experiments and downy mildew maintenance, this variety was tested for seed-borne infection of plant pathogens. A subset of seeds from a 500g package was surface sterilised with 2% sodium hypochlorite for 2-3 minutes and grown in pots filled with JIC Arabidopsis mix. These pots were placed in a covered propagator in a CER under the following conditions: 10-hour day, 20°C. After ten days, these uninoculated seedlings were moved to a different CER set at the optimum conditions for *H. brassicae* infection, 12-hour day, 15°C. After two weeks, the seedlings did not show any sign of infection. This seed package was used for culturing and multiplying all the isolates or for phenotyping experiments as a control.

An *in vitro* method was used to maintain all fields and single spore *H. brassicae isolates*. UK isolates of *H. brassicae* of the *Brassica oleracea* pathotype were collected by CN seeds from different customers, and field trials in six growing locations (Table 2.1). USA isolates of *H. brassicae* of the *Brassica oleracea* pathotype were sent by post from three fields in the USA. (Table 2.1). Most customers were growing different kale varieties purchased from CN seeds, so DB was used as a universal susceptible variety for culturing and maintaining the isolates.

B. oleracea seeds were surface sterilised for 1 min with 2% sodium hypochlorite and sowed in pots at 20°C in a CER After 10-14 days, the cotyledons are big enough for inoculation, but this depends on the *Brassica* species and

cultivar; 10-14 days was sufficient for all *B. oleracea* cultivars tested, but for many *B. napus* lines, 7-10 days was preferable. All downy mildew work was performed inside a Class 2 Microbiological Safety Cabinet to avoid the dispersal of spores in the laboratory. Inside the cabinet, previously cultured cotyledons showing downy mildew sporulation were transferred to a 5mL Eppendorf tube containing 2mL 0.05% sterile tween 20 solution using sterilised tweezers. The tube was shaken or vortexed gently for a few seconds to produce spore suspension (1-3 fully sporulated cotyledons are usually enough to inoculate 30 cotyledons). The resulting suspension was inoculated by pipetting two 10µl droplets on each side of the 10-14 days old seedling cotyledons (Abaxial side) placed on water agar plates with benzimidazole (6g Agar + 900 mL distilled water + 100 mL Benzimidazole [1g/L]). After inoculation, the plates were sealed with micropore tape to ensure high humidity. These plates are then incubated for 7-14 days in a growth room at 15°C with a day length of 12h. The cabinet must be carefully swabbed with 70% ethanol and allow 10-20 minutes between subcultures to prevent cross-contamination. Forceps were sterilised by dipping in 70% ethanol and flaming. Depending on cotyledon collapsing, every 7-10 days, the isolates were subcultured routinely and incubated at the optimum maintenance conditions described earlier. For long-term storage, 5-10 entire cotyledons exhibiting pathogen sporulation were collected in a 5 ml Eppendorf tube and stored at -80°C for up to one year.

2.2.2 *Hyaloperonospora brassicae* single spore isolates derivation

Single spore isolates were produced from field isolates collected by transferring a single conidium to the surface of an excised cotyledon. The

primary goal was to achieve successful infection and profuse sporulation from the single conidium so that genetically homogeneous isolates could be subsequently propagated and maintained over time.

Single spore isolates were derived using nine field isolates (Table 2.1) collected between 2016, 2018 and 2019 from the U.K. and U.S.A. Nine-day-old susceptible kale seedlings (cv. Dwarf Blue) were used for this experiment. These isolates were maintained on cotyledons detached from the seedlings using sharp, sterilised tweezers. Two of these cotyledons were placed upside down in square sealed plastic boxes (4x2cm) filled with water agar with benzimidazole (6g Agar + 900 mL distilled water + 100mL Benzimidazole [1g/L]), one of them was infected by single spore, and the other one was used as a control (Figure 1). For every isolate, a total of 12 small individual plates were prepared.

2.2.3 Inoculum preparation

Inside a Class 2 Microbiological Safety Cabinet, cotyledons showing downy mildew sporulation were used as the source of inoculum. Three cotyledons were transferred into a 5mL Eppendorf tube containing 1mL of 0.05% sterile tween 20 solutions. The tube was gently shaken to produce spore suspension and filtered using two layers of Miracloth. The spore concentration of the resulting suspension was calculated using a hemocytometer and adjusted to a final concentration of 0.5 spores/µl.

2.2.4 Single spore isolation and inoculation

A single 2% water agar Petri plate was prepared, and several agar plugs were cut using a 3mm ⌀ biopsy puncher. Those plugs were placed on top of a microscopy slide (Figure 2.1-a). Two microliters of the inoculum prepared

in section Inoculum preparation were pipetted on each one of them (Figure 2.1-a). Immediately after pipetting, the whole surface of the plug was scanned using a microscope (40x magnification), discarding plugs with 0, 2 or more conidiospores and solely using the ones with one conidiospore (Figure 2.1-b, c).

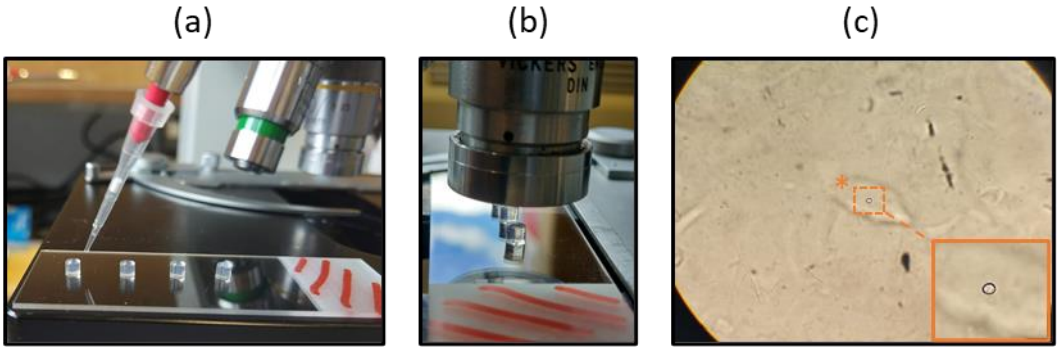


Figure 2.1. (a) Pipetting 2µl of inoculum on top of the agar plug; (b) Inoculated agar plug under the microscope; (c) Single conidiospore on an agar plug (*single conidiospore on top of an agar plug).

To favour the infection process on the inoculated cotyledon, a 10µl droplet of sterile distilled water was pipetted on the surface. Plugs with a single conidiospore were transferred spore side down onto the droplet on the surface of one of the two cotyledons placed on the small plates (Figure 2.2-a, b). Twelve plates were prepared and placed inside a bigger sealed squared plastic plate for each isolate (Figure 2.3). These plates were placed inside a CER programmed under the following conditions: 12 hours daylight, 15°C. Every isolate was closely monitored every day until sporulation was visible. When cotyledons were fully colonised (Figure 2.3), individual cotyledons were subcultured. Excellent sporulation was seen six to eight days after inoculation.

Figure 2.2. (a) Transferring agar plugs with a single conidiospore using tweezers; (b) Agar plug spore side down onto one of the cotyledons.

The spore isolation procedure was repeated one more time for every isolate. After this, five plates full of fresh Dwarf Blue cotyledons (≈100 cotyledons) were prepared and inoculated for all nine isolates. One of the plates was used to maintain the single spore isolates, and the rest were transferred into six different 5 mL Eppendorf tubes and stored at -80°C. Single spore isolates were named with their original names plus "S.S-2" (Single Spore).

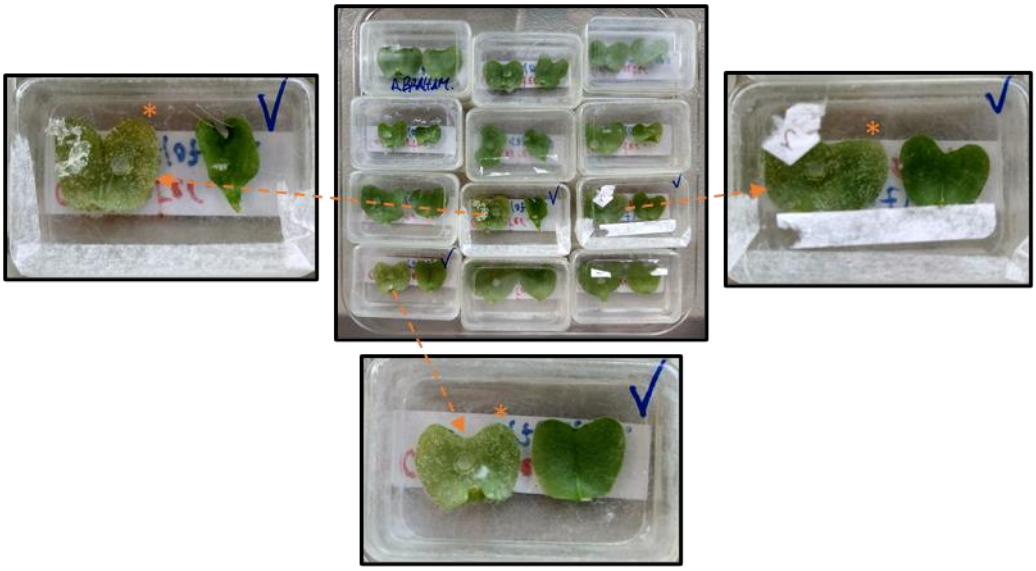


Figure 2.3. Square plate showing all 12 small, inoculated plates using agar plugs with a single conidiospore (Left cotyledon on every plate) for one of the isolates. *3 out of 12 plates successfully sporulated.

2.3 Results

2.3.1 *Hyaloperonospora brassicae* culture collection

A comprehensive single spore isolate set from different commercial fields in the UK and the USA was created from field isolates collected in 2016, 2018 and 2019 (Table 2.1). All isolates represented in Table 2.1 successfully infected the universally susceptible *Brassica oleracea* commercial line Dwarf Blue (CN seeds). Therefore, this line was used to maintain all single spore and field isolates in the laboratory on a two-week basis. Due to low infection levels, the derivation of three other field isolates sent by CN seeds customers from the USA. was impossible even using the original commercial variety.

Table 2.1. List of all *Hyaloperonospora brassicae* field isolates collected between 2016-2019 from commercial *B. oleracea* fields.

Isolate Name	Origin	Year	Crop
Colchester field	Colchester field, UK	2016	*B. oleracea* var. *acephala*
Intercrop field	Deal field, UK	2016	*B. oleracea* var. *acephala*
Martin's field	Spalding field, UK	2016	*B. oleracea* var. *acephala*
Malcom's field	Thetford field, UK	2016	*B. oleracea* var. *acephala*
Lettuce field	West Sussex field, UK	2016	*B. oleracea* var. *acephala*
Vel fresco field	Evesham field, UK	2016	*B. oleracea* var. *acephala*
1047 field	Salinas, California, USA	2018	*B. oleracea* var. *acephala*
1048 field	Salinas, California, USA	2018	*B. oleracea* var. *acephala*
Happyland field	Salinas, California, USA	2019	*B. oleracea* var. *acephala*

2.3.2 Maintenance of *Hyaloperonospora brassicae* single spore and field isolates

The most important environmental factors influencing Downy mildew infection for *B. oleracea* in the laboratory were air temperature and relative humidity. Downy mildew sporulation could not occur on seedlings without relatively high environmental humidity, but this was even more critical for *in vitro* assays. At 20-28°C, symptoms appeared quickly after inoculation, but sporulation was limited. Cotyledons become senescent at this temperature range and collapse after 5-7 days. At low temperatures between 5-10°C, Downy mildew infection growth was slowed, and sporulation was delayed to 10-14 days after inoculation. For susceptible lines, cotyledon senescence was slowed considerably at this temperature range of 2-3 weeks after inoculation. Cotyledon senescence was minimal at 5-15°C; susceptible lines would start collapsing after 12-16 days, and resistant lines or not inoculated susceptible lines would start yellowing after 21 days, and some cotyledons will start rooting in the water agar plate. After inoculation, high relative humidity was achieved by sealing the plates with parafilm or micropore tape.

Based on these findings, the laboratory's most favourable temperature for maintenance and disease assays was 14-16°C. This range was the balance between downy mildew growth and host senescence, allowing both the host and the pathogen to grow at a slower pace, resulting in minor damage and, therefore, a more widespread disease development. Under these environmental conditions, profuse sporulation of *H. brassicae* was achieved 6-8 days after inoculation (Figure 2.4).

Figure 2.4. Sporulated *B. oleracea* cotyledon from *H. brassicae* single spore.

Single spore derivation methodology from downy mildew field isolates resulted in infection rates ranging from 25-50%. Successful infection relied on the single spore being entrapped between the water agar plug and the abaxial surface side of the cotyledon. Adding a 10µl droplet of water to the infection site allowed the agar plug to stay in place and further boost the infection. Small individual sealed plates (Figure 2.4) were highly effective at maintaining the high humidity required for Downy Mildew infection and sporulation.

After five days, single spore isolates, which showed the first sporangia emergence, were monitored until day eight, when they were bulked by inoculating a plate with susceptible DB cotyledons. Figure 2.3 depicts the profuse sporulation of three cotyledons out of twelve plates used to produce single spore cultures of the Colchester field isolates. Isolates were grown for several generations by subculturing at regular intervals. After inoculum preparation, the spores concentration was higher when spores were collected using 0.05-0.1% Tween 20 solution rather than distilled water.

2.4 Discussion

Downy mildews are biotrophic organisms with very sophisticated nutritional requirements that cannot be supplied when the host plant dies. Therefore, early work on this pathogen concentrated on different strategies for cultivating it in the laboratory with varying levels of success (Ingram, 1969; Krober, 1969; Krober, 1970; McMeekin, 1981). Guttenberg and Schmoller (1958) attempted, with minimal success, to culture downy mildew isolates in the absence of living plant tissue. Ingram (1969) was able to grow and maintain cultures of downy mildew effectively on callus tissues for more extended periods. This methodology, however, was not widely used because calluses could not be transferred directly (Ingram, 1969). Moreover, callus culture could not be used to research plant resistance since the actual response to the pathogen did not represent that of the seedling or the adult plant.

Therefore, the most cost-effective and widely used way to maintain and produce inoculum is to culture downy mildew on host seedlings or detached cotyledons *in vitro*. Early investigations by Davison (1968) and Pinckard (1942) produced significant information on the optimal conditions that promote infection and sporulation of downy mildew. In summary, high humidity, low temperature and a dark-light cycle favour downy mildew culture maintenance. The possibility to freeze the inoculum on host plants for more extended periods made it easier to have a ready source of inoculum when necessary (Satou & Fukumoto, 1993; VH & E, 1993). In this study, the temperature range for maximum infection of seedlings and detached cotyledons in vitro using a known susceptible *B. oleracea* host was 15°C. The water agar plate provided high relative humidity and hydration

on in vitro assays. Under these settings, sporulation was observed 6-8 days after inoculation.

In the long run, it was crucial to find a universally susceptible *B. oleracea* host for all the isolates before initiating a downy mildew maintenance routine. The time and effort involved in the maintenance decreased considerably by not having to sow different susceptible varieties/lines for different isolates. The biotrophic nature of downy mildew made it necessary to find ways to preserve the culture collection. One necessary measure to avoid the loss of isolates due to a climatic chamber or freezer malfunction was to back up the cultures in different freezers and locations. Special measures had to be taken during shipping for the importation of isolates from the USA to the John Innes Centre. Large amounts of fresh sporulating imported material were necessary due to the very low viability of the spores. Higher spore viability was achieved by sending infected cotyledons 1-2 days after inoculation in a sealed water agar plate. High viability was also achieved using frozen sporulated cotyledons in dry ice.

In this study, an *H. brassicae* collection was successfully bulked and maintained using the universally susceptible commercial variety Dwarf Blue. All isolates were collected from 9 different *B. oleracea* commercial fields. Six of these fields were in different geographical locations in the UK and three separate fields from a single location in the USA. Many single spore isolates were created for every isolate from the fields. This provides a suitable *H. brassicae* isolate collection with the aim of identifying new sources of resistance to *H. brassicae* in *B. oleracea* germplasm, which is the core of this work and will be discussed in the subsequent chapters.

Chapter 3

Finding new sources for *H. brassicae* resistance in *B. oleracea*

3.1 Introduction

Downy mildew, caused by the oomycete *Hyaloperonospora brassicae* (Gäum.) (Göker et al., 2003), is the most widely reported foliar disease in horticultural and agricultural *Brassicaceae* species including *Brassica oleracea* (Coelho et al., 2012; D. Silue et al., 1996). It can be very destructive at all developmental stages, but young seedlings are more vulnerable (Koike *et al.*, 2007), causing heavy losses in plant nurseries (Channon, 1981). In the fields, yield losses from this disease can be very high (Butler, 1918) or completely fatal depending on climatic situations (Monot et al., 2010). Severe Downy Mildew disease can severely impair the final market product quality for specific brassica crops where the external appearance of the products is very important like broccoli/cauliflower heads and crops like kale or collard leaves (Coelho et al., 2012; Dickson & Petzoldt, 1993; Koike et al., 2009).

Downy mildew control with fungicides or cultural strategies has only

achieved limited or partial control at a considerable expense (Neik et al., 2017; Vicente et al., 2012). Therefore, the adoption of genetically resistant plants is one of the most efficient, low-cost, and environmentally friendly ways for plant disease prevention. Several sources of genetic resistance to downy mildew have been found in *B. oleracea* seedlings and adult plants (Branca et al., 2005; Carlsson et al., 2004; Coelho et al., 1998; Coelho & Monteiro, 2003b; Dickson & Petzoldt, 1996; Dickson & Petzoldt, 1993; Farnham et al., 2002; Hoser-Krauze et al., 1995; Jensen et al., 1999; B. D. Jensen et al., 1999; Kontaxis et al., 1979; Mahajan et al., 1991; Mahajan et al., 1995; Ménard et al., 1999; Mohammed et al., 2019; Monteiro et al., 2005; Monteiro & Williams, 1989; Natti & Atkin, 1960; Natti et al., 1967; Neik et al., 2017; Pandey et al., 2001; Rooster et al., 1999; Sharma et al., 1991; Sharma et al., 1995; Silué et al., 1995; Singh et al., 1987; Singh et al., 2013; C. E. Thomas & E. Jourdain, 1990; C. E. Thomas & E. L. Jourdain, 1990; Vicente et al., 2012). However, there have been limited studies on the genetic mapping of these resistances (Carlier et al., 2012; Farinho et al., 2004; Gao et al., 2003; Giovannelli et al., 2002; Saha et al., 2020; Singh et al., 2012), and to date, no R gene has been cloned from *Brassica* species.

The resistant and moderately resistant genotypes identified in this study against *H. brassicae* offers opportunities to map these sources of resistance and/or develop new highly resistant commercial varieties.

3.2 Material and Methods

3.2.1 *Brassica oleracea* plant material

In this study, a total of 529 genotypes from three different sources were used to assess them for *H. brassicae* resistance at the cotyledon stage. For every experiment, *cv.* Dwarf Blue (CN seeds) was used as positive control by

inoculating it alongside the other genotypes, and as a negative control by inoculating it with distilled water.

3.2.1.1 A12xGD fixed DH population

The first source is a double haploid (DH) population derived from a cross between *Brassica oleracea* ssp *alboglabra* (A12DHd) and *B. oleracea* ssp *italica* (Green Duke GDDH33), also known as the A12xGD population. Seed material of 75 DH genotypes, including thirteen substitution lines (Ramsay et al., 1996) and the two parental lines A12DHd and GDDH33 (Appendix D1), were made available from the John Innes Centre, Norwich, UK.

3.2.1.2 *Brassica oleracea* Diversity Fixed Foundation Set (Bol DFFS)

The second source is a subset of 58 fixed genotypes (DH) from the *B. oleracea* diversity fixed foundation set, from Warwick University (Walley et al., 2012). This seed set includes many different *B. oleracea* cultivars, including broccoli, cauliflower, cabbage, brussels sprout, kale, calabrese, kohlrabi, etc (Appendix D2). This set includes TO1000, a genotype used to create one of the reference genomes available for *B. oleracea*. Seeds were made available by Rachel Wells from the John Innes Centre, Norwich, UK. All data for the DH lines are given in the Appendix D1 table.

3.2.1.3 *Brassica oleracea* accessions from the Warwick genetic resources unit and breeding genotypes from CN seeds

The third source consists of 390 *Brassica oleracea* var. *acephala* accessions from the Warwick genetic resources unit (WIEWS instcode-GBR006) gently provided by Charlotte Allender (Warwick gene bank) and six breeding lines from CN Seeds. The country origin of all accessions is summarised in Table 3.1. Full details for all the accessions and breeding lines are given in the

Appendix D3 table.

Table 3.1. Country of origin for the 390 *B. oleracea* var. *acephala* accessions from the Warwick GRU and the 6 breeding lines from CN seeds.

Country	Number of accessions
Portugal	189
Great Britain	120
Spain	23
Japan	15
France	13
Unknown	7
Ireland	6
Germany	4
Italy	3
New Zealand	3
Netherland	2
Sweden	2
Yugo Slovakia	2
Denmark	1
Syria	1
USA	1

3.2.2 *Hyaloperonospora brassicae* isolates

For all three seed sets described in section 3.2.1, a mixture of nine single spore isolates or a single field isolate was used to screen the material for downy mildew resistance. Due to isolate availability, the first seed set of 389 *Brassica oleracea* var. *acephala* accessions was initially screened for downy mildew resistance using a single *H. brassicae* isolate from the Colchester field (Table 2.1). Later, a subset of the most resistant accessions to downy mildew was screened again but this time in vitro using a mixture of all nine single spore isolates. For the other two seed sets, a mixture of nine single-spore *H. brassicae* isolates derived from field isolates was used (Table 2.1).

All *H. brassicae* isolates used in this study were maintained individually in

vitro on 10-14-day old seedlings of the universally susceptible *B. oleracea* cv. Dwarf Blue.

3.2.3 Inoculum preparation

To ensure the viability of the maximum number of spores, fresh inoculum from infected cotyledons must be used every time. Therefore, if fresh inoculum is not available, frozen spores must be transferred at least once into fresh cotyledons.

For all *H. brassicae* isolates used in this study, cotyledons that displayed abundant sporulation were collected and placed in a 50 mL tube containing 40 mL of 0.05% sterile tween 20 solution inside a Class 2 microbiological safety cabinet. The tubes were gently shaken/vortex to dislodge the sporangia. The resulting suspension was filtered into a new 50 mL tube using a funnel and two layers of Miracloth to remove suspended plant material and other contaminants. The spore suspension concentration was calculated using a hemocytometer and adjusted to a final concentration of $4x10^4$ spores/mL. Due to time, space and seed availability limitations, all individual isolates could not be tested individually for all *B. oleracea* genotypes used in this study. Therefore, the purpose of using a mixture of isolates from different geographical locations in the UK and the USA (Table 2.1) was to minimize any potential unique pathotype-specific interaction that may occur when screening *B. oleracea* genotypes for *H. brassicae* resistance.

3.2.4 *Hyaloperonospora brassicae* inoculation

In this study, seedlings were inoculated in two ways. The first set, including all 389 accessions from Warwick and the six breeding lines from CN Seeds, were inoculated for downy mildew resistance at the seedling stage by

pipetting the *H. brassicae* inoculum directly on the seedlings cotyledons growing on the trays (Figure 3.1). For the other two sets (A12xGD DH; Bol DFFS) and the subset from the Warwick accessions, all the genotypes were screened *in vitro*.

3.2.4.1 Seedling inoculation on trays

Each *B. oleracea* accession was sown in 40 cell seedling trays containing JIC Arabidopsis Mix. After seven days every cell was reduced to two seedlings. All the accessions were grown at 22°C day, 15°C night, 16-hour photoperiod and 70-90% humidity in a CER. All trays were watered daily to field capacity. Ten days after sowing all seedlings were inoculated with a single *H. brassicae* field isolate (Colchester field) fresh spore suspension adjusted to $4x10^4$/mL, by pipetting two 10µl droplets to the top side of each cotyledon (Figure 3.1-a,b). Immediately after inoculation, trays were watered and covered with a transparent propagator lid to maintain high relative humidity and moved to a CER at 15°C day, 10°C night and 70-90% humidity for 17 days (Figure 3.1-c).

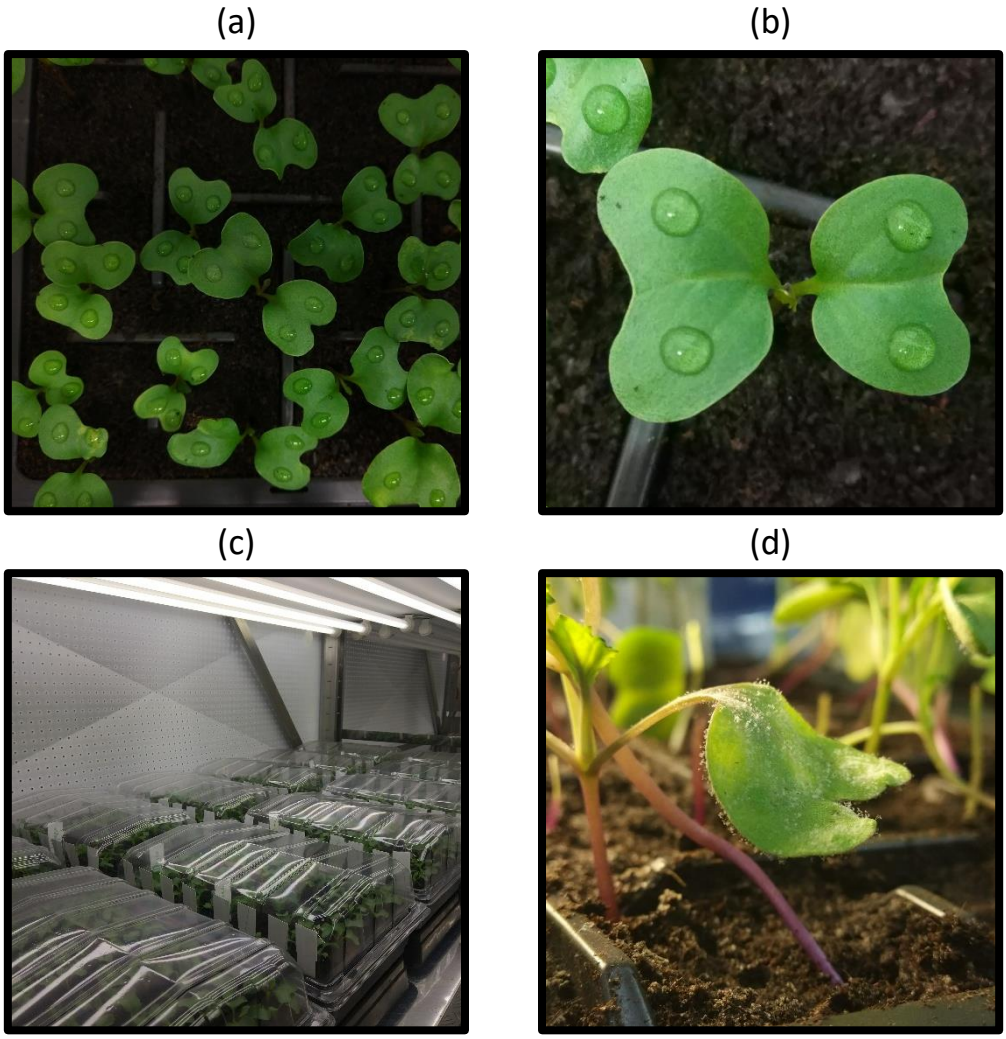


Figure 3.1. (a) Seedlings inoculated with *H. brassicae*. (b) Closeup of an inoculated seedling. (c) Trays covered with a transparent propagator lid to maintain high relative humidity in a CER at 15°C day, 10°C night. (d) Susceptible *B. oleracea* genotype with visible *H. brassicae* sporangias at 17 dpi.

3.2.4.2 In vitro cotyledon inoculation

Seeds were sown in 40 cell seedling trays containing JIC Arabidopsis Mix. After sowing the trays were moved to a CER at 20°C and a 10-hour day. After ten days trays were taken out of the CER into a Class 2 Microbiological Safety Cabinet and cotyledons were detached from the seedlings using a

sterilised tweezer. For every experiment and genotype, cotyledons were placed upside down randomly into three different plates containing water agar with benzimidazole (6g Agar + 900 mL distilled water + 100 mL Benzimidazole [1g/L]). All genotypes were inoculated using a fresh inoculum mixture of all nine single spore isolates (Table 2.1) and pipetting two 10µl droplets on each side of the cotyledon. After inoculation plates are sealed using micropore tape and incubated at 15°C with a day length of 12h for 14 days. The plates are placed in a fully randomized design over a metallic grill to allow for air circulation to try to minimise condensation on the plates (Figure 3.2).

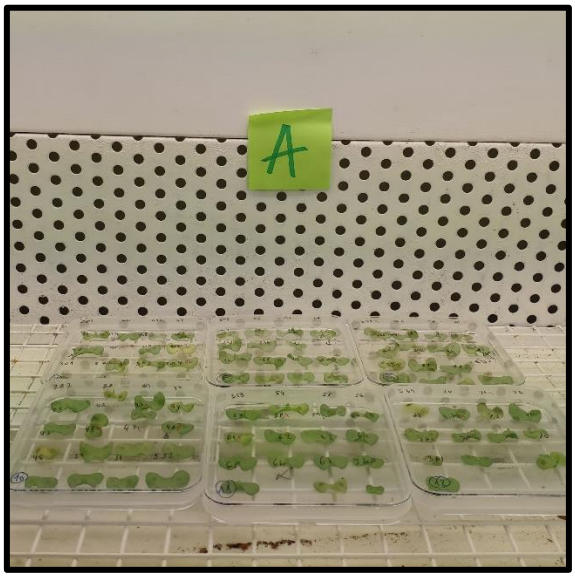


Figure 3.2. Sealed water agar plates with benzimidazole on top of a metallic grill to allow for air circulation around the plate and avoid excessive condensation on the plate lid.

3.2.4.3 In vitro leaf disk inoculation

To assess downy mildew resistance at the adult plant stage of selected very resistant genotypes from the BolDFFS, 22 mm width leaf discs were cut using a cork-borer from different leaves at the 6-leaf stage. To allow the downy mildew inoculum to adhere properly to the leaf surface the wax was

removed by submerging the disks in 2% distilled water with tween 20 for 2 minutes. After this, disks were submerged again in distilled water to remove the tween 20 and dried gently with a paper roll. Physically damaged disks were discarded and no damaged ones were placed upside down on plates containing water agar with benzimidazole (6g Agar + 900 mL distilled water + 100 mL Benzimidazole [1g/L]). Three plants were used per genotype and 16 disks were assessed per plant. All leaf disks were inoculated by spraying the Downy Mildew inoculum (68) using a hand spray bottle. After inoculation, plates are sealed using micropore tape and incubated at 15°C with a day length of 12h for 14 days. Genotypes were assessed for Downy Mildew resistance and categorized into resistant or susceptible based on leaf sporulation after 14 dpi.

3.2.5 Cotyledon disease assessment

The *H. brassicae* disease severity was assessed at two different time points after inoculation. For seedlings on trays (389 accessions from Warwick and the six breeding lines), the disease was assessed at 10 and 17 dpi for every cotyledon. For *in vitro* studies, cotyledons were assessed at 7 and 14 dpi. The infection was evaluated based on a category scale from 0 to 3: 0 = no symptoms; 1 = necrotic flecks at the point of inoculation, no sporulation (Figure 3.4-a); 1.5 = very sparse sporulation (one to five sporangia) localized at the points of inoculation (Figure 3.4-b); 2 = sparse sporulation localised at the points of inoculation (Figure 3.4-c); 2.5 = abundant sporulation, sporangia are not localised to the points of inoculation; 3 = heavy sporulation, most of the surface of the cotyledon is covered with sporangia (Figure 3.3). The Townsend– Heuberger formula (Townsend & Heuberger, 1943) was used to calculate the disease index (DI%) using the disease scores for all replicates per genotype. The equation used to calculate the disease index (DI%) = Σ (n

x V/Z x N) x 100. Where n is the number of cotyledons with a specific disease score on the scale, V is the scale value, Z is the highest scale value and N is the total amount of cotyledons scored. All genotypes were categorised using the DI% from the latest day, 14 dpi for in vitro and 17 for seedlings Downy Mildew assessment. Genotypes with a DI% between 0-33.3 were categorised as highly resistant, 33.3-70 as moderately resistant and >71 as susceptible.

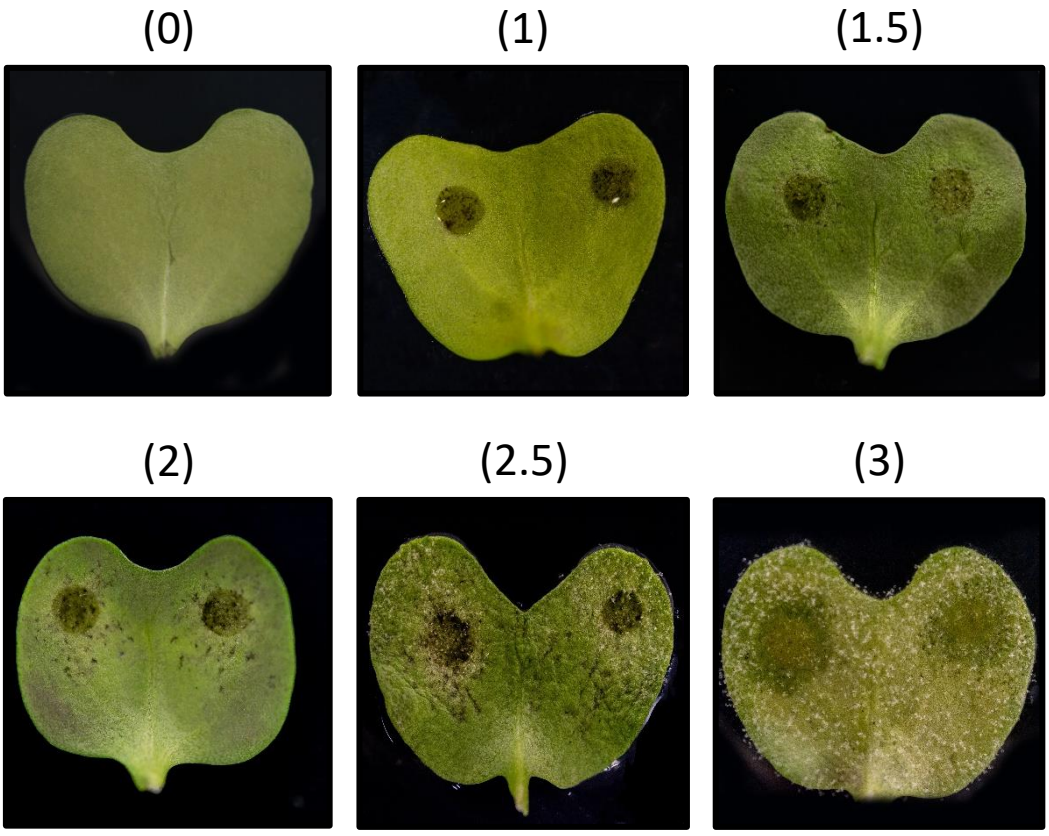


Figure 3.3. All six categories for visual *H. brassicae* dissease assessment.

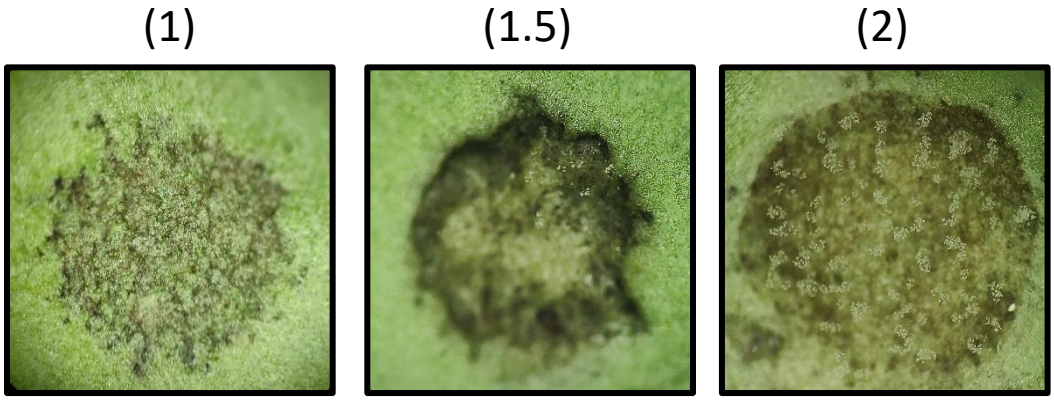


Figure 3.4. Closeup for one of the inoculation points for category 1, 1.5 and 2. All three have visual necrotic flecks. Category 1.5 has very few Downy Mildew conodiophores visible localised at the point of inoculation. Category 2 has more abundant quantity of visible Downy Mildew sporangias but localised to the point of inocualtion.

3.2.6 Creation of F_2 and BC_1 populations

F_1 hybrids were created by manually crossing identified resistant genotypes to downy mildew with identified susceptible genotypes to downy mildew. All selected accessions were selfed by hand once before using them to create the F_1 crosses. Fixed DH genotypes were not selfed before creating the F_1 crosses. Only closed buds were used for crossing to prevent self-pollination. These closed flower buds were emasculated by removing all unmatured anthers from the bud, leaving the stigma out and hand-pollinating it with other genotype anthers pollen. F_1 hybrids were self-pollinated by hand to produce the F_2, while BC_1 were developed by crossing back F_1 hybrids with the susceptible parent. The majority of the F_2 populations were created by Sujit Tha Shrestha at the CN Seeds facilities and some of the F_2 populations and all the BC_1 were created at the JIC facilities.

3.2.7 Multiplex PCR assay for rapid identification of *Brassica* species

DNA isolation was performed by Richard Goram from the genotyping service platform at the John Innes Centre using leaf material from three different seedlings. Initially, as a proof of concept, a *Brassica* diversity panel including close relatives was tested and DNA material was extracted from all 26 different genotypes and commercial varieties, including very diverse genotypes for all 6 *Brassica* genomes and also 3 Brassicaceae close relatives (Table 3.2). DNA material was also extracted from all 390 *B. oleracea* accessions from the Warwick GRU (Section 0) and the 6 breeding lines from CN seeds. Genomic DNA was quantified using the NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. DNA concentrations (ng/µl) were adjusted in sterile distilled water to concentrations of ≈ 100ng/µl.

According to the methodology described by Koh et al. (2017); 5 primers were used (Table 3.3). Two A genome-specific markers, A6-1 (253 bp) and A6-2 (186 bp) (both located on chromosome A06 of *B. rapa*, separated by 30 kbp), two C genome-specific markers, C1 (157 bp, located on chromosome C01 of *B. oleracea*) and C9 (612 bp, located on chromosome C09 of *B. oleracea*) and one B genome marker, B (331bp).

PCR reactions were performed on an Eppendorf Mastercycler® pro in 15 µl volumes containing 2 µl genomic DNA, 7.5 µl HotStartTaq Master Mix (Qiagen, Germany), primer mix (forward and reverse each) of 0.8 µl (10 µM) C1 and C9, 0.4 µl (10 µM) A6-1 and A6-2, and 0.3 µl (10 µM) B; 0.1 µl of water. PCR cycling conditions were as follows: 96°C for 15 min followed by 35 cycles at 95 °C for 30 s (Ramp 50%), 58 °C for 20 s (Ramp 60%) and 72 °C

for 30 s, then a final extension step at 72 °C for 2 min. PCR products were separated by electrophoresis in 2% (w/v) agarose in 1× TBE (Tris–Borate–EDTA).

Table 3.2. Small *Brassica* diversity panel including 20 *Brassica* genotypes and 6 close *Brassica* relatives.

Code name	Specie
Dwarf blue	*Brassica oleracea* [C genome]
Kale 1028	*Brassica oleracea* [C genome]
53	*Brassica oleracea* [C genome]
57	*Brassica oleracea* [C genome]
2	*Brassica rapa* [A genome]
16	*Brassica rapa* [A genome]
24	*Brassica rapa* [A genome]
28	*Brassica rapa* [A genome]
1127	*Brassica nigra* [B genome]
1142	*Brassica nigra* [B genome]
1156	*Brassica nigra* [B genome]
49	*Brassica napus* [A+C genome]
50	*Brassica napus* [A+C genome]
65	*Brassica napus* [A+C genome]
32	*Brassica juncea* [A+B genome]
36	*Brassica juncea* [A+B genome]
38	*Brassica juncea* [A+B genome]
61	*Brassica carinata* [BC]
1212	*Brassica carinata* [BC]
1213	*Brassica carinata* [BC]
163	*Diplotaxis tenuifoilia*
2443(168)	*Diplotaxis tenuifoilia*
24	*Eruca sativa*
25	*Eruca sativa*
26	*Eruca sativa*
Columbia	*Arabidopsis thaliana*

Table 3.3. List of primers used for the mPCR *Brassica* identification (Koh et al., 2017).

Primer	Sequence (5'-3')	Length	Tm (°C)	Species	Chromosome	Gene	Genomic location	Product (bp)
A6-1	F: CCAGCGAAGGATTTGACGAC	20	59.3	*B. rapa*	A06	Bra019579	A06:12049397-120493649	*253*
	R: GACGAATCGAGTGCCCTG	18	57.9					
A6-2	F: GTTTTGGCCGTAAATCCCAC	20	57.6	*B. rapa*	A06	Bra019582	A06:12019296-12019481	*186*
	R: GTTACGGGTAGCGTGTGTC	19	58.3					
C1	F: TGCTGCGCCGAACAATAG	18	58.5	*B. oleracea*	C01	Bo1g016520	C1:5373673-5373829	*157*
	R: CCGATCGTGGTTCATATTGC	20	57.2					
C9	F: GTTAACGCACTTAAGGACCATG	22	57.7	*B. oleracea*	C09	Bo9g098720	C9:32340576-32341187	*612*
	R: ATTGACAACACCACCTCCCG	20	60.3					
B	F: GGCATCTGAAGAGAGAGTC	19	54.4	*B. nigra*	All	-	-	*331*
	R: CCATCTTCTTCTTGCCATG	19	53.7					

3.3 Results

For all the experiments, no disease was found after 14-17 dpi on the negative control (DB) inoculated with sterile water. On the other side, the positive control (DB) showed symptoms of *H. brassicae* infection at 3 dpi and was heavily sporulated after 7-10 dpi.

3.3.1 A12xGD DH population

Initial downy mildew resistance assessment of both parental lines of the A12xGD population showed moderate resistance to downy mildew at 7 and 14 dpi in the broccoli type (*B. oleracea* ssp *italica*) GDDH33 parental line and Downy Mildew susceptibility in the Chinese kale type (*Brassica oleracea* ssp *alboglabra*) A12DHd parental line at 7 and 14 dpi. No resistant lines were found in this population at 7 dpi or 14 dpi. At 7 dpi 32 lines were moderately resistant and 40 were susceptible to downy mildew. At 14 dpi only DH lines 2068 (Figure 3.6-a), 2208 and 5080 were considered moderately resistant to downy mildew with a DI% below 70. The other 70 lines showed a susceptible phenotype with a DI% above 71. DH line 1012, which is the genotype used as a model for brassica *Agrobacterium tumefaciens* mediated transformations had a susceptible phenotype after 14 dpi (Figure 3.6-b). At 7 dpi 30 lines that showed a moderate resistant phenotype to downy mildew showed a susceptible phenotype at 14 dpi (Figure 3.7; Figure 3.8). Significant differences were found when comparing the DI% average for all the genotypes in the population at 7 and 14 dpi (Figure 3.5). Therefore, having a second downy mildew assessment at 14 dpi after the early assessment of downy mildew resistance at 7 dpi was crucial to verify if a line was resistant to downy mildew.

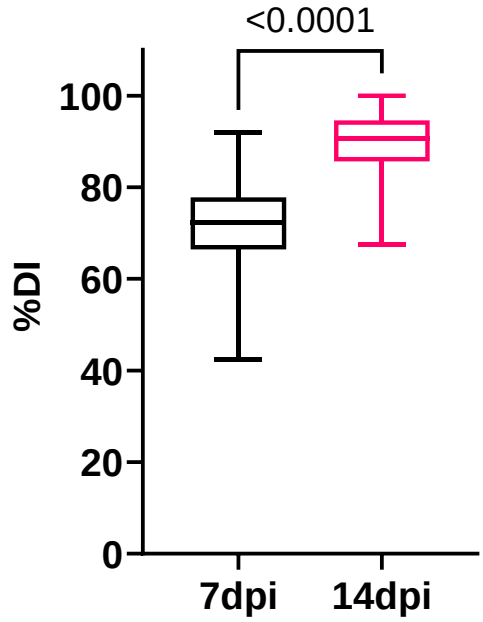


Figure 3.5. Boxplot illustrating the distribution of the DI% to Downy Mildew for the 75 genotypes in the A12xGD DH population at 7 and 14 dpi (P-value <0.0001).

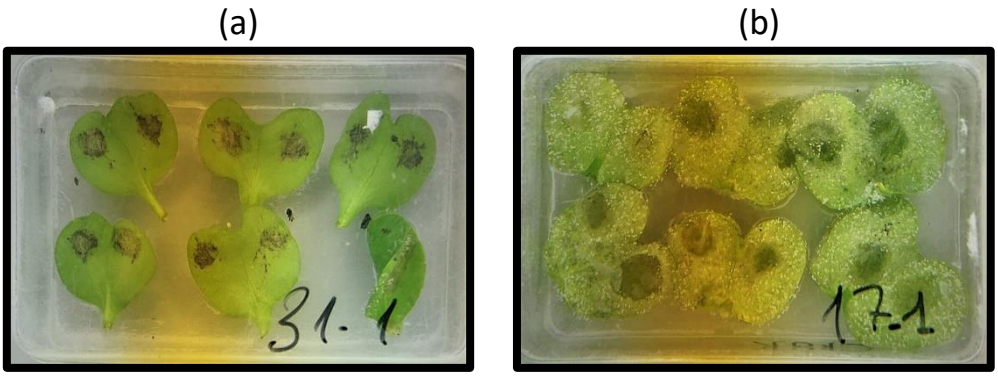


Figure 3.6. Illustrates the phenotypes of a moderate resistant and a susceptible genotypes to Downy Mildew after 14 dpi for visual comparison. (a) Moderate resistant DH line 2068. (b) Susceptible DH line 1012.

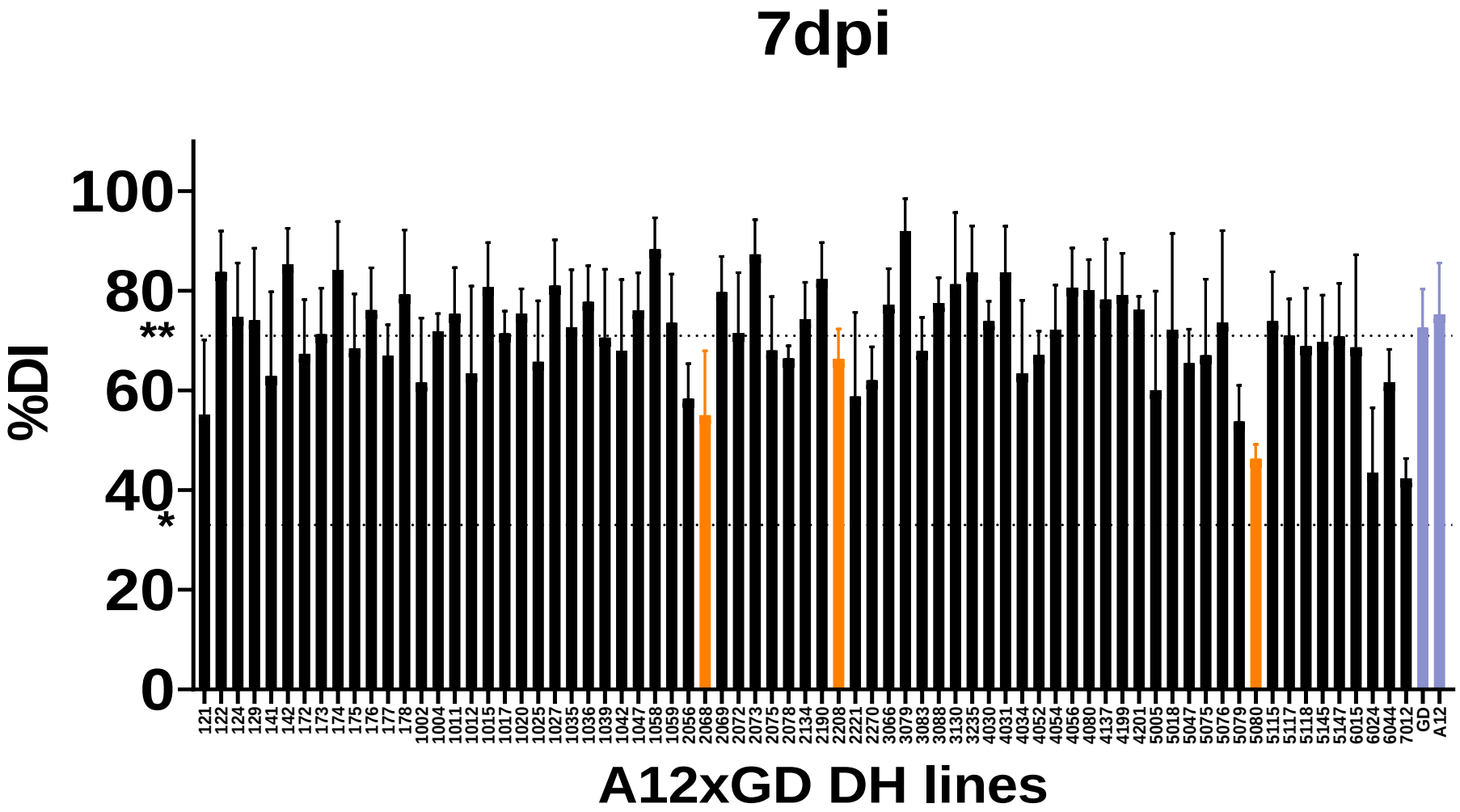


Figure 3.7. Downy Mildew levels of resistance for 75 genotypes of the A12xGD population, expressed by the disease index percentage of *H. brassicae* disease at 7 dpi. Parental lines are shown in violet-coloured bars and the most resistant genotypes in the population after 14dpi are in orange-coloured bars. *Limit between Resistant and moderately resistant genotypes to Downy Mildew. **Limit between moderately resistant and susceptible genotypes to Downy Mildew.

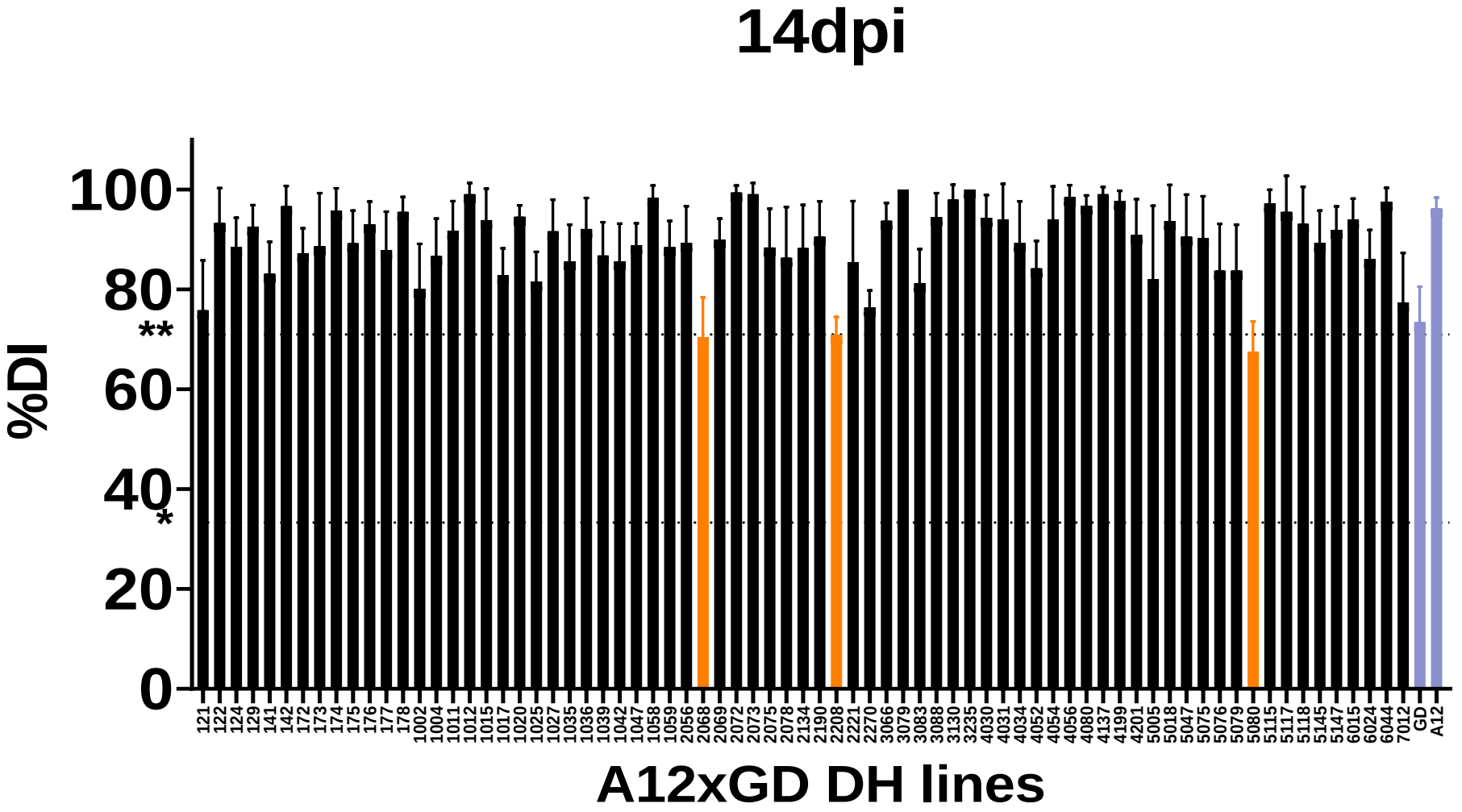


Figure 3.8. Downy Mildew levels of resistance for 75 genotypes of the A12xGD population, expressed by the disease index percentage of *H. brassicae* disease at 14 dpi. Parental lines are shown in violet-coloured bars and resistant genotypes in the population after 14dpi in orange-coloured bars. *Limit between Resistant and moderately resistant genotypes to Downy Mildew. **Limit between moderately resistant and susceptible genotypes to Downy Mildew.

3.3.2 BolDFFS DH set from Warwick GRU

A total of 58 fixed genotypes were assessed for Downy Mildew resistance at 7 and 14 dpi against an inoculum mix of 9 isolates from the UK and USA. This set was very biodiverse including several different *B. oleracea* cultivars (Appendix D2). Compared to the biparental population A12xGD (79) this seed set was very heterogeneous for Downy Mildew resistance phenotypes. Genotypes HRIGRU008658 (Waltham 29) and HRIGRU005108 (Kailan) had total resistance to Downy Mildew. None of these two genotypes showed Downy Mildew sporulation after 14 dpi, at this time point, HRIGRU008658 showed a bit of discolouration but no necrotic flecks at the points of inoculation and HRIGRU005108 showed necrotic flecks at the points of inoculation (Figure 3.9-a,b). The broccolo nero spigariello type line HRIGRU004885 was not considered resistant to Downy Mildew because very few sporangias were visible at the points of inoculation at 14 dpi, but the overall Downy Mildew resistance was very high compared to the other genotypes of the set (Figure 3.9-c). At 7 dpi a total of 15 genotypes were considered moderately resistant to Downy Mildew but after 14 dpi only 5 were moderately resistant. For susceptible genotypes, at 7 dpi 41 genotypes were susceptible and the number increased to 51 after 14 dpi (Figure 3.11; Figure 3.12).

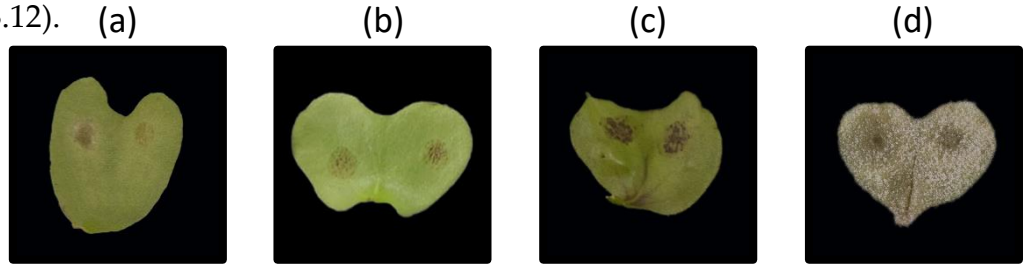


Figure 3.9. (a) Resistant genotype HRIGRU008658 (Waltham 29) at 14 dpi. (b) Resistant genotype HRIGRU005108 (Kai lan) at 14 dpi. (c) Moderate resistant genotype HRIGRU004885 broccolo nero spigariello at 14 dpi. (d) Susceptible genotype TO1000DH3 at 14 dpi.

Genotypes HRIGRU008658, HRIGRU005108 and HRIGRU004885 were also assessed for Downy Mildew resistance at the adult plant stage. All three were highly resistant to Downy Mildew at this stage. As previously mentioned, Downy Mildew was able to sporulate at the points of inoculation on genotype HRIGRU004885 but was not able to spread to the rest of the cotyledon. On leaves, no Downy Mildew sporulation was observed on any of the disks for this genotype. Necrotic flecks were observed for genotypes HRIGRU005108 and HRIGRU004885, but not on HRIGRU008658. Significant differences were found when comparing the DI% average for all the genotypes in the population at 7 and 14 dpi (Figure 3.10). Although the difference was not as big as the one observed in the A12xGD population (79) it was useful to verify the resistant or susceptible phenotype at a later stage.

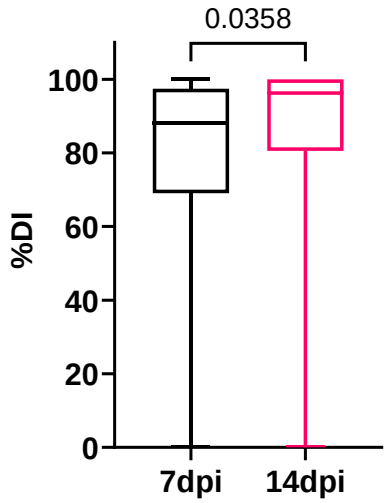


Figure 3.10. Boxplot illustrating the distribution of the DI% to Downy Mildew for 58 DH lines included in the *B. oleracea* diversity fix foundation set from Warwick at 7 and 14 dpi (P-value <0.0358).

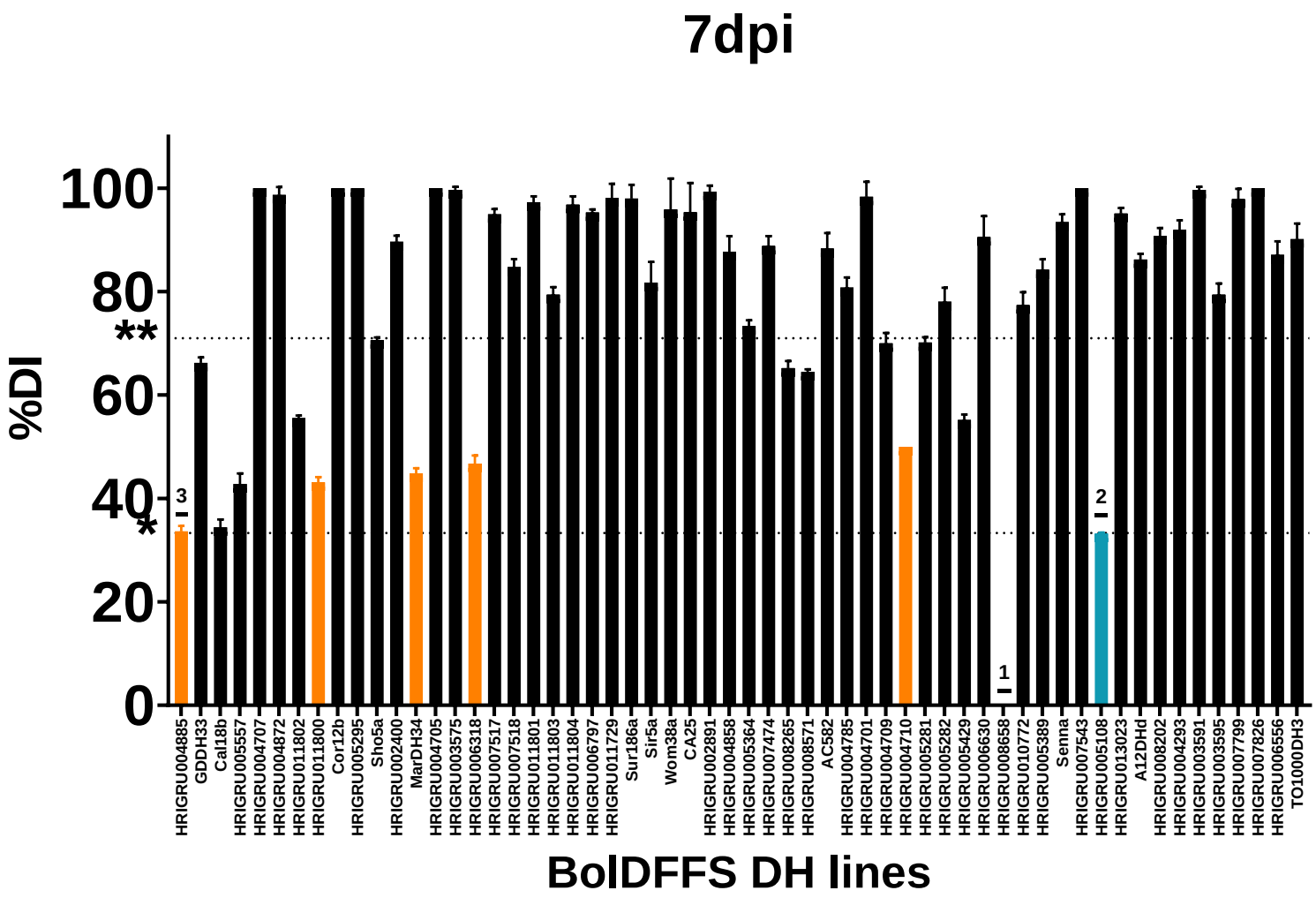


Figure 3.11. Downy Mildew levels of resistance for 58 DH lines included in the *B. oleracea* diversity fix foundation set from Warwick, expressed by the disease index percentage of *H. brassicae* disease at 7 dpi. (1 and 2) Resistant lines to Downy Mildew at 14 dpi coloured blue bars. (3) Moderate resistant lines to Downy Mildew at 14 dpi coloured orange bars. *Limit between Resistant and moderately resistant genotypes to Downy Mildew. **Limit between moderately resistant and susceptible genotypes to Downy Mildew.

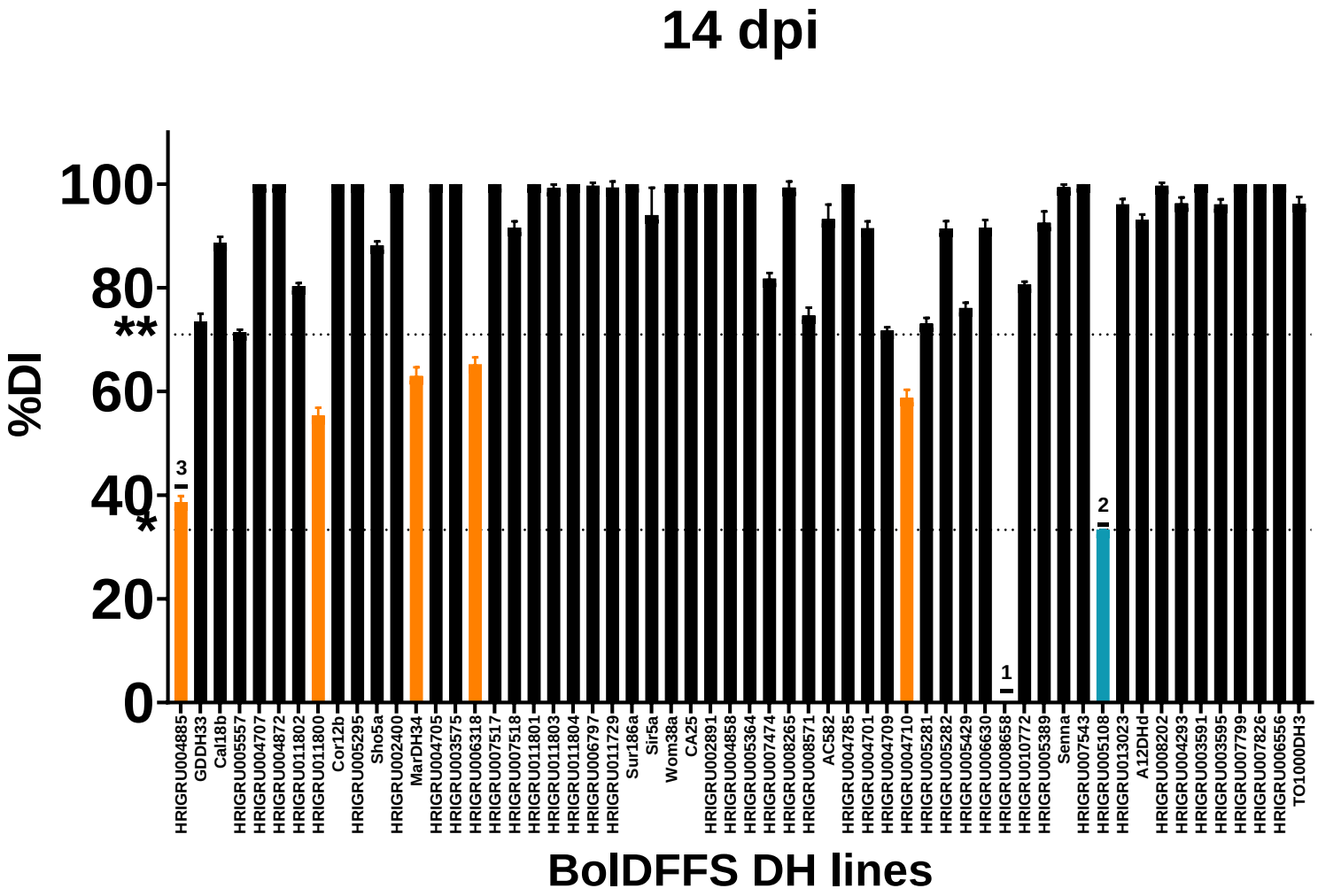


Figure 3.12. Downy Mildew levels of resistance for 58 DH lines included in the *B. oleracea* diversity fix foundation set from Warwick, expressed by the disease index percentage of *H. brassicae* disease at 14 dpi. (1 and 2) Resistant lines to Downy Mildew at 14dpi coloured blue bars. (3) Moderate resistant lines to Downy Mildew at 14dpi coloured orange bars. *Limit between Resistant and moderately resistant genotypes to Downy Mildew. **Limit between moderately resistant and susceptible genotypes to Downy Mildew.

3.3.3 *Brassica oleracea* accessions from the Warwick genetic resources unit and breeding lines

This Downy Mildew resistance assessment on *B. oleracea* accessions and breeding lines was performed at the CN Seeds facilities in Pymoor, Ely, UK. The CER where the plants were incubated after Downy Mildew inoculations were set at a lower temperature compared to the one used at the JIC facilities. Therefore the Downy Mildew scorings were adjusted using the universally Downy Mildew susceptible line DB from 7 and 14 dpi for the in-vitro assays at the JIC to 10 and 17 dpi at CN seeds achieving very similar phenotypes at both 7-10 dpi and 14-17 dpi.

A minimum of ten seeds were sown for all 390 accessions and 6 breeding genotypes. Due to the low emergence, accessions with four or fewer biological replicates (eight technical replicates) were eliminated from the final analysis. After 10 dpi, 106 accessions were highly resistant meaning no sporulation was observed, 159 accessions were moderately resistant and 67 accessions were susceptible. In comparison, after 17 dpi, the total number of highly resistant accessions was 28, 151 accessions were moderately resistant and 153 accessions were susceptible (Figure 3.14). After retesting the 28 resistant accessions against a mixture of all nine Downy Mildew isolates in vitro at the JIC following the methodology explained in Section 3.2.4.3 (71), all accessions showed an identical phenotype after 14 dpi *in vitro* compared to the one observed at 17 dpi in seedlings, so the change in the conditions of the CER didn't affect the observed phenotype.

Furthermore, as evidenced by the DI% values, there were overall varied ranges in comparative resistance responses to Downy Mildew depending on the country of origin after 17 dpi. From the top five countries with more

accessions (Portugal, UK, Spain, Japan and France), Portuguese accessions were the most moderately resistant on average with a DI% of 47.35. Spanish and British accessions were also moderately resistant to Downy Mildew on average with a DI% of 62.98 and 60.96 respectively. Japanese and French accessions were very susceptible to Downy Mildew on average with a DI% of 90.94 and 81.38 respectively (Figure 3.15).

Comparing this accession seed set with the other two with DH fixed lines, higher variability between replicates was observed. This observed phenotypic variability for Downy Mildew resistance could be due to genetic variability on the seed packages for some of these accessions.

As observed in the A12xGD DH population and the BolDFFS, significant differences were found when comparing the DI% average for all the genotypes in the set at both Downy Mildew assessment times at 10 and 17 dpi (Figure 3.13).

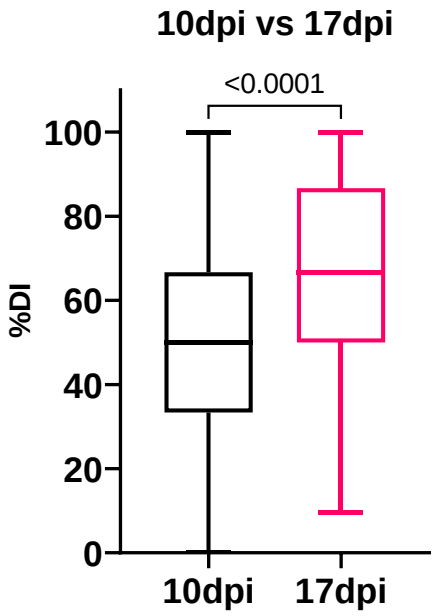


Figure 3.13. Boxplot illustrating the distribution of the DI% to Downy Mildew for 390 *B. oleracea* accessions from the Warwick GRU and 6 *B. oleracea* breeding lines from CN seeds (P-value <0.0001).

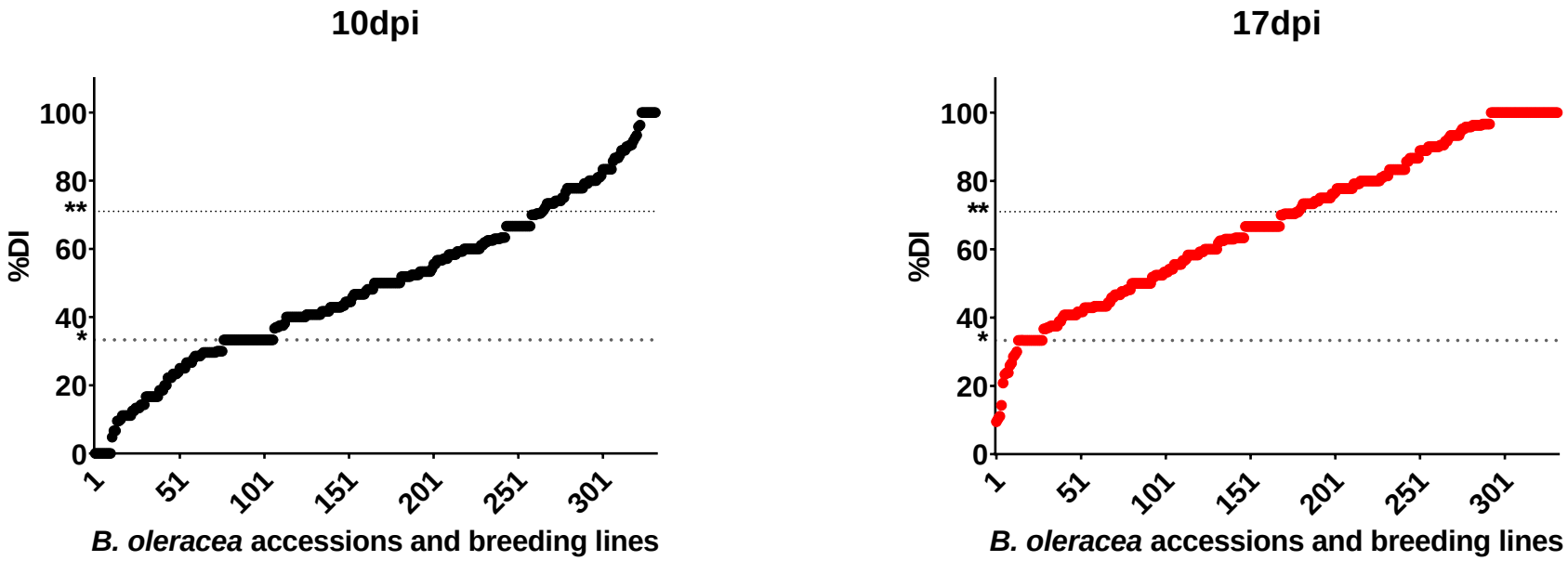


Figure 3.14. Disease index resistance percentage to *H. brassicae* at 10 and 17 dpi of 390 *Brassica oleracea* var. *acephala* accessions from the Warwick genetic resources unit (WIEWS instcode-GBR006) and six breeding lines from CN Seeds. *Limit between Resistant and moderately resistant genotypes to Downy Mildew. **Limit between moderately resistant and susceptible genotypes to Downy Mildew.

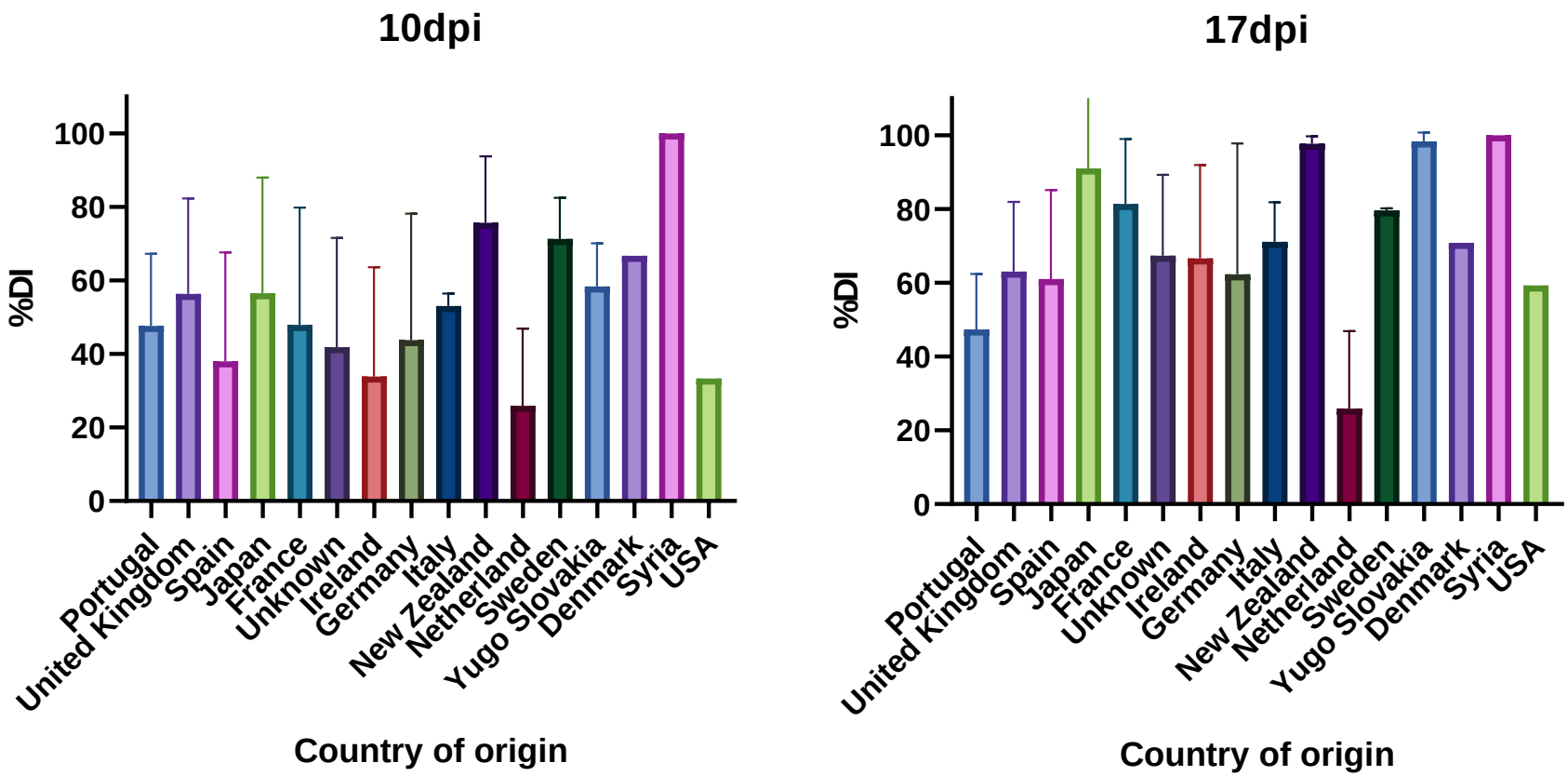


Figure 3.15. Bar chart illustrating the average and SD for the DI% to *H. brassicae* of all 396 *B. olerecea* genotypes tested at the cotyledon stage according to their country of origin (Portugal, United Kingdom, Spain, Japan, France, Unknown, Ireland, Germany, Italy, New Zealand, Netherland, Sweden, Yugo Slovakia, Denmark, Syria and USA). Countries were ordered based on the number of genotypes available.

3.3.4 Multiplex PCR assay for rapid *Brassica* identification

Downy mildew host specificity has long been reported for many isolates on different *Brassica* species (Byford, 1967; Gaumann, 1918; Gaumann, 1923; Lebeda & Syrovátko, 1988; Sherriff & Lucas, 1990). And there are only a few cases where a downy mildew isolate appears to have a large range of hosts (Kenneth, 1981; Runge et al., 2011). For this reason and the morphological variability observed for the individuals in some of the *Brassica oleracea* var. *acephala* accessions from the Warwick genetic resources unit. I tested all of them plus the six breeding lines from CN Seeds to verify if they were *B. oleracea*.

To achieve this I used a multiplex PCR assay designed by Koh et al. (2017). This PCR consists of 5 primers specific to all 3 brassica genomes (Table 3.3). After testing all 26 genotypes included in the *Brassica* panel (Table 3.2), this mPCR assay was able to identify correctly all 20 accessions from the *Brassica* panel producing distinct and discernible gel banding patterns. A small subset of the bands observed in a 2% (w/v) agarose in TBE gel for all 6 brassica genomes can be seen in Figure 3.16. No PCR product was detected in the *Eruca sativa*, *Diplotaxis erucoides* and *Arabidopsis thaliana* samples also present in the panel. Primers used for the reaction were specific to their corresponding genomes and no spurious amplification was detected on all gDNA samples. These results suggest that these PCR markers are specific to each *Brassica* A, B and C genome and can distinguish between diploid and polyploid *Brassica* species.

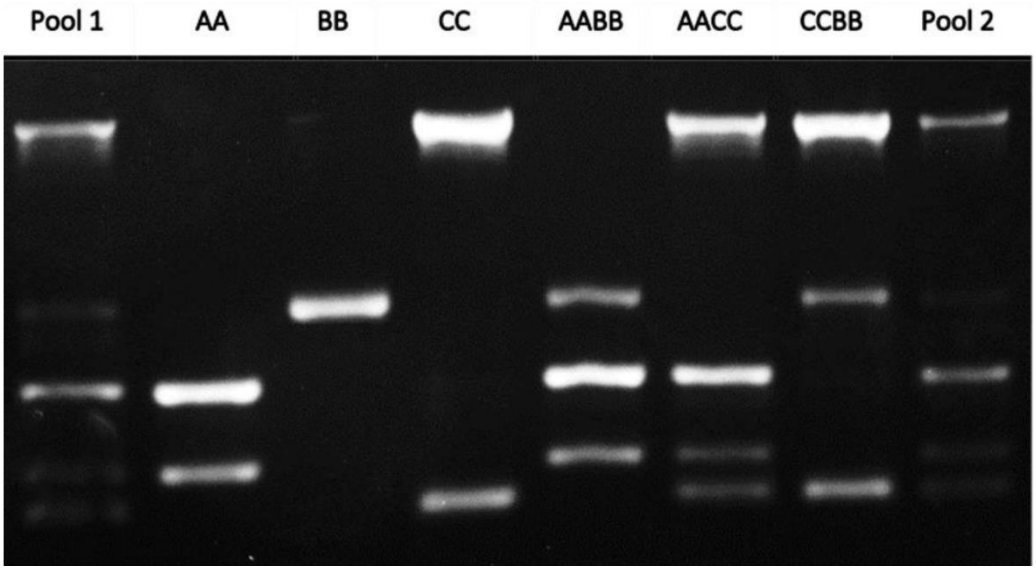


Figure 3.16. mPCR amplification of 6 control samples with known species identity plus 2 gDNA pools including DNA from all 6 *Brassica* species (Pool 1 and Pool 2). Diploid *Brassica* species: AA: *B. rapa*, BB: *B. nigra* and CC: *B. oleracea*. Polyploid *Brassica* species: AABB: *B. juncea*, AACC: *B. napus* and CCBB: *B. carinata*. PCR products were resolved on 2% TBE agarose gel.

From all 390 accessions, 5 accessions were identified as *B. rapa*, 3 as *B. napus* and the other 382 accessions were identified as *B. oleracea* which is the species they were initially labelled. All 6 breeding lines from CN seeds, including the resistant genotype 1028 were identified as *B. oleracea*.

Interestingly, all 8 incorrectly labelled accessions included in the WGRU material were classified as highly resistant to the UK Downy Mildew isolate from Colchester (Section 3.3.3-87) tested in previous screenings, denoting the high host specificity of the isolates collected from a *B. oleracea* host in the UK (Appendix D3).

3.3.5 Investigating Downy mildew host specificity of the H. *brassicae* single spore isolates collected from *B. oleracea* fields

To investigate the host specificity observed on the 8 non *B. oleracea* lines from section 3.3.4 (92) a total of 232 Brassicaceae genotypes were assessed for resistance against an inoculum mix of all 9 *H. brassicae* isolates collected originally from *B. oleracea* hosts on fields in the UK and USA. The plant material consists of the *Brassica* diversity panel used for the mPCR (Table 3.2) assay plus 12 extra genotypes for every *Brassica* species (Diploid: *Brassica oleracea*, *Brassica rapa* and *Brassica nigra*; Polyploid: *Brassica napus*, *Brassica carinata* and *Brassica juncea*) (Table 3.4). In total this new *Brassica* diversity panel consists of 98 genotypes, which were assessed using the mPCR assay from section 3.3.4 (92) to confirm the *Brassica* species that they belong to. In addition to these 98 genotypes, I also included an extra set of 134 *B. napus* genotypes but they were not assessed using the mPCR assay to confirm their *Brassica* species. All genotypes in the *Brassica* diversity panel (98) and the *B. napus* set (134) were assessed *in vitro* for downy mildew resistance following the same methodology used for the A12xGD DH and the BolD-FFS from section 3.2.5 (73), at separate times. A minimum of 20 seeds were sown and 10-16 cotyledons were assessed for every genotype at 7 and 14 dpi.

After 14 dpi, none of the 9 single spore downy mildew isolates included in the inoculum mix was able to sporulate on any of the lines shown in (Table 3.2) for *B. rapa, B. nigra, B. napus, B. carinata*, *B. juncea*, *E. sativa*, *D. erucoides* and *A. thaliana* genotype hosts included in the *Brassica* diversity set and they were considered as highly resistant to Downy Mildew. On the other side

sporulation was observed on all *B. oleracea* genotypes included in the *Brassica* diversity panel except for genotype 1028 which was considered highly resistant.

I increased the number of *B napus* lines in the screen to confirm whether all accessions were resistant. I screened an additional set of 134 lines and after 14 dpi 9 genotypes were susceptible to *H. brassicae* showing profuse sporulation. The other 125 genotypes showed a highly resistant phenotype and no sporulation was observed after 14 dpi. The resistant phenotype observed was identical to the resistant phenotypes on the other brassica species, meaning dark black/brown flecks at the two points of inoculation for each cotyledon tested and no sporulation. Although the 134 *B. napus* genotypes used in this study were previously used for breeding purposes, the mPCR assay should be performed at least on the susceptible genotypes to confirm which brassica species they belong. This experiment highlights the physiological specialization present in *H. brassicae* previously reported by many authors.

Table 3.4. Ampliation of the *Brassica* diversity panel (Table 3.2) to investigate *H. brassicae* host specificity.

Genome	Species	Accesion name	Crop type	Origin	Source
AA	*Brassica rapa*	Brocaletto 372	Brocaletto	Italy	JIC
AA	*Brassica rapa*	Chi Hi Li	Chinese Cabbage	China	UKVGB, Warwick
AA	*Brassica rapa*	Mizuna 371	Mizuna	Japan	JIC
AA	*Brassica rapa*	Sonja	Oilseed	UK	JIC
AA	*Brassica rapa*	Pak Choi 375	Pak Choi	China	JIC
AA	*Brassica rapa*	Lange weisse Rotkopfige	Turnip	Germany	UKVGB, Warwick
AA	*Brassica rapa*	Zeltomiannaja	Unknown	Russia	VIR, St. Petersburg
AA	*Brassica rapa*	R500	Yellow sarson	USA	JGI
AA	*Brassica rapa*	genes (99)	Unknown	Unknown	JIC
AA	*Brassica rapa*	Ryou	Pak Choi	China	CN Seeds
AA	*Brassica rapa*	Komatsuna 9490F	Komatsuna	Japan	Johnsons Seeds
AA	*Brassica rapa*	Tat soi 9489/D	Tat soi	China	Johnsons Seeds
BB	*Brassica nigra*	Turkey 1127	Black Mustard	Turkey	JIC
BB	*Brassica nigra*	Ethiopia 1142	Black Mustard	Ethiopia	JIC
BB	*Brassica nigra*	India 1119	Black Mustard	India	JIC
BB	*Brassica nigra*	Puerto rico 1131	Black Mustard	Puerto rico	JIC
BB	*Brassica nigra*	Ethiopia 1130	Black Mustard	Ethiopia	JIC
BB	*Brassica nigra*	USA 1152	Black Mustard	USA	JIC
BB	*Brassica nigra*	India 1156	Black Mustard	India	JIC
BB	*Brassica nigra*	Yugoslavia 1149	Black Mustard	Yugoslavia	JIC

BB	*Brassica nigra*	Afganistan 1132	Black Mustard	Afganistan	JIC
BB	*Brassica nigra*	Yugoslavia 1145	Black Mustard	Yugoslavia	JIC
BB	*Brassica nigra*	Pakistan 1151	Black Mustard	Pakistan	JIC
BB	*Brassica nigra*	Poland 1147	Black Mustard	Poland	JIC
CC	*Brassica oleracea*	Purple Vienna	Kohl rabi	Austria	Mr Fothergill's
CC	*Brassica oleracea*	Green Curled	Kale dwarf	Unknown	Thompson and Morgan
CC	*Brassica oleracea*	All Year Round	Cauliflower	Unknown	Thompson and Morgan
CC	*Brassica oleracea*	Walcheren	Cauliflower	Unknown	Mr Fothergill's
CC	*Brassica oleracea*	Robinsons Champion	Giant cabbage	Unknown	W. Robinson and Son
CC	*Brassica oleracea*	AAA	Kale	Unknown	CN seeds
CC	*Brassica oleracea*	January King 3	Savoy cabbage	Unknown	Thompson and Morgan
CC	*Brassica oleracea*	Nero di Toscana	Kale nero	Italy	Thompson and Morgan
CC	*Brassica oleracea*	Rubine	Brussel sprout	Europe	Mr Fothergill's
CC	*Brassica oleracea*	Summer Purple	Broccoli	Unknown	Thompson and Morgan
CC	*Brassica oleracea*	Romanesco	Broccoli romanesco	Italy	Thompson and Morgan
CC	*Brassica oleracea*	A12	Kai lan	China	JIC
AABB	*Brassica juncea*	J136	Mustard	Russia	USDA
AABB	*Brassica juncea*	J151	Mustard	Turkey	USDA
AABB	*Brassica juncea*	Poorbi raya	Mustard	Pakistan	USDA
AABB	*Brassica juncea*	Canaklijski	Mustard	Poland	USDA
AABB	*Brassica juncea*	Blaze	Mustard	Canada	USDA
AABB	*Brassica juncea*	Pan xi da Zhai you Cai	Mustard	China	USDA
AABB	*Brassica juncea*	J149	Mustard	India	USDA

AABB	*Brassica juncea*	BJ58	Mustard	India	UKVGB, Warwick
AABB	*Brassica juncea*	BJ16	Mustard	Germany	USDA
AABB	*Brassica juncea*	K-1072	Mustard	Afghanistan	USDA
AABB	*Brassica juncea*	J153	Mustard	Puerto Rico	USDA
AABB	*Brassica juncea*	Japanese greens	Greens	Japan	UKVGB, Warwick
BBCC	*Brassica carinata*	Jimma kefa 5002 5	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Gonclar bagem 5002 16	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Bekedjamor 5002 5	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Kuni harege 5002 1	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Rabi bade 5002 21	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	HAYK welo 5002 40	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Addis 1212	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Sao tome isle 5002 47	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Negelli 5002 99	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Sire Arhus 5002 26	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	Eth chambre 1213	Mustard	Ethiopia	JIC
BBCC	*Brassica carinata*	5003 172	Mustard	Ethiopia	JIC
AACC	*Brassica napus*	Ningyou 1	Semi-winter oilseed rape	China	OCRI, Wuhan
AACC	*Brassica napus*	Canard	Winter forage rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	Tapidor	Winter oilseed rape	France	UKVGB, Warwick
AACC	*Brassica napus*	Jet Neuf	Winter oilseed rape	France	UKVGB, Warwick
AACC	*Brassica napus*	Lesira	Winter oilseed rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	N02D-1952	Spring oilseed rape	Australia	UKVGB, Warwick

AACC	*Brassica napus*	Topas	Spring oilseed rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	Palmedor	Winter oilseed rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	Ningyou 7	Semi-winter oilseed rape	China	OCRI, Wuhan
AACC	*Brassica napus*	Temple	Winter oilseed rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	Westar	Winter oilseed rape	Unknown	UKVGB, Warwick
AACC	*Brassica napus*	Tina	Swede	Unknown	UKVGB, Warwick

3.4 Discussion

This study identified several *B. oleracea* genotypes from three sources of plant material with very high pathotype-independent resistance to *H. brassicae* from six different locations in the UK and three in the USA. These identified resistant genotypes were included in the CN seeds breeding programs to create new downy mildew-resistant kale varieties and also to develop F_2 and BC_1 segregant populations for further analysis. Apart from the three resistant phenotypes from the BolDFFS which were tested at the adult plant stage, this study mainly focused on cotyledon tests. Nevertheless, resistance expression on cotyledons is generally very similar to that on leaves in lab conditions (Wang et al., 2000) or field tests (Jensen et al., 1999; B. Jensen et al., 1999; Monteiro & Williams, 1989; Natti, 1958), therefore using cotyledons to assess Downy Mildew resistance at the seedling stage is a good indicator of overall resistance at the adult plant stage.

The higher variability observed on the accession plant material compared to the fixed DH genotypes from the A12xGD and the BolDFFS could be explained by genetic variability on the seed packages. Seed banks usually bulk accessions by placing several plants of a similar assumed genetic background in cages allowing them to fertilize each other. If these accessions are not screened individually before placing them in the cages, some of the accessions could be a mix of different genotypes or in some cases even different species if the family or genre have very similar morphological attributes. To investigate this further, I used an mPCR assay designed by Koh et al. (2017) to identify all the accessions included in the plant material at the species level. This assay identified 8 of the 28 downy mildew-resistant accessions from section 3.3.3 (87) to be a different brassica species, 5 of which

were identified as *B. rapa* and 3 as *B. napus*. This data suggested that the physiological specialization previously reported by many authors is also prevalent in the isolates used for this thesis. Although DNA for the mPCR assay was extracted from 3 different individuals for every accession in the accession plant material, other brassica species apart from *B. oleracea* could be mixed in small quantities or the identified *B. napus* accessions could be a mix of a *B. oleracea* and *B. rapa* genotypes. Therefore, this physiological specialization could explain the phenotypic variability previously observed for the downy mildew disease assays.

No other *Brassica* species apart from *B. oleracea* was identified as being susceptible to downy mildew in the accession plant material. I investigated the *H. brassicae* physiological specialization by infecting a *Brassica* diversity panel consisting of 98 different genotypes of all the *Brassica* species in the triangle of U and close relatives (16 *B. oleracea*, 16 *B. rapa*, 15 *B. nigra*, 15 *B. napus*, 15 *B. juncea* and 15 *B. carinata*; 2 *D. erucoides*, 3 *E. sativa* and 1 *A. thaliana*), and a set of 134 *B. napus* genotypes against an inoculum mix of all 9 single spore isolates from the UK and USA. After 14 dpi only 15 *B. oleracea* genotypes and 8 *B. napus* genotypes were identified as susceptible out of the 232 genotypes tested. The rest of the genotypes from all brassica species and close relatives showed a very resistant phenotype against the *H. brassicae* isolates mix, all collected from *B. oleracea* hosts in the UK and the USA.

These results highlight the physiological specialization present in *H. brassicae* for isolates collected from specific *Brassica* hosts, in this thesis *B. oleracea*. If, like in this study, some isolates can infect other brassica hosts, these new sources of resistance present in other *Brassica* hosts could be used to improve commercial varieties on other Brassicas species because all six brassica species share the same 3 diploid genomes. The susceptible genotypes

to *H. brassicae* found in this study are from *B. oleracea* and *B. napus* genotypes, which both share the C brassica genome. Therefore, to investigate *H. brassicae* host specificity further, more genotypes should be included, especially from *B. rapa, B. nigra, B. carinata* and *B. juncea* in order to test if they can also infect and sporulate on this *Brassica* species or if this physiological specialization observed for this specific isolates is linked to the three brassica genomes.

As observed on all three plant materials, having a second Downy Mildew assessment at 14 or 17 dpi after the early assessment of resistance at 7 or 10 dpi was very useful to verify if a line was resistant to Downy Mildew (Figure 3.5; Figure 3.10; Figure 3.13). Many genotypes had this latent period when the visible damage caused by Downy Mildew compared to other genotypes was much lower at the first time of assessment, but were finally overcome by the pathogen at the second time of assessment. Therefore the more interesting phenotypes are the ones that are resistant or moderately resistant consistently over a long period, in this case of 14 or 17 dpi.

Pathotype independent resistances are extremely valuable for breeders and farmers to fight effectively against all different pathotypes of *H. brassicae* present in the fields in all *Brassica* growing countries. These identified Downy Mildew host resistances provide great potential to improve the management of this disease in the fields of many *B. oleracea* crops, especially vegetable crucifers. In the short term, some of these resistant genotypes could be directly deployed as new commercial varieties, if they are in genotypes with agronomically adequate backgrounds. In the long term, these genotypes could be used as a source of Downy Mildew resistance in breeding programs to produce new resistant commercial varieties or to gene pyramid several of them in a single more durable resistant hybrid variety. Even

though nine different Downy Mildew isolates from two countries were used to test all the genotypes, there is the possibility that some other isolates from other regions could be able to overcome the qualitative resistance observed in these individuals. Therefore, another approach could be to use some of the most interesting genotypes identified to have moderate resistance to Downy Mildew as a quantitative source of resistance. This could be used in combination with one or more qualitative ones to increase the non-race specificity and achieve a long-lasting resistance to Downy Mildew in the fields.

Another approach to managing Downy Mildew in fields is using chemical products such as pesticides, however, they only give irregular control when applied to genotypes with limited resistance to Downy Mildew (Barbetti et al., 2011). On top of this, for many leafy *B. oleracea* cultivars where the actual product is the immature plant, like baby leaves, there is a very short time frame where the pesticides can be applied, therefore the use of agrochemicals is not usually recommended. For that reason, there is an immediate need to use new sources of Downy Mildew resistance to develop and commercialize new resistant varieties to Downy Mildew. Finally, at least some of the resistances revealed in the current investigation across all the materials tested may comprise new sources and/or kinds of host resistance not previously recognized.

Chapter 4

Mapping a new dominant resistance to *Hyaloperonospora brassicae* by Bulk-Segregant Analysis on Chromosome 2 of *Brassica oleracea* L.

4.1 Introduction

The use of genetically resistant plants is one of the most successful, low-cost, and environmentally friendly techniques for plant disease control. Therefore, resistance to plant pathogens such as the oomycete *Hyaloperonospora brassicae*, also known as downy mildew, is one of the most desired traits in *Brassica* breeding programs. The identification of new sources of

resistance, mapping and cloning of useful QTLs and candidate genes would help *Brassica* breeders improve their commercial varieties against downy mildew in the fields. The availability of next-generation long-read sequencing and affordable Illumina short-read sequencing has created new opportunities to develop large amounts of genomic data and resources, all of which make it easier to map and clone new sources of resistance to many pathogens, including downy mildew. These new sources of resistance could then be used to improve the efficiency and precision of conventional brassica breeding via Marker Assisted Selection (MAS) and allow them to gene pyramiding different sources of resistance to create superior and more durable brassica cultivars.

In this chapter, an F_2 population created using one of the resistant genotypes identified on the plant materials screened for downy mildew resistance in Chapter 3 was selected to further investigate the observed resistance. This F_2 population was derived from a cross between Downy Mildew resistant genotype 1028 (Kale variety - CN seeds) and the susceptible commercial variety Dwarf Blue (Kale variety - CN seeds) and it was selected to map the resistance to Downy Mildew observed in the 1028 genotype using a Bulk Segregant Analysis approach. A small region containing an NLR cluster was identified to be linked with the resistance to Downy Mildew observed on the 1028 genotype in chromosome 2 of the pangenome *B. oleracea* assembly. A *de novo* draft genome assembly using long reads sequencing of the 1028 genotype revealed a copy-number variation of the NLR genes between the pangenome, HDEM and this draft assembly. Using this 1028 draft genome assembly KASP markers highly linked with the Downy Mildew resistance were designed for MAS. These markers will be very useful for brassica breeders looking to incorporate new sources of downy mildew

resistance into new *B. oleracea* commercial varieties or to incorporate them into new ones.

4.1.1 Segregating F_2 populations for genetical mapping

An F_2 segregating population is a genetically diverse progeny from a cross between two parental genotypes that differ for one or more genetically controlled traits. To discover and map which regions are associated with a specific trait this type of segregant population is normally used. To create an F_2 segregant population, the first step is to cross two parental lines with different alleles. Assuming homozygosity for these alleles, parent A has a dominant resistant gene "R/R" (1028 resistant genotype), and parent B has a recessive "s/s" (DB susceptible genotype). For this type of cross, all F_1 segregant individuals will have a combination of both alleles "R/s" and all of them will express the dominant trait because the recessive allele is masked by the dominant allele (Mendelian segregation). After this, one or more F_1 individuals are self-pollinated to produce an F_2 segregant population that will segregate in a 1:2:1 (RR:Rs:ss) or 3:1 (Resistant:Susceptible) ratio. When a trait is mapped to a specific genomic region, genetic molecular markers that detect polymorphism between different genotypes or alleles can be used by scientists and breeders to test if a certain allele and subsequently a trait is present in a particular individual without the need to phenotype it.

4.1.2 Bulk Segregant Analysis (BSA)

Bulk Segregant Analysis (BSA) is a molecular technique used to identify variants (genetic markers) associated with a trait by bulking the DNA of specific individuals from a segregant population that differ for that specific trait (Giovannoni et al., 1991; Michelmore et al., 1991). Using this approach one can find a region/s that is overrepresented or enriched in the respective

bulk by merging many independent individuals with the same trait. This allows the identification of the causative variants in the gene or locus of interest. In a single dominant gene segregant population, the region associated with the dominant trait will be enriched for the allele present on both parental genotypes, meaning it will be homozygote for the susceptible allele in the susceptible bulks and a mix of heterozygote and homozygote for the resistant allele. Regions not associated with the trait will be heterozygotes on both bulks. Depending on the complexity of the trait mapping by BSA approach can be done with different NGS methods like RNASeq (R. H. Ramirez-Gonzalez et al., 2015), DNA whole genome sequencing (This chapter - section 5.3.4-123), exome capture (Martinez et al., 2020), or a combination (Sharma et al., 2021).

To find variants associated with a trait in a diploid organism, variants must be called from NGS data on both contrasting bulks (Resistant and Susceptible) against a reference genome, preferably contiguous. As can be seen in Figure 4.1., first, the SNP-Index (Takagi et al., 2013) is calculated by dividing the times a specific allele is observed on a genomic position (allele count) by the sum of times both alleles have been observed for that position (coverage). The SNP index is calculated for both the Alternative and Reference alleles. And second, depending if the reference is carrying the allele for the trait, the SNP index for the alternative or the reference will be used to calculate the bulk frequency ratio. Variants with high BFR are linked with the trait being mapped (Trick et al., 2012). Neighbouring regions for the variants with high BRF are then examined using a genome browser with the alignments for all the bulks to look for possible candidate genes and structural variations that could be responsible for the trait. As mentioned in section 5.3.4 (123), a limitation of this approach is that for variants where the

allele counts for one of the bulks is zero due to no escapes in the bulks or because of a region or gene that is not present in one of the bulks, the BFR can't be calculated and therefore that specific variant won't be included in the analysis and the Manhattan plots, which are graphical ways of demonstrating associations between traits and genes on a chromosome. For this reason, careful analysis of all the variants after calculating the BFR is necessary to understand the reason why that is happening and if this could be a limitation or not for our specific analysis.

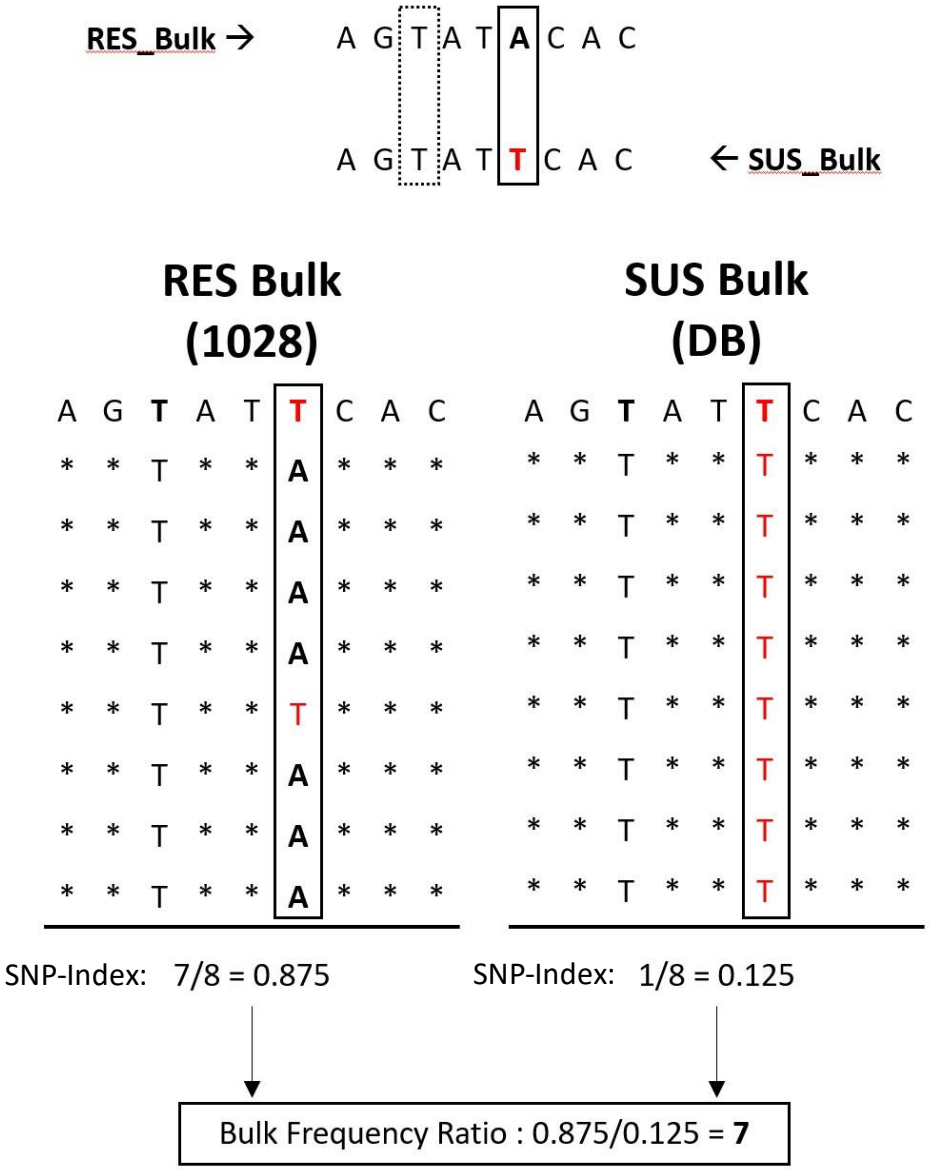


Figure 4.1. Schematic representation for the calculation of the Bulk Frequency Ratio for an SNP in a diploid species like *Brassica oleracea*. The figure represents the reads aligned to a reference genome for the resistant and susceptible bulks. The rectangle box marks an SNP for the recessive allele (T) present in the susceptible parent that matches the reference and the SNP for the allele (A) carrying the allele from the resistant parent. SNP-Index is calculated by dividing the allele count for the alternative resistant allele (to the reference) by the allele coverage for that SNP position. BFR is calculated by dividing the resistant bulk SNP-Index by the susceptible SP-Index.

4.2 Material and Methods

Some sections in this chapter required the use of some standard tools. All the code used for this chapter is available on the GitHub repository: AbrahamGoGu/PhDThesis_2022 (github.com)

4.2.1 F_1, F_2 and BC_1 Downy Mildew disease assessment

F_1, F_2 and BC_1 individuals derived from a cross between the resistant Downy Mildew breeding line 1028 (CN seeds) and the susceptible Downy Mildew commercial variety Dwarf Blue (CN seeds) were assessed for Downy Mildew resistance at 7 and 14 dpi following the methodology described in Chapter 3. If sampling was needed for DNA extraction, both cotyledons were assessed for every individual *in vitro* at separate times. The first cotyledon was assessed for Downy Mildew resistance at 14 dpi and the second cotyledon at 21 dpi. All cotyledons were placed in individual sealed small plates (Figure 4.2) containing water agar with benzimidazole (6g Agar + 900 mL distilled water + 100 mL Benzimidazole [1g/L]). The reasons to do it this way were: one, by detaching a single cotyledon instead of both at the same time, the seedlings could still grow at a normal pace; two, by detaching the cotyledons and assessing the resistance to downy mildew *in vitro*, tissue from susceptible seedlings was not contaminated and grow speed was not decreased by a systemic downy mildew infection; three, two technical replicates were available to verify the phenotype for every seedling. If the phenotype was not confirmed with the second cotyledon, the individual was not used for further DNA extraction and analysis.

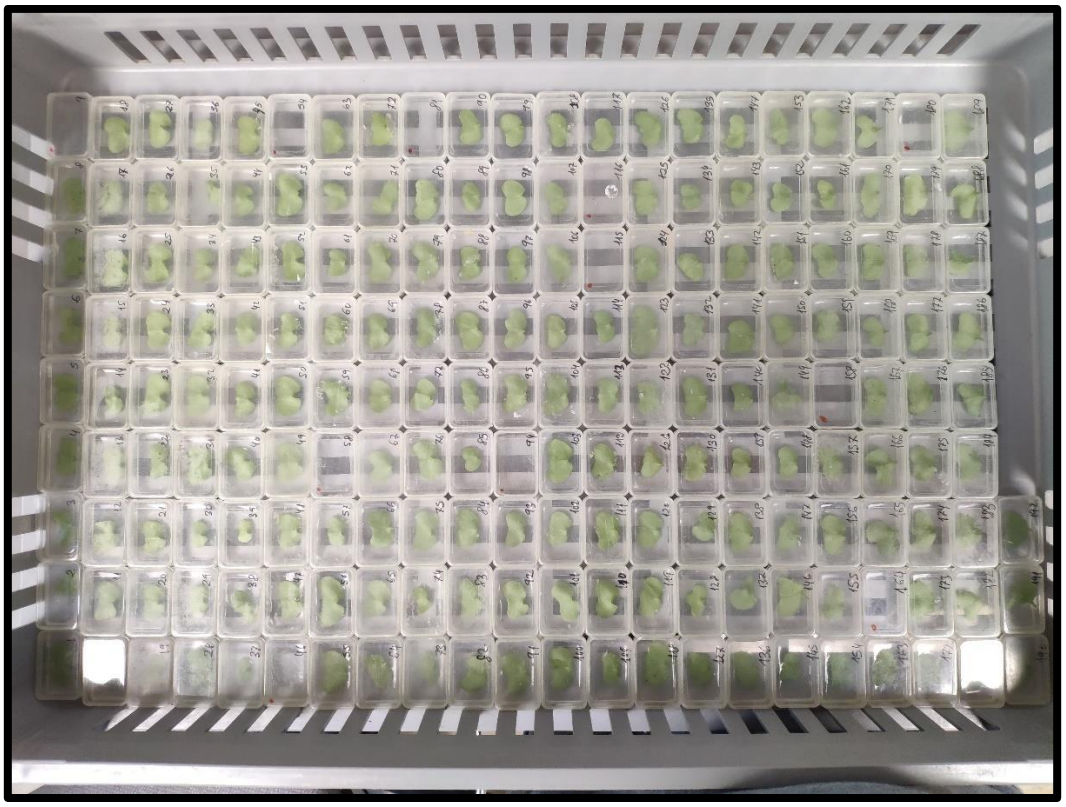


Figure 4.2. First cotyledons from F_2 individuals were in small sealed plates with water agar and benzimidazole before Downy Mildew inoculation.

Both parental lines were assessed for Downy Mildew resistance at the adult plant stage. A total of six plants per genotype were inoculated with all 9 Downy Mildew isolates (Table 2.1) by spraying both sides of the leaves at the 6-leaf stage. Plants were watered and incubated on sealed trays with transparent lids at 15°C with a day length of 12h for 14 days. Plants were assessed for Downy Mildew resistance at 7 and 14 dpi.

4.2.2 Bulk creation and high-quality DNA isolation for Illumina sequencing

The phenotypic data of the F_2 population was used to produce six bulks, three susceptible and three resistant to Downy Mildew. To ensure that a

similar amount of DNA is present for every individual in the bulk, 0.05 g of leaf tissue from five-week-old F_2 individuals were sampled and snap freeze in liquid nitrogen. For every bulk three 50 mL, falcon tubes were produced. Plant material from both parental genotypes was collected at the same time. DNA was isolated from all six bulks and both parental lines using an in-house protocol kindly supplied by Rachel Wells and created by Lorelei Bilham "*DNA preparation from Brassica for Illumina Sequencing*" (Appendix A). Alongside the bulks, DNA from both parental lines was isolated. After DNA isolation, every DNA sample was measured using nanodrop, qubit BR and run on 1% agarose gel to visualize for possible degradation and/or RNA contamination.

4.2.3 DNA Illumina sequencing

All six DNA bulks and the Downy Mildew susceptible parent DB were sent to NOVOGENE for whole-genome sequencing using Illumina, PE150 with consistent coverage of 100x (50Gb output) for every bulk and 10x (5Gb output) for the parental line. None of the bulks had more than 60 individuals in them which ensures that each DNA sample was sequenced between 1-2 times per bulk.

4.2.4 Sequencing reads cleaning and Quality Control

Trimmomatic (Bolger et al., 2014) was used to clean the Illumina reads and remove the Illumina paired-end adaptor sequences from every sample. The quality of the sequencing was assessed with FastQC v0.11.9 (Babraham-Bioinformatics, 2012).

4.2.5 Alignment DNA reads to reference genomes

Illumina DNA reads were aligned using BWA v0.7.12 (Li & Durbin, 2009)

to reference and draft genomes (Belser et al., 2018; Golicz et al., 2016; Parkin et al., 2014). Using Samtools v1.9 (Li et al., 2009), alignment SAM files were sorted and stored as a BAM file per sample against a given reference, which allows compression and random access to the alignments.

4.2.6 Variant calling and Variant filtering

Freebayes v1.3.1 (Erik Garrison, 2012) was used to call variants for every reference genome using the alignment "sorted.bam" files. The tool was run in parallel using a custom Python script for all the samples at the same time by dividing the reference genome into 100000bp chunks and processing them independently using different cores in the cluster. This improved the calling efficiency and decreased the total run time in the cluster. Variants called using all six bulks were stored in a single VCF file for every reference genome used.

A hard filter was applied for all VCF files using bcftools v1.8 (Danecek et al., 2021). Filtering out variants with a depth of <40, a quality score of <1000 and variants that were identical on both, the Resistant and Susceptible bulks.

4.2.7 Bulk Frequency Ratio calculation and mapping

The bulk frequency ratios (BFRs) were calculated using the methodology described by Trick et al. (2012). Using R studio, information from all variant calls from the three Resistant bulks and the three Susceptible bulks were combined into a single Resistant and Susceptible bulk. BFRs were calculated for each variant call by dividing the frequency of the reference variant in the Resistant bulk by the frequency of the reference variant in the Susceptible bulk, and the same for the alternative variant. A database with the BFRs

and the QUAL, DEPTH, REFERENCE ALLELE COUNT (RO) AND ALTERNATIVE ALLELE COUNT (AO) for all the variants and their position in the reference chromosomes was created to navigate the regions easily and find the most significant markers associated with Downy Mildew resistance in the F_2 population.

Using R studio, ggplot2 was used to plot chromosome-level Manhattan plots with all the BFRs values for each variant position in the *B. oleracea* genome.

4.2.8 KASP primer design and assay

All KASP markers were designed using the tool PolyMarker (Ricardo H. Ramirez-Gonzalez et al., 2015). To use references not available on the online tool, PolyMarker had to be installed and run on the cluster. All the primers include the standard FAM or HEX compatible tails (FAM tail: 5′ GAAGGTGACCAAGTTCAT-GCT 3′; HEXtail: 5′ GAAGGTCGGAG-TCAACGGATT 3′) and were ordered from Sigma Aldrich.

To run the KASP assay, DNA was extracted from all F_2 individuals previously phenotyped for Downy Mildew resistance (110) and both parental lines. DNA isolation and KASP validation for the SNPs were performed by Richard Goram from the genotyping service platform at the John Innes Centre.

4.2.9 NLR annotator

NLR-annotator tool was run on the reference genomes in order to annotate and predict possible loci associated with NLRs in all reference genomes used in this chapter (Steuernagel et al. (2018); Steuernagel et al. (2020); GitHub - steuernb/NLR-Annotator: NLR-Annotator upload). This tool

searches for amino acid motifs within all 6 frames of a DNA sequence. After this, an NLR locus is defined from the first to the last motif that could be associated with a possible NLR.

4.2.10 RNA sequencing and quality control

Four RNA samples isolated from fourteen-day-old seedling cotyledons from the Downy Mildew susceptible parental line DB and Resistant parental line 1028 were sent for RNA sequencing. A total of six cotyledons were pooled and snap-frozen in liquid nitrogen from six different seedlings per sample. These samples were ground into a fine powder using a mortar and pestle while keeping the cotyledons frozen using liquid nitrogen. RNA was extracted using the RNEasy plat mini kit from QIAGEN® according to the manufacturer's instructions. After RNA isolation all samples were treated with TURBO DNA-free™ kit from INVITROGEN according to the manufacturer's instructions. Nanodrop was used to measure the final RNA concentration and contamination of the samples. To check for RNA degradation samples were run on 1% agarose gel.

Non-degraded or non-contaminated RNA samples were sent for sequencing to Novogene (Cambridge, UK). Sequencing was performed using the Illumina HiSeq platform PE 150 for 20 million paired-end reads (6 Gb) per sample.

4.2.11 Alignment of RNA paired reads to reference genomes

RNAseq paired reads were aligned using HISAT2 to *B. oleracea* assemblies. Using Samtools v1.9 (Li et al., 2009), alignment SAM files were sorted and stored as a BAM file per sample against a given reference, which allows compression and random access to the alignments.

4.2.12 Resistant parental line 1028 Draft Genome assembly

High molecular DNA isolated from the resistant parental line 1028 was sent to the Earlham Institute (Norwich, UK) for long-read whole-genome sequencing using Pacbio HiFi technology, aiming for 18X coverage and 15kb insert sizes assuming a 620Mb *B. oleracea* genome size.

4.2.13 Parental line 1028 Draft Genome annotation

The pipeline used to identify gene models in the 1028 draft genome was provided by Matthew Moscou (TSL). Both samples were aligned to the draft genome using HISAT2 (Kim et al., 2015) and gene models were identified and merged using Cufflinks (Trapnell et al., 2010). Finally, (TransDecoder) was used to identify candidate coding regions using the merged transcript sequences.

4.3 Results

4.3.1 F_2 Mapping population

4.3.1.1 Cotyledon Downy Mildew resistance assessment and bulks creation

The F_2 mapping population was developed by crossing the resistant Downy Mildew breeding line from CN seeds 1028 and the susceptible Downy Mildew commercial variety Dwarf Blue from CN seeds (Figure 4.3). Fifty F_1 individuals, 576 F_2 individuals and 240 BC_1 individuals were assessed for downy mildew resistance. For the F_2 individuals, a total of 576 seeds from a single F_1 plant were sown in three batches of 192 seeds and were assessed *in vitro* for downy mildew resistance at 7 and 14 dpi following the methodology from section 4.2.1. Tissue sampling for every individual was performed 35 days after sowing. The segregation expected from a single dominant gene in an F_1 population is 1:0 resistant and susceptible individuals, for an F_2 population is 3:1 of resistant and susceptible and for BC_1 is 1:1 resistant and susceptible. All F_1 individuals were highly resistant to downy mildew showing the phenotype of the resistant parent 1028. From the 576 F_2 individuals assessed for downy mildew resistance 21 were eliminated because of differences in the resistance observed between the first and the second cotyledon. From the remaining 555 F_2 individuals, a total of 417 were resistant and 138 were susceptible to the pathogen (χ^2: 0.01), which matches almost perfectly the segregation for a single dominant gene. Finally, from the 240 BC_1 individuals, 124 were resistant and 116 were susceptible, which deviates slightly from the ratio expected from a single dominant gene (χ^2: 0.27). This lower number of susceptible individuals in the BC_1 and the higher number of resistant individuals could be explained by escapes where inoculum failed to infect in the Downy Mildew resistance assays. DNA was

extracted from all the individuals to perform KASP assays and selected individuals from all three batches were bulked into 6 different bulks, 3 resistant and 3 susceptible.

Figure 4.3. Timelapse of downy mildew infection over 15 days. Genotype on the left is the resistant parental line 1028 and on the right is the susceptible parental line Dwarf Blue (Right-click → open in a browser to visualize).

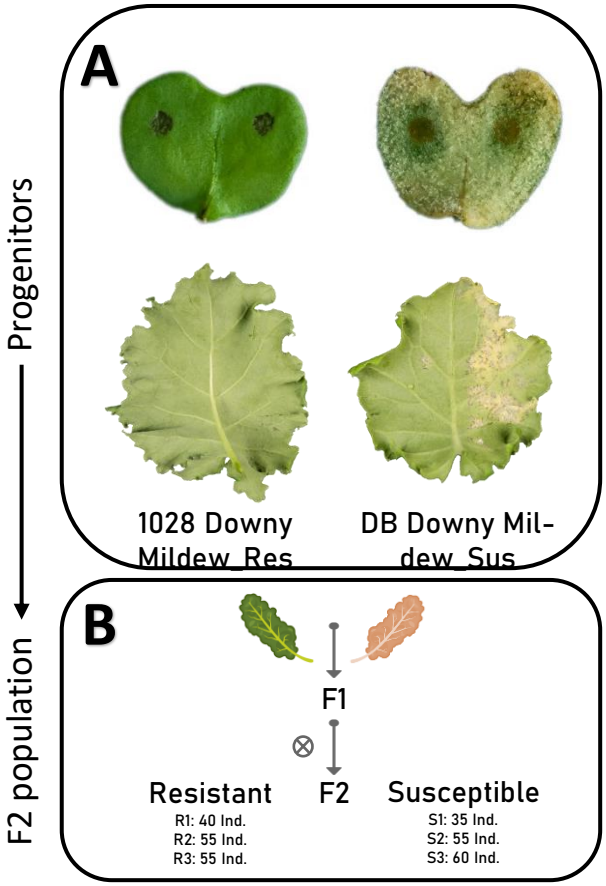


Figure 4.4. (a) Phenotype of the Downy Mildew resistant parental line 1028 (left) and the Downy Mildew susceptible parental line Dwarf Blue (right); 14 days post inoculation at seedling and adult plant stage. (b) F2 population developed by crossing 1028 and Dwarf Blue, followed by self inoculation of one F1 plant. F2 individuals were phenotyped for Downy Mildew resistance at 7 and 14 dpi and bulked based on their resistant or susceptible phenotype.The ⊗ represent self-pollination.

4.3.1.2 Adult plant downy mildew resistance assessment

Both parental lines were assessed for Downy Mildew resistance by spraying Downy Mildew inoculum (9 isolates) on both sides of the leaves at the 6-leaf stage. Every leaf was assessed individually at 7 and 14 dpi by visual screening and categorized into susceptible or resistant. Susceptible parental line Dwarf Blue showed signs of infection and sporulation with Downy Mildew at 7 and 14 dpi (Figure 4.5). On the other side, resistant parental line 1028 showed no signs of infection to Downy Mildew at 7 or 14 dpi (Figure 4.6). At the seedling stage, genotype 1028 was highly resistant but necrotic flecks were visible on the points of inoculation. In comparison, at the adult plant stage, no signs of damage caused by Downy Mildew were visible for the 1028 genotype. It is worth mentioning that the damage and sporulation caused by Downy Mildew in the susceptible parental line DB on leaves was significantly less compared to the damage and sporulation observed on the cotyledons. After 14 dpi cotyledons were fully colonised by Downy Mildew and tissue started to collapse not long after, but on leaves, the disease was more localised, and tissue took 21-28 dpi to start yellowing and collapse. However, this could be caused by the different methods of inoculation and incubation and not increase resistance to downy mildew of this tissue.

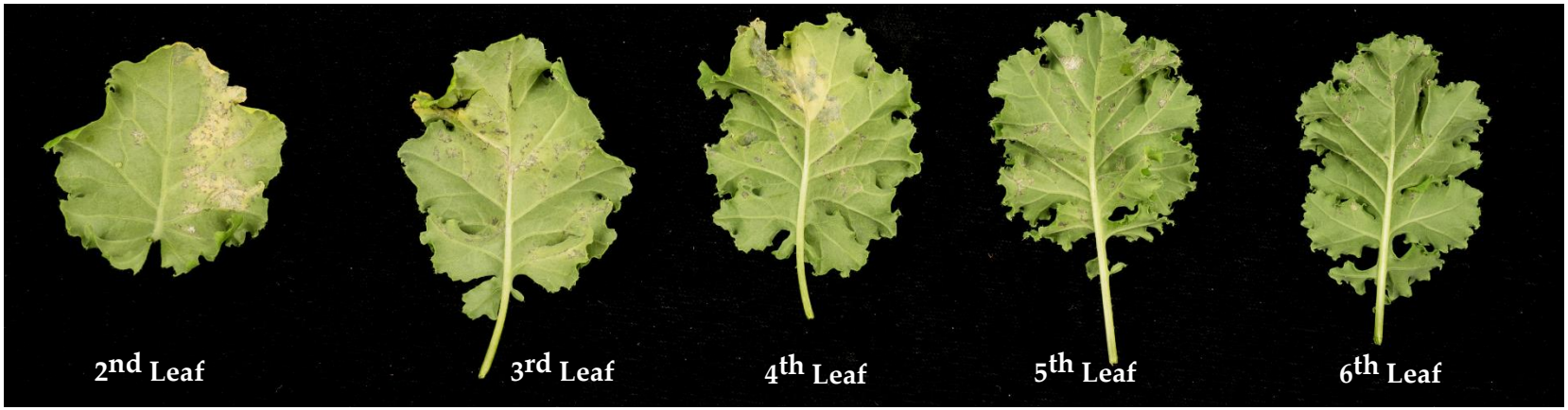


Figure 4.5. Downy Mildew susceptible parental genotype Dwarf Blue at 14 dpi. Leaves were sprayed with Downy Mildew inoculum on both sides.

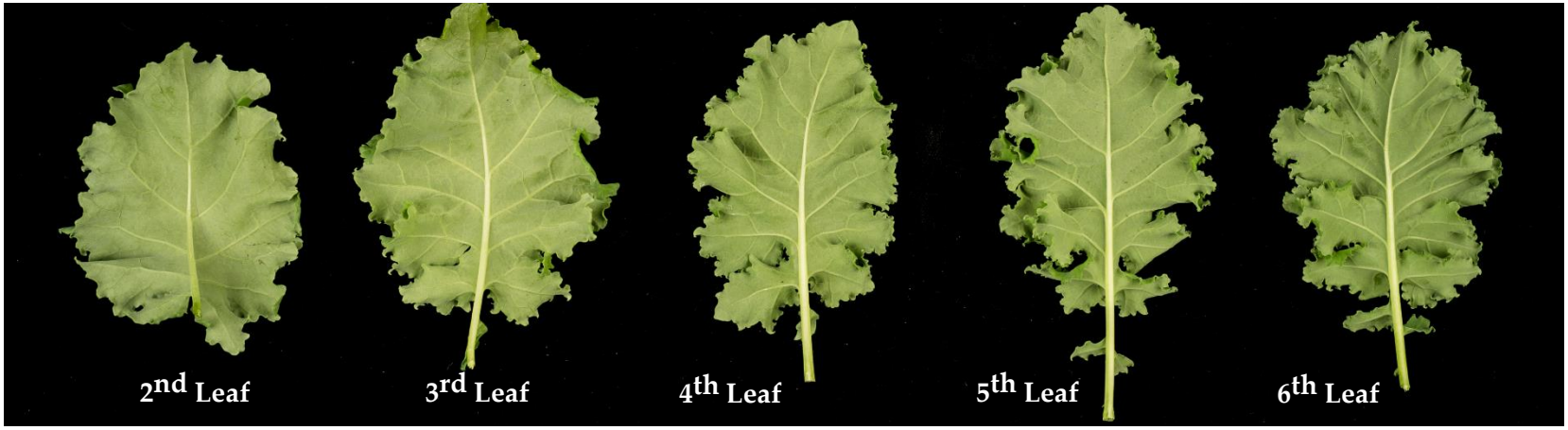


Figure 4.6. Downy Mildew resistant parental genotype 1028 at 14 dpi. Leaves were sprayed with Downy Mildew inoculum on both sides.

4.3.2 Sequencing and Mapping

High molecular weight DNA from all 6 bulks and both parental lines were successfully isolated using a custom CTAB method (Appendix A). All DNA samples were sent to Novogene for Illumina sequencing. The bulks aimed for 100x coverage and the parental line DB for 10x. FastQC was run with its default parameters for each of the FASTQ files and a summary of the sequencing data and QC analysis for all the samples can be seen in Appendix B. Raw read sequence reads and GC content % for every sample can be seen in Table 4.1. An average low GC (%) was observed for all the samples (37-38%) which is similar to the 36.76% reported by Sun et al. (2019) for the cauliflower draft genome.

Table 4.1. Raw reads, Raw data (G), Effective % and GC% for the 3 resistant and the 3 susceptible F_2 bulks and Dwarf blue parental genotypes were sequenced using Illumina reads.

Sample	Raw reads	Raw data(G)	Effective %	GC(%)
RES_BULK1	230,013,470	69.0	99.90	37.55
SUS_BULK1	204,344,821	61.3	99.88	37.08
RES_BULK2	372,161,110	55.8	99.24	36.73
SUS_BULK2	463,217,314	69.5	99.22	37.25
RES_BULK3	378,816,152	56.8	99.20	37.25
SUS_BULK3	442,824,430	66.4	99.19	37.25
DB_SUS	37,668,020	5.7	98.91	38.38

When the experiment was undertaken, several *B. oleracea* reference genomes were available: TO1000 (rapid-cycling Chinese kale; *B. oleracea* var. *alboglabra*) (Parkin et al., 2014), 02-12 (cabbage; *B. oleracea* var. *capitata*) (Liu et al., 2014), Pangenome (Golicz et al., 2016) and HDEM (broccoli type; *Brassica oleracea* var. *italica*). From these four assemblies, the more genetically similar to my parental kale lines (*B. oleracea* var. *acephala*) is the pangenome which is an assembly based on TO1000 but it also integrated additional sequences

from other *B. oleracea* cultivars like cabbage, cauliflower, broccoli, brussels sprout, kohlrabi and the kale DH line ARS_18 (Arsis DH) (Golicz et al., 2016). Therefore, the pangenome was selected as the main assembly to perform all the initial analyses. Using BWA v0.7.12 (Li & Durbin, 2009) all the samples were aligned to both the pangenome and HDEM assemblies and the resulting alignments were sorted using Samtools (Li et al., 2009).

4.3.3 Variant calling and filtering

To call for variants Freebayes v1.3.1 (Erik Garrison, 2012) was used on all six bulk alignments and the resulting variants were stored in a single VCF file for every assembly used. Using bcftools v1.8 (Danecek et al., 2021) a hard filter was applied for depth and quality (DP<60, QUAL<1000). To remove non-informative variants for the analysis, identical variants on both the resistant and the susceptible bulks were excluded from all VCF files. After applying the hard filters to the vcf file, a total of 7,574,161 variants were found on the pangenome assembly. 7,298,803 of them on all 9 *B. oleraceca* chromosomes, and the other 275,358 spread out along the scaffolds not anchored to chromosomes on the assembly (Table 4.2.). Types of variants and the total amount on the assembly can be seen on

Table 4.3.

Table 4.2. Total chromosomal variant number and rate for the *B. oleracea* Pangenome assembly.

Chromosome	Length	Variants	Variants rate
C1	46,852,435	1,419,503	33
C2	56,582,104	663,327	85
C3	70,374,523	603,720	116
C4	57,422,787	1,481,487	38
C5	49,931,166	1,085,026	46
C6	42,849,748	353,046	121
C7	51,750,011	955,163	54
C8	44,999,818	295,431	152
C9	58,447,134	442,100	132
Total	**479,209,726**	**7,298,803**	**86**

Table 4.3. The total number of variants by type for all chromosomes and scaffolds on the *B. oleracea* pangenome assembly.

Type	Total
SNP	2,902,792
MNP	1,181,026
INS	207,848
DEL	216,525
MIXED	3,065,970
Total	**7,574,161**

4.3.4 Bulk Frequency Ratios

After calling and filtering the variants present on the resistant and susceptible DNA bulks the objective was to find which variants were enriched for the resistant parental allele (1028 genotype) compared to the susceptible (Dwarf Blue genotype). Based on the segregation observed for the Downy Mildew resistance on the F_2 and BC_1 populations (single dominant gene), the region linked to the phenotype is expected to be

homozygote/heterozygote for the resistant allele (1028) on the resistant bulks and entirely homozygote for the susceptible allele (DB) on the susceptible pools. Therefore, to score this enrichment, the BFR, which is the ratio between the frequency of each allele (alternative; reference) for every variant was calculated between the resistant and the susceptible bulks (R. H. Ramirez-Gonzalez et al., 2015; Trick et al., 2012). Thus, a higher BFR value is indicative of an allele that is present in high numbers in the resistant bulk but has very low numbers or it is absent in the susceptible bulk.

In the pangenome assembly, the reference allele was expected to be linked with the susceptible phenotype and the alternative allele to the resistant phenotype. The pangenome assembly was mainly based on the previous TO1000 reference genome, and this genotype was very susceptible in the Downy Mildew resistance assessments performed in Chapter 3. Therefore, the BFRs for the alternative allele (AO) were selected and plotted against their chromosomal position in the *B. oleracea* pangenome.

A very narrow peak representing this enrichment was observed in the *B. oleracea* chromosome 2 on the pangenome assembly (Figure 4.7). To verify the position of the peak the HDEM assembly (broccoli type) was used with the same methodology. In the HDEM broccoli-type assembly, the previously observed peak on C2 on pangenome assembly was also observed in chromosome 2 and a very similar region on the chromosome (Figure 4.8). As can be seen on the close-up peak region in Figure 4.10, the peak encompasses a region of approximately 100.000 bp, from 53.275 kb to 53.385 kb on chromosome 2 of the pangenome assembly. While analysing all the variants (High-Low BFR) on the peak region, some of these variants perfectly matched the expected segregation for a single dominant gene (homozygote for the reference allele on the susceptible bulks and

heterozygote/homozygote on the resistant bulks). But these aforementioned ideal variants were unfortunately not included in the initial Manhattan plots because their BFR values tended to infinity due to having a value of zero for the allele count of one of the bulks, in this case, the susceptible bulk. Therefore, to include these ideal variants in the analysis, variants with a value of 0 for the alternative allele count on the susceptible bulks were changed from 0 to 1 on the pangenome and HDEM assembly. This change made the peak located on chromosome 2 more defined as can be seen in Figure 4.10. So even though in this scenario where there is a lack of coverage or complete homozygosity for a variant it is impossible to use the data, it can be resolved by changing the values from 0 to 1. Nevertheless, this is not ideal and is a limitation of the Bulk Segregant Analysis pipeline used in this chapter. Based on the number of homozygote variants found on the susceptible bulks we can assume that no escapes from the Downy Mildew resistant essays were included at least in the susceptible bulks. A summary of all the variants with a BFR value above 300 can be seen in Table 4.4. SNPs from this table were used to create KASP markers for Marker Assisted Selection (MAS) for the breeder on the pangenome assembly.

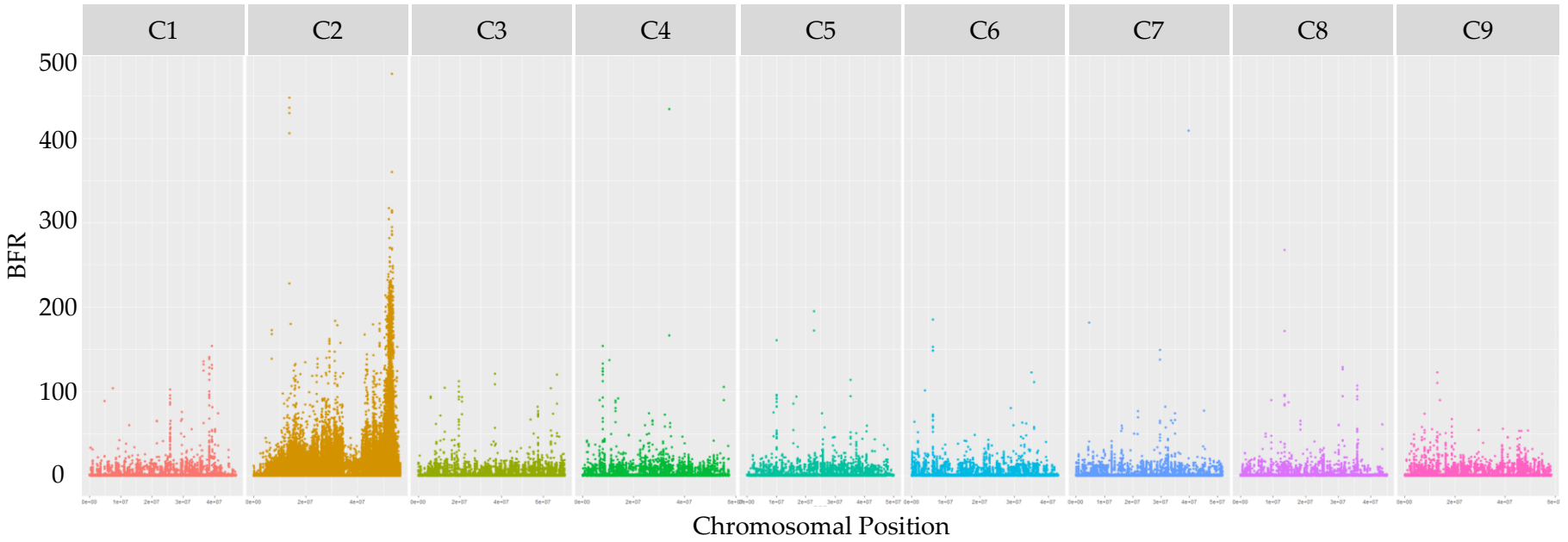


Figure 4.7. Variants Bulk Frequency Ratio (BFR) values plotted against their position on each chromosome on *B. oleracea* pangenome assembly. Higher BFR values are indicative of an allele that is highly associated with Downy Mildew resistance. C2 peak pinpoints the region with the variants that are highly associated with the resistance to Downy Mildew observed in the 1028 parental genotype.

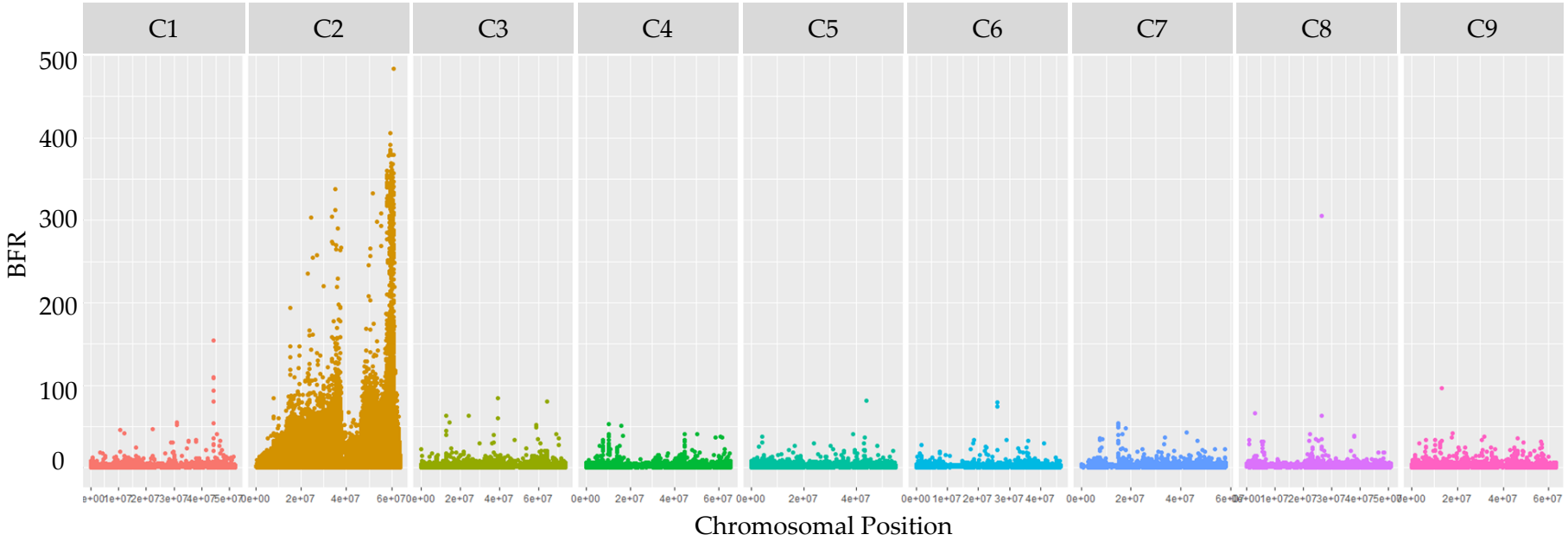


Figure 4.8. Variants Bulk Frequency Ratios (BFR) plotted against their position on each chromosome on *B. oleracea* HDEM assembly. Higher BFR values are indicative of an allele that is highly associated with Downy Mildew resistance. C2 peak is also present in the HDEM (broccoli) assembly and in the same region pinpointing the region assiciated to the resistance Downy Mildew observed the 1028 parental genotype.

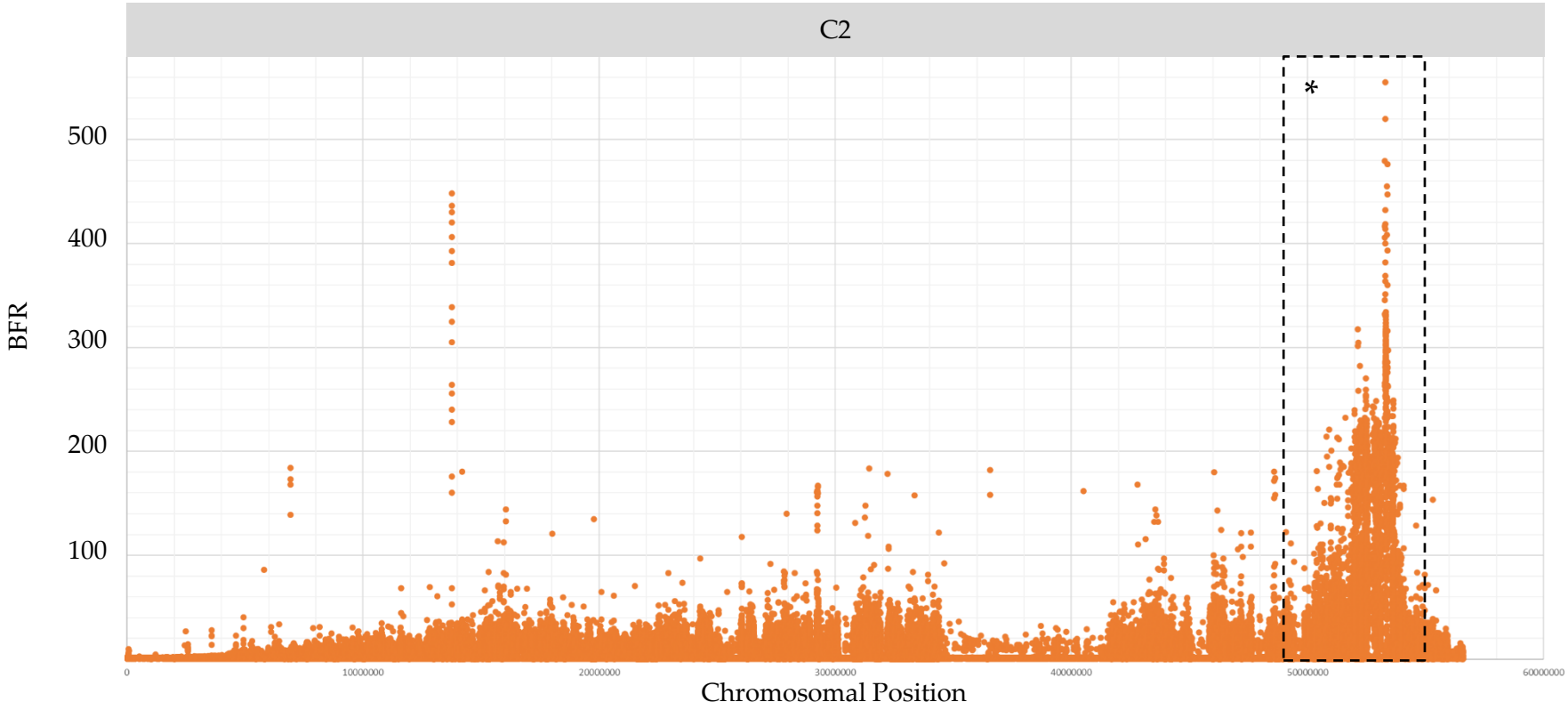


Figure 4.9. Chromosome 2 variants Bulk Frequency Ratios (BFR) plotted against their position on *B. oleracea* pangenome assembly. Higher BFR values are indicative of an allele that is highly associated with Downy Mildew resistance. * Location of the peak on 49,000,000-55,000,000 C2 position (Pangenome).

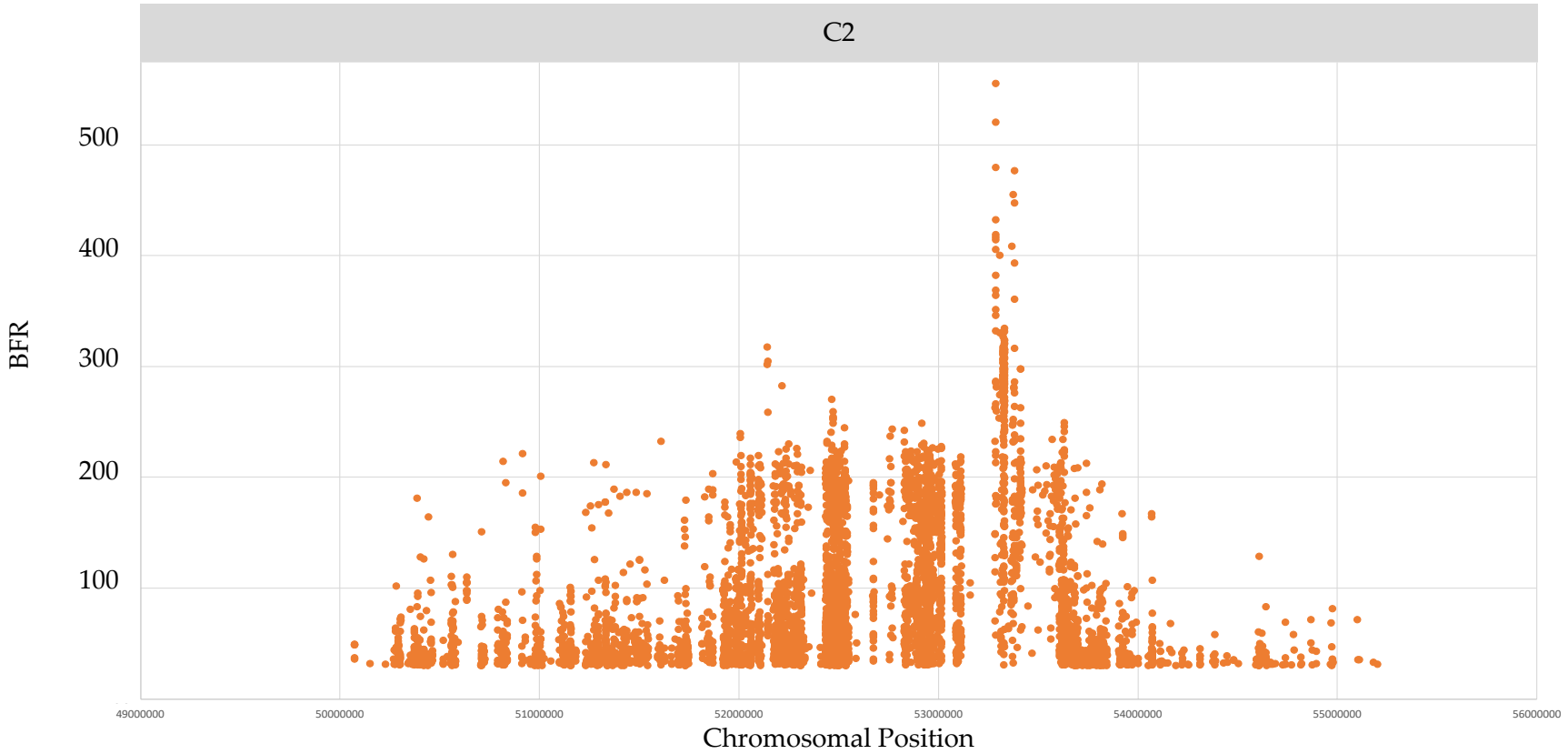


Figure 4.10. Zoom on the Chromosome 2 peak (* Figure 4.9) variants Bulk Frequency Ratios (BFR) plotted against their position on *B. oleracea* pangenome assembly (49,000,000-55,000,000).

Table 4.4. Variants with a BFR_AO (alternative allele count) value above 300 mapped to the *B. oleracea* pangenome assembly (Highly associated with Downy Mildew resistance).

CHROM	POS	REF	ALT	TYPE	QUAL	DP	BFR_AO
C2	53287602	GT	AA	mnp	13153	2217	555.2
C2	53287593	C	T	snp	11875	2132	520.1
C2	53285674	GTAAAT	GTAACT	snp	5668	2339	479.4
C2	53381916	C	A	snp	12156	2804	476.6
C2	53374760	T	G	snp	34645	2286	454.9
C2	53381937	A	G	snp	10646	2685	447.4
C2	53288943	C	T	snp	4076	2271	432.2
C2	53288559	G	T	snp	3017	2504	418.6
C2	53286315	A	T	snp	4423	1758	416.6
C2	53288569	A	G	snp	2951	2501	413.9
C2	53368396	T	C	snp	12099	1290	408.1
C2	53286304	G	A	snp	4454	1757	405.5
C2	53309024	A	G	snp	22312	1787	400.1
C2	53379955	TTGA	ATGA	snp	12049	3011	393.1
C2	53286571	C	G	snp	6566	1392	381.9
C2	53288618	A	G	snp	2130	2390	368.9
C2	53288437	G	C	snp	2493	2533	363.8
C2	53381834	C	A	snp	4930	2387	360.3
C2	53288902	T	A	snp	2277	2173	351.3
C2	53286202	G	C	snp	4614	1616	345.8
C2	53330178	C	T	snp	22151	2253	334.4
C2	53286147	A	G	snp	4840	1468	331.8
C2	53329540	ATTTTCTTGT	CTTTTCTTGT	snp	20850	2343	331.3
C2	53309004	C	A	snp	21569	1651	330.3
C2	53325024	A	C	snp	21402	2258	330.2
C2	53325767	T	G	snp	23403	2411	326.9
C2	53330384	G	T	snp	23095	2432	323.6
C2	53330199	AAA	GAC	complex	21880	2331	323.1
C2	53330415	G	A	snp	22718	2323	320.8
C2	53330455	A	C	snp	22777	2397	320.3
C2	53329179	A	T	snp	21755	2521	318.1
C2	53329188	CTTACC	CTTGCT	complex	22220	2443	317.6
C2	52143173	G	A	snp	6349	1006	317.4
C2	53325674	A	C	snp	23187	2419	317.1
C2	53381965	C	T	snp	6776	2215	316.1
C2	53325794	G	C	snp	23283	2377	315.5
C2	53330449	C	T	snp	22506	2362	315.1
C2	53330311	A	G	snp	22015	2399	315.0
C2	53326185	AAC	GAG	complex	22166	2375	314.8
C2	53329417	T	C	snp	21287	2251	314.7

C2	53325860	G	A	snp	22600	2420	313.6
C2	53330421	G	T	snp	21906	2302	313.2
C2	53324956	T	C	snp	21030	2159	313.1
C2	53331122	T	C	snp	21411	2330	312.5
C2	53331087	T	A	snp	21119	2328	312.3
C2	53325688	ATGT	GTGC	complex	23689	2375	312.2
C2	53325276	GTT	ATC	complex	20913	2409	311.0
C2	53329399	G	A	snp	22416	2215	311.0
C2	53325865	G	A	snp	21876	2408	310.1
C2	53329657	G	A	snp	21210	2338	307.8
C2	53329778	T	G	snp	22755	2306	307.0
C2	53325744	ACAAAC	TAAAAT	complex	22123	2282	306.8
C2	53324985	CATG	TATA	complex	21778	2180	306.2
C2	53324604	A	C	snp	20676	2235	304.8
C2	52144059	CACGA	TACGC	complex	6848	987	304.4
C2	53325057	T	G	snp	20873	2203	303.5
C2	53329394	T	C	snp	22037	2245	303.1
C2	53325217	A	G	snp	20366	2346	302.6
C2	53329832	G	C	snp	22043	2250	302.4
C2	53329939	T	C	snp	21538	2228	302.0
C2	52143062	C	T	snp	6138	1010	301.5
C2	53326062	G	A	snp	21258	2251	300.5

4.3.5 Pangenome peak region and candidate genes analysis

The peak encompasses a 100.000 bp region on chromosome 2 of the *B. oleracea* pangenome assembly (Figure 4.10). Based on the annotation, 12 genes are included in the region and all these genes are catalogued as un-characterized proteins. Therefore, I blasted all 12 annotated genes against the *A. thaliana* genome and annotated their corresponding homologue and function (Table 4.5).

Table 4.5. *A. thaliana* homologues and function description of all candidate genes associated with Downy Mildew resistance in the chromosome 2 peak region (*B. oleracea* pangenome).

CDS Name	*A. thaliana* ID	*A. thaliana* description
Bo2g159380	AT5G24450.1	Transcription factor IIIC, subunit 5
Bo2g159390	AT5G24340.1	3'-5' exonuclease domain-containing protein
Bo2g159400	AT1G59620.2	Disease resistance protein (CC-NBS-LRR class)
Bo2g159410	AT1G58807.2	Disease resistance protein (CC-NBS-LRR class)
Bo2g159420	AT1G59620.2	Disease resistance protein (CC-NBS-LRR class)
Bo2g159430	AT5G24450.2	Transcription factor IIIC, subunit 5
Bo2g159440	AT1G59620.2	Disease resistance protein (CC-NBS-LRR class)
Bo2g159450	AT1G58400.1	Disease resistance protein (CC-NBS-LRR class)
Bo2g159460	AT1G58807.2	Disease resistance protein (CC-NBS-LRR class)
Bo2g159470	AT4G32180.2	Pantothenate kinase activity protein
Bo2g159480	AT5G24350.2	Haloacid dehalogenase-like hydrolase (HAD)
Bo2g159490	AT1G58807.2	Disease resistance protein (CC-NBS-LRR class)

Seven of these genes have high homology with three different *A. thaliana* disease-resistant related homologues. Genes Bo2g159400, Bo2g159420 and Bo2g159440 have high homology to the *CW9* (At1g59620) gene, coding for a CC-NBS-NLR and a probable disease-resistance gene. Genes Bo2g159410, Bo2g159460, and Bo2g159490) have high homology with AT1G58807.2 (Bo2g159450) and gene Bo2g159450 with AT1G58400.1, which both encode for a probable disease resistance-related protein in *A. thaliana*. All three *A.*

thaliana homologous genes are in a similar region and have very high homology with each other. Moreover, all three genes have high homology with RPP7 (AT1g58602) or recognition of Peronospora parasitica 7 (*A. thaliana* downy mildew). This data suggests that these 7 genes in chromosome 2 of *B. oleracea* could be part of an NLR cluster related to Downy Mildew resistance. To confirm if these 7 genes or some of them are functional NLRs I ran the tool NLR annotator (Steuernagel et al., 2018; Steuernagel et al., 2020) on the assemblies to predict possible loci associated with NLRs. Out of the 7 genes, 5 of them are predicted to be completely functional NLRs and one of them (C2_nlr_25; PANG_NLR4) is not included in the annotation (Figure 4.12). NLR clusters are known to have drastic structural variations between different genotypes, therefore I compared the NLR cluster on the pangenome assembly to investigate if all 5 predicted NLRs were also present in the HDEM assembly. Analysing the two NLR clusters, only 3 NLRs were present in the HDEM assembly compared with the pangenome, and in this case two of them were predicted to be non-functional NLRs so only one is predicted to be complete and functional if expressed (Figure 4.12). In addition to this, some motifs on the predicted NLR genes were highly enriched compared to other motifs from the same gene (more reads aligned to these motifs). This could be due to the high sequence similarity between NLR genes in the cluster and reads could be aligned to more than one place in the genome. This was the case for some motifs but for others, all the reads were unique to that specific motif. This observation suggests that more NLRs could be present in the 1028 parental genotype but the reads from these genes can't be mapped to the assemblies because they don't exist in the genotypes that were used to create the assemblies or they couldn't be properly resolved and assembled due to the use of short reads to create it.

Therefore, due to the high sequence similarity observed between the NLRs in the cluster, some of the reads from other NLRs not present in the assemblies could be mapped to similar regions in the NLRs present in the assemblies, enriching those sites. Moreover, analysing the NLR cluster and neighbouring regions using a genome browser, big deletions were detected on both assemblies for the reads coming from the 3 susceptible bulks and the susceptible DB parental genotype. These deletions were mostly on non-coding regions around the NLRs in the cluster but also on the coding regions of 2 of these NLRs. No deletions were observed on the resistant bulks. Therefore, to fully resolve the NLR cluster region present in the downy mildew resistant genotype 1028, a long-read sequencing assembly was performed.

It is worth mentioning the presence of the gene Bo2g159320 in the neighbouring regions of the NLR cluster. This gene is a homologue of the AT5G24530.1 *A. thaliana* gene which encodes for a 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein also known as Downy MildewR6. Downy MildewR6 was identified by Van Damme et al. (2005) as a susceptibility factor in *A. thaliana*. The effects of the inactivation of Downy MildewR6 resulted in increased salicylic acid levels, conferring resistance to different plant pathogens (Zeilmaker et al., 2015). Analysing the variants identified in the coding and nearby regions of Bo2g159320 (Downy MildewR6 homologue) between the susceptible and resistant bulks and susceptible parental genotype Dwarf Blue, no variants were found to have an effect for the resistance observed on the 1028 genotype. Analysing these gene on the pangenome and HDEM assembly some variants were identified but all of them were identical in the resistant and susceptible bulks.

4.3.6 1028 Draft genome assembly using PacBio HiFi long reads

To fully resolve the NLR cluster identified using the bulk segregant analysis pipeline previously described a high molecular weight DNA sample of the Downy Mildew resistant 1028 genotype was sent to the Earlham Institute (Norwich, UK) for long-read whole-genome sequencing using Pacbio HiFi technology. This sequencing generated 9.5Gb of Hifi data (>=Q20 quality) with about 15.3x coverage assuming a 620Mb *B. oleracea* genome size. The average insert size is 22Kb with reads up to 30-35Kb (Appendix C3).

To create the assembly I used the tool hifiasm v0.16.1 to *de novo* assemble the PacBio HiFi reads. The genome assembly resulted in a 685Mb *B. oleracea* var *acephala* (kale) genome size with 1542 contigs, with a contig N50 length of 2.97 Mb and a 22.8Mb biggest contig. The genome size of this draft assembly (685Mb) is a bit bigger compared with other *B. oleracea* assemblies like D134 (cabbage) with 530Mb (Lv et al., 2020), 02-12 (cabbage) with 514Mb (Liu et al., 2014), TO1000 (kale-like) with 489Mb (Parkin et al., 2014), pangenome with 587Mb (Golicz et al., 2016), and the HDEM (broccoli) with 555Mb (Belser et al., 2018).

4.3.6.1 NLR prediction and gene models identification

The 1028 draft assembly was analysed using NLR annotator (Steuernagel et al., 2018; Steuernagel et al., 2020) to predict all possible loci associated with NLRs. The gene models were identified using RNA-Seq data derived from two samples coming from cotyledon and leaf tissue from six different plants each. The pipeline used to identify gene models in the 1028 draft genome was kindly provided by Matthew Moscou (TSL). Both samples were aligned to the draft genome using HISAT2 (Kim et al., 2015) and gene models were

identified and merged using Cufflinks (Trapnell et al., 2010). Finally, TransDecoder was used to identify candidate coding regions using the merged RNA-Seq alignments to the 1028 Draft genome generated by Cufflinks.

4.3.7 NLR cluster analysis

The *de novo* 1028 draft assembly was processed using the same methodology previously used with the pangenome and HDEM assemblies to calculate and plot the BFR using the 3 Downy Mildew resistant bulks and the 3 Downy Mildew susceptible bulks. In contrast to the pangenome assembly, the expected allele carrying the resistance to Downy Mildew in the 1028 draft assembly is the one from the reference. Therefore, the BFRs for the reference allele (RO) were selected and plotted against their chromosomal position (Figure 4.11). After analysing the data, the peak and NLR cluster observed on chromosome 2 on the pangenome and HDEM assemblies was localised to a single 888kb contig on the 1028 draft genome (utg000149l).

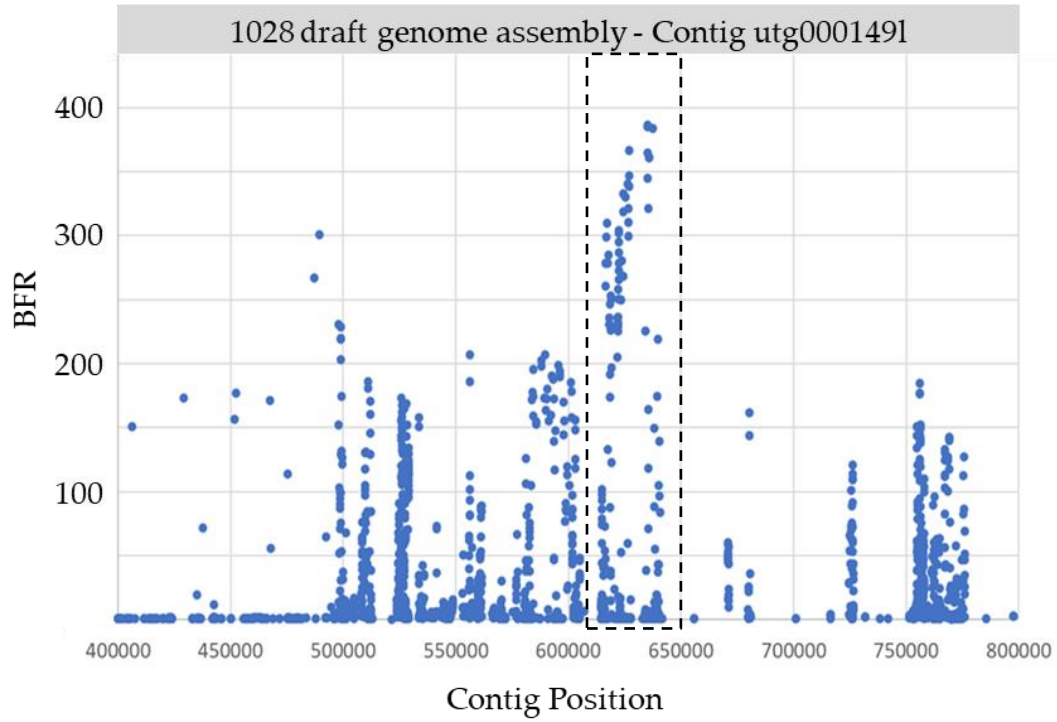


Figure 4.11. Variants Bulk Frequency Ratio (BFR) values plotted against their position on the utg000149l contig from the 1028 draft genome assembly. Higher BFR values are indicative of an allele that is highly associated with Downy Mildew resistance. It represents a section (400kb-800kb) of the contig utg000149l (888kb).

The NLR cluster in the 1028 draft genome contig contained a total of 12 predicted NLRs from which 9 are predicted to be complete by NLR annotator (Figure 4.12) in a 350kb region. From these 3 partial NLRs, only one is expressed in the 1028 genotype the other 2 are in the 1028 genotype genome but they are not expressed. In comparison, as previously mentioned, the NLR cluster in the pangenome assembly had 5 predicted complete NLRs and 3 NLRs in the HDEM assembly, out of which just one was predicted to be complete (Figure 4.12). Even though gene clustering is uncommon in eukaryotes (Lee & Sonnhammer, 2003), it is more frequent for NLR genes (van Wersch & Li, 2019). In *A. thaliana* NLR clusters range from small NLR clusters, containing only two NLR genes (Narusaka et al., 2009), to very large

with 11 NLR genes (Holub, 2001). Assuming that the observed structural variations for this NLR cluster are real and not due to assembly issues due to NLR sequences similarities that prevented a proper assessment of the CNV and diversification of duplicated sequences. This NLR gene cluster associated with Downy Mildew resistance shows a significant structural diversification among all three *B. oleracea* assemblies, with gene copy number variation ranging from 3 to 12 NLRS. From the three assemblies tested the only one that could have had technical limitations to assess the NLR present in the cluster is the pangenome due to the use of short reads to create it. For the other two assemblies, long reads sequences were used, HDEM (Nanopore PromethION) and 1028 draft genome (Pacbio HiFi), which increases the probability of assessing regions with high gene structural variation, like NLR clusters properly. To assess how similar these NLRs sequences are to each other, I ran a progressive pairwise alignment using default settings on Geneious prime for all the NLRs predicted by NLR annotator on all three assemblies (pangenome; HDEM; 1028 Draft genome). Data is presented as identity percentage in

	1028_NLR1_partial	**1028_NLR2**	**1028_NLR3**	**1028_NLR4**	**1028_NLR5**	**1028_NLR6_partial**	**1028_NLR7_partial**	**1028_NLR8**	**1028_NLR9**	**1028_NLR10**	**1028_NLR11**	**1028_NLR12**	**HDEM_NLR1_partial**	**HDEM_NLR2**	**HDEM_NLR3_partial**	**PANG_NLR1**	**PANG_NLR2**	**PANG_NLR3**	**PANG_NLR4**	**PANG_NLR5**
1028_NLR1_partial		81	72	73	86	0	**96**	0	86	80	85	79	83	86	84	**90**	81	83	83	84
1028_NLR2	81		66	72	86	0	79	0	79	77	**93**	75	70	86	73	74	81	74	80	81
1028_NLR3	72	66		70	68	0	81	0	82	72	67	72	74	70	74	70	70	**98**	79	78
1028_NLR4	73	72	70		76	0	84	0	79	72	75	72	71	75	76	72	67	77	**99**	75
1028_NLR5	86	86	68	76		0	85	0	84	82	88	79	72	**97**	78	82	78	80	84	85
1028_NLR6_partial	0	0	0	0	0		0	**100**	0	0	0	0	0	0	0	0	0	25	0	0
1028_NLR7_partial	**96**	79	81	84	85	0		0	86	87	83	86	83	85	89	84	79	82	83	85

1028_NLR8	0	0	0	0	0	**100**	0		0	0	0	0	0	0	0	0	0	25	0	0
1028_NLR9	86	79	82	79	84	0	86	0		**91**	84	**91**	72	85	82	82	71	81	81	**92**
1028_NLR10	80	77	72	72	82	0	87	0	**91**		80	88	75	81	83	76	84	85	84	**90**
1028_NLR11	85	**93**	67	75	88	0	83	0	84	80		78	70	**90**	76	81	82	80	84	84
1028_NLR12	79	75	72	72	79	0	86	0	**91**	88	78		73	79	79	77	75	82	81	**98**
HDEM_NLR1_partial	83	70	74	71	72	0	83	0	72	75	70	73		72	78	72	70	74	71	73
HDEM_NLR2	86	86	70	75	**97**	0	85	0	85	81	**90**	79	72		79	81	75	80	84	85
HDEM_NLR3_partial	84	73	74	76	78	0	89	0	82	83	76	79	78	79		75	74	78	78	79
PANG_NLR1	**90**	74	70	72	82	0	84	0	82	76	81	77	72	81	75		67	73	81	83
PANG_NLR2	81	81	70	67	78	0	79	0	71	84	82	75	70	75	74	67		61	74	79
PANG_NLR3	83	74	**98**	77	80	25	82	25	81	85	80	82	74	80	78	73	61		80	84
PANG_NLR4	83	80	79	**99**	84	0	83	0	81	84	84	81	71	84	78	81	74	80		81
PANG_NLR5	84	81	78	75	85	0	85	0	**92**	**90**	84	**98**	73	85	79	83	79	84	81	

Table 4.6 and NLRs are named with the assembly to which they correspond (PANG; HDEM; 1028) + a number in the order of appearance based on their position in the NLR cluster; partial NLRs are also included in the comparison. The pairwise percentage identity for all 20 NLRs is very high (79.6%). Every NLR sequence identified in the assemblies is identical or almost identical to 1 or 2 NLRs on the same or the other assemblies. Starting with all 12 NLRs identified in the 1028 draft genome assembly with a % identity above 90: 1028_NLR1_partial sequence is 96% similar to 1028_NLR7 and 90% to the pangenome nlr1 (PANG_NLR1). 1028_NLR2 sequence is 93% similar 1028_NLR11. 1028_NLR3 sequence is 98% similar to PANG_NLR3. 1028_NLR4 sequence is 99% similar to PANG_NLR4. 1028_NLR5 sequence is 97% similar to HDEM_NLR2. 1028_NLR6_partial sequence is 100% similar to 1028_NLR8. 1028_NLR7_partial sequence is 98% similar to PANG_NLR3. 1028_NLR8 sequence is 100% similar to 1028_NLR6. 1028_NLR9 sequence is 91% similar to 1028_NLR10 and 1028_NLR10, and 92 to PANG_NLR5. 1028_NLR10 sequence is 90% similar to PANG_NLR5 and 91% to 1028_NLR9. 1028_NLR11 sequence is 90% similar to HDEM_NLR2 and 93% to 1028_NLR2. 1028_NLR12 sequence is 91% similar to 1028_NLR9 and 98% to PANG_NLR5.

For the 3 NLRs in the HDEM assembly with a %identity above 90: For

HDEM_NLR1_partial sequence, no other NLR sequence has above 90% %identity but 1028_NLR1_partial and 1028_NLR7_partial have 83% identity. HDEM_NLR2 sequence is 90% similar to 1028_NLR11 and 97% to 1028_NLR5. For HDEM_NLR3_partial no other sequence had above 90% identity but as well as HDEM_NLR1 has 84% and 89% identity to 1028_NLR1_partial and 1028_NLR7_partial respectively.

For the 5 complete NLRs in the pangenome assembly with a %identity above 90: PANG_NLR1 sequence is 90% similar to 1028_NLR1_partial. For PANG_NLR2 no other sequence had above 90% identity but has 84% identity to 1028_NLR10. PANG_NLR3 sequence is 98% similar to 1028_NLR3. PANG_NLR4 sequence is 99% similar to 1028_NLR4. And PANG_NLR5 sequence has a 90%, 92% and 98% identity to 1028_NLR10, 1028_NLR9 and 1028_NLR12 respectively.

This data suggests that the identified NLR cluster could have had more than one event of contraction and/or expansion creating the different CNV NLR numbers observed in all 3 assemblies. The resolved NLR cluster using the *de novo* assembly of the 1028 genotype complicates the search for the gene responsible for the high Downy Mildew resistance observed at the seedling and adult plant stage increasing the total number of possible NLR genes from 1-5 in the HDEM and pangenome assembly to 9 in the 1028 draft assembly. The initial idea of using RNA-seq data was to possibly eliminate some of the candidate genes if some of the predicted NLR genes were not expressed but based on the annotation created with the RNA-Seq reads, 8 out of the 9 complete NLRs are expressed. In addition to this, even though the F_2 population matches perfectly the segregation for a single dominant gene, 2 or more genes could be functioning together (van Wersch & Li, 2019) and don't segregate as 2 separate genes because of how physically close

they are in the 1028 genome.

Table 4.6. Pairwise comparison for all predicted NLR (complete and partial) sequences from the Pangenome, HDEM and 1028 draft genome assemblies. All values are represented as percentage identity.

	1028_NLR1_partial	1028_NLR2	1028_NLR3	1028_NLR4	1028_NLR5	1028_NLR6_partial	1028_NLR7_partial	1028_NLR8	1028_NLR9	1028_NLR10	1028_NLR11	1028_NLR12	HDEM_NLR1_partial	HDEM_NLR2	HDEM_NLR3_partial	PANG_NLR1	PANG_NLR2	PANG_NLR3	PANG_NLR4	PANG_NLR5
1028_NLR1_partial		81	72	73	86	0	**96**	0	86	80	85	79	83	86	84	**90**	81	83	83	84
1028_NLR2	81		66	72	86	0	79	0	79	77	**93**	75	70	86	73	74	81	74	80	81
1028_NLR3	72	66		70	68	0	81	0	82	72	67	72	74	70	74	70	70	**98**	79	78
1028_NLR4	73	72	70		76	0	84	0	79	72	75	72	71	75	76	72	67	77	**99**	75
1028_NLR5	86	86	68	76		0	85	0	84	82	88	79	72	**97**	78	82	78	80	84	85
1028_NLR6_partial	0	0	0	0	0		0	**100**	0	0	0	0	0	0	0	0	0	25	0	0
1028_NLR7_partial	**96**	79	81	84	85	0		0	86	87	83	86	83	85	89	84	79	82	83	85
1028_NLR8	0	0	0	0	0	**100**	0		0	0	0	0	0	0	0	0	0	25	0	0
1028_NLR9	86	79	82	79	84	0	86	0		**91**	84	**91**	72	85	82	82	71	81	81	**92**
1028_NLR10	80	77	72	72	82	0	87	0	**91**		80	88	75	81	83	76	84	85	84	**90**
1028_NLR11	85	**93**	67	75	88	0	83	0	84	80		78	70	**90**	76	81	82	80	84	84
1028_NLR12	79	75	72	72	79	0	86	0	**91**	88	78		73	79	79	77	75	82	81	**98**
HDEM_NLR1_partial	83	70	74	71	72	0	83	0	72	75	70	73		72	78	72	70	74	71	73
HDEM_NLR2	86	86	70	75	**97**	0	85	0	85	81	**90**	79	72		79	81	75	80	84	85
HDEM_NLR3_partial	84	73	74	76	78	0	89	0	82	83	76	79	78	79		75	74	78	78	79
PANG_NLR1	**90**	74	70	72	82	0	84	0	82	76	81	77	72	81	75		67	73	81	83
PANG_NLR2	81	81	70	67	78	0	79	0	71	84	82	75	70	75	74	67		61	74	79
PANG_NLR3	83	74	**98**	77	80	25	82	25	81	85	80	82	74	80	78	73	61		80	84
PANG_NLR4	83	80	79	**99**	84	0	83	0	81	84	84	81	71	84	78	81	74	80		81
PANG_NLR5	84	81	78	75	85	0	85	0	**92**	**90**	84	**98**	73	85	79	83	79	84	81	

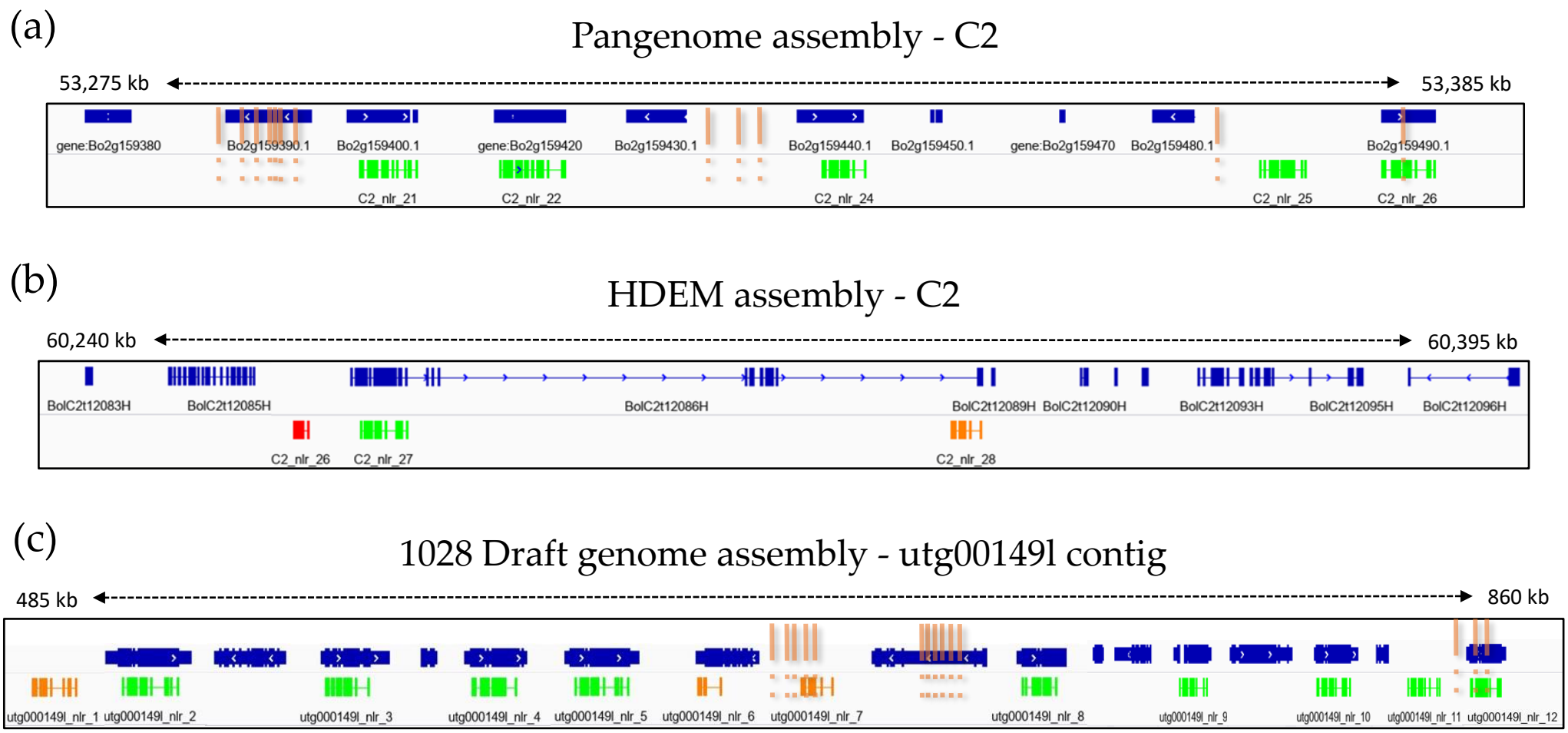


Figure 4.12. NLR cluster region on all 3 assemblies: (a) Pangenome (C2), (b) HDEM (C2) and (c) 1028 de novo draft genome (utg00149l). For all three assemblies: black dotted line represent the size of the NRL cluster and the initial and end position in the chromosome or contig; annotated genes are represented in blue; Predicted complete NLR are represented in light green and partial NLR represented in orange and red; KASP markers designed using high BFR SNPs for the Pangenome and the 1028 draft genome are represented as vertical orange lines with two dots. Space between genes in the regions have been compressed for space efficiency.

4.3.8 KASP marker analysis of the NLR cluster

4.3.8.1 SNPs validation

To validate the mapped Downy Mildew resistance loci (Section 5.3.5-132), KASP assays were designed using SNPs with high BFR from the BSA in both the Pangenome and the 1028 draft genome assembly. To automate the design and validation of specific primers for every assembly, the tool Poly-Marker was used (Ricardo H. Ramirez-Gonzalez et al., 2015). The main criteria to select the SNPs in the mapped Downy Mildew loci region for marker creation was the BFR, only selecting the ones above 300 for the pangenome assembly SNPs and above 200 for the 1028 draft genome assembly (Table 4.4. Variants with a BFR_AO (alternative allele count) value above 300 mapped to the *B. oleracea* pangenome assembly (Highly associated with Downy Mildew resistance). The selected SNPs should originate from the allele coming from the 1028 parental genotype, for the pangenome assembly is the AO (alternative allele) and for the 1028 assembly is RO (reference allele).

A total of 26 SNPs were selected using SNP data from the BSA and markers were designed for them. In total, 12 KASP markers were designed for the Pangenome assembly (Table 4.7), and 14 for the 1028 draft genome assembly (Table 4.8). As a control, 2 KASP markers were designed based on SNPs outside of the mapped Downy Mildew resistant loci for both the pangenome (KASP markers 1 and 14) and the 1028 genome assembly (KASP markers 16 and 17). To validate if the 30 SNPs for both assemblies were polymorphic and diagnostic of the Downy Mildew resistance loci, a subset of 96 F_2 individuals plus both parental lines and four Downy Mildew susceptible breeding lines from CN seeds were tested. For both assemblies, markers outside of the mapped Downy Mildew resistance loci (KASP

marker 1, 14, 16 and 17) were polymorphic but they were not diagnostic of the Downy Mildew resistance. For the 12 SNPs on the pangenome assembly, only 2 were polymorphic and diagnostic (KASP markers 3 and 12) of the Downy Mildew resistance loci (17%). Of the 14 SNPs on the 1028 assembly, 9 were polymorphic and diagnostic (KASP markers 18, 19, 20, 22, 23, 24, 26, 28 and 29) of the Downy Mildew resistance loci (64%).

All the polymorphic and diagnostic SNPs for both the pangenome and the 1028 assembly were further tested on the whole F_2 population consisting of 555 individuals with phenotypic information for the resistance to downy mildew. Of those 555 89 were removed from the analysis due to having more than 2 missing data points. Therefore a total of 466 F_2 individuals (119 susceptible; 347 resistant) were assessed with the 11 most closely linked KASP markers for the Downy Mildew resistance. After running the assay, all 11 KASP markers were highly associated with the resistance on both assemblies. Nevertheless, false positives were observed, for example, one completely resistant individual was scored as susceptible due to being homozygote for the susceptible allele for all 11 SNPs tested. Because of the consistency for all the KASP markers assessed this was probably an escape while phenotyping the individuals for Downy Mildew resistance or a wrongly labelled sample. On the other side of the spectrum, some lines phenotypically identified as susceptible were scored as resistant (heterozygote) for KASP markers 3, 22 and 28 with 1, 1 and 3 lines respectively. For the rest of the markers, all resistant individuals carried at least one copy of the resistant 1028 haplotype and all the susceptible were homozygotes for the susceptible dwarf blue haplotype (KASP markers 12, 18, 19, 20, 23, 24, 26, and 29). From these highly associated SNPs, it is worth annotating that KASP markers 20, 24, 26, and 29 have less than 2 missing data points out of

466 individuals and no false positives.

Table 4.7. Primer details and genome position in the pangenome assembly (Chromosome 2) for the markers to validate. Primers in dark blue are outside of the mapped downy mildew resistant locus and markers in bold are markers validated to be polymorphic and diagnostic of the Downy Mildew resistance.

KASP ID	Position	Contig Regions	A	B	Common	Orientation
1	16412166	C2:16412038-16412290	acaaagaaaggtgctctgacaT	acaaagaaaggtgctctgacaC	gatagaggggcaagctcact	forward
2	53285674	C2:53285565-53285797	aacatcctgccaagtaggtaaA	aacatcctgccaagtaggtaaC	cgtgcatcacctgaggacaa	forward
3	**53286147**	**C2:53286018-53286264**	**tccactcttttaccctcttgtgtaT**	**tccactcttttaccctcttgtgtaC**	**gatgcagactatgttgacacgg**	**reverse**
4	53287593	C2:53287476-53287697	aacttgcatacatccatgaacttC	aacttgcatacatccatgaacttT	tggttgcagggaactcagta	forward
5	53288437	C2:53288328-53288549	atgtgctcaagtgcaagacC	atgtgctcaagtgcaagacG	ccaaaaggggcatgatcct	reverse
6	53288559	C2:53288440-53288665	ggtcaattcctgggaagcaG	ggtcaattcctgggaagcaT	gcgcaaaatcattccttgcc	forward
7	53288569	C2:53288456-53288673	tgggaagcagtttgtcagcA	tgggaagcagtttgtcagcG	gcgcaaaatcattccttgcc	forward
8	53288943	C2:53288829-53289033	cgtaaaaacctcagaaacacaaaaC	cgtaaaaacctcagaaacacaaaaT	gggagtggtggacctaaaaca	forward
9	53325767	C2:53325650-53325875	acaaacacccatgtattgacattT	acaaacacccatgtattgacattG	tgcgttgttttgttatatgcga	forward
10	53329540	C2:53329427-53329664	cccggcccactaaacaaatattA	cccggcccactaaacaaatattC	acatagtgtcaaacaacttattgca	forward
11	53330449	C2:53330334-53330566	agtgtacaagttgttccttaggT	agtgtacaagttgttccttaggG	caggtttggacgtatgtgtgt	reverse
12	**53368396**	**C2:53368328-53368497**	**tcatcattggtgcaacatggT**	**tcatcattggtgcaacatggC**	**gcgcatggacttcattgctg**	**forward**
13	53379955	C2:53379848-53380072	atacaaaacaccgtaaaccttcaA	atacaaaacaccgtaaaccttcaT	tagagggagtgggtctgacg	reverse
14	54570284	C2:54570153-54570411	ttctacctggttctgatgaaacT	ttctacctggttctgatgaaacA	gtggtgtgatccctgcacta	forward

Table 4.8. Primer details and genome position in the 1028 draft genome assembly (utg00149l contig) for the markers to validate. Primers in dark blue are outside of the mapped downy mildew resistant locus and markers in bold are markers validated to be polymorphic and diagnostic of the Downy Mildew resistance.

KASP ID	Position	Contig Region	A	B	Common	Orientation
16	129517	utg000149l:129385-129646	atatatgcatcaactttgcttgttG	atatatgcatcaactttgcttgttT	agaatgaaatgggaaacgttgtagt	forward
17	304207	utg000149l:304075-304339	cacaacggatctgataaagctcT	cacaacggatctgataaagctcA	ggctaacagtggctcttcct	reverse
18	**617669**	**utg000149l:617538-617799**	**agattatagaggggaatttcactgC**	**agattatagaggggaatttcactgT**	**tgaaccctaccagctcatca**	**reverse**
19	**624162**	**utg000149l:624031-624295**	**ttcagtttgatgttaacagtgcaT**	**ttcagtttgatgttaacagtgcaC**	**actgtggataggcatctcca**	**reverse**
20	**624401**	**utg000149l:624274-624531**	**gtgcttccactagctccaT**	**gtgcttccactagctccaC**	**cttcggtggacgtcctctaa**	**forward**
21	624449	utg000149l:624326-624578	tggtatatgattatgctcgcttcG	tggtatatgattatgctcgcttcA	tccatgaaatcggcatcgcc	reverse
22	**626432**	**utg000149l:626320-626559**	**gctgaagagtttcatgtcactttcC**	**gctgaagagtttcatgtcactttcT**	**gctcagacatgacgagggaa**	**reverse**
23	**635067**	**utg000149l:634949-635187**	**atgtgctcaagtgcaagacG**	**atgtgctcaagtgcaagacC**	**ccaaaaggggcatgatcct**	**reverse**
24	**635189**	**utg000149l:635082-635303**	**ggtcaattcctgggaagcaT**	**ggtcaattcctgggaagcaG**	**cgcaaaaccattccttgcca**	**forward**
25	635199	utg000149l:635094-635315	tgggaagcattttgtcagcG	tgggaagcattttgtcagcA	cgcaaaaccattccttgcca	forward
26	**635248**	**utg000149l:635126-635377**	**tttcttggatggcaaggaatgG**	**tttcttggatggcaaggaatgA**	**ggctatcagttcatcttttggcg**	**forward**
27	635673	utg000149l:635555-635799	agttgatatctcggtagttccaagT	agttgatatctcggtagttccaagA	gggcgagcgttgtttttact	forward
28	**637412**	**utg000149l:637298-637536**	**ccatgttttatctcagggcttgtG**	**ccatgttttatctcagggcttgtA**	**cctgcaaatatcttctggtctgc**	**forward**
29	**854513**	**utg000149l:854386-854643**	**accagatgcttataaaacaactctG**	**accagatgcttataaaacaactctC**	**caatccccacacagcgtaga**	**reverse**
30	855123	utg000149l:855009-855249	tgaccagtcactacattgaagcA	tgaccagtcactacattgaagcT	cctgcattaatgtacaggttttgt	forward
31	856973	utg000149l:856843-857104	cgattctcgtttcggttccT	cgattctcgtttcggttccC	tagaacacgaggggcagaac	forward

4.4 Discussion

In this chapter, I identified and map a new source of resistance to *H. brassicae* on Chromosome 2 in an F_2 population derived from the cross between the downy mildew resistant genotype 1028 and the susceptible commercial variety Dwarf Blue. To date, four different loci conferring resistance to downy mildew have been mapped in *B. oleracea*, three of them with a known chromosomal location in the genome. BoDowny Mildew1 in chromosome 5 (Gao et al., 2003; Giovannelli et al., 2002), Pp523 in chromosome 8 (Carlier et al., 2012; Coelho & Monteiro, 2003b; Farinho et al., 2004) and Ppa207 in chromosome 2 (Saha et al., 2020). The loci associated with downy mildew resistance presented in this chapter were identified from a kale breeding variety 1028 (CN seeds). Despite the fact that Ppa207 and the QTL identified in this study were both located on C02, Ppa207 is located in the extreme of the longest arm (Shaw et al., 2021) and the resistance from this study in the extreme of the short arm, so their physical placements indicate that they are two separate QTLs (

Figure 4.13).

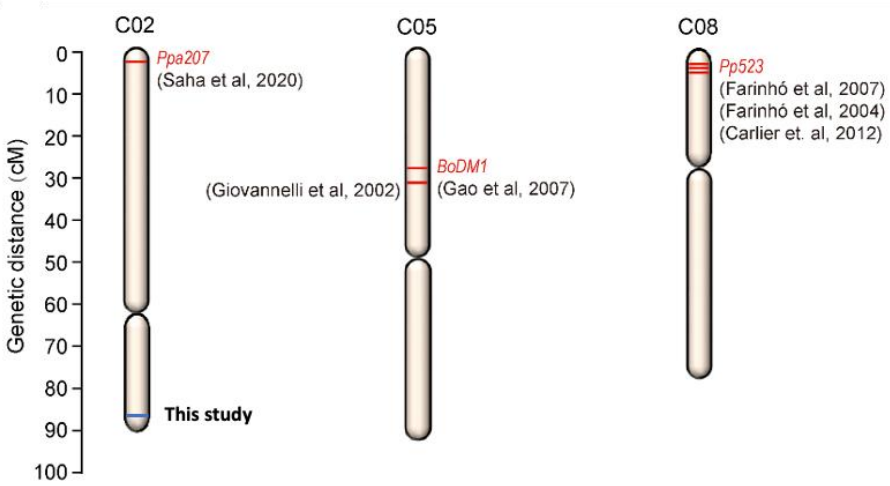


Figure 4.13. Chromosomal locations of mapped downy mildew resistance loci on *B. oleracea*. The blue line pinpoints the location of the resistance mapped in this chapter. Adapted from (Shaw et al., 2021).

The bulk segregant analysis approach I used to map the resistance present

in the 1028 genotype in an F_2 population was very cost-effective, locating the QTL to a 100.000 bp region on the chromosome 2 of the *B. oleracea* pangenome assembly (Figure 4.10). While designing the initial BSA experiment I ran into the question of how many recombinants to include in the bulks and the average sequencing coverage per individual. I decided to aim for coverage of 1-2x per individual, therefore I could not include in the bulks all the F_2 recombinants I assessed, so I selected recombinants with a very clear resistant or susceptible phenotype to Downy Mildew in the bulks. To compare what is the minimum amount of recombinants (1-2x coverage per individual) to be able to map the loci observed in C2, I run the BSA analysis using the sequencing data from RES_BULK1 (40-F_2 recombinants) and SUS_BULK1 (35-F_2 recombinants) and as it can be seen in Figure 4.14 even though the noise has increased and the BFRs are lower compared to using 6 bulks (Figure 4.9), the peak is located in the same region in C2. This data suggests that it is possible to map a trait with a low number of recombinants, but this will greatly depend on several other factors like the complexity of the trait, the quality of the phenotyping, the coverage per individual, the assembly and the quality of the assembly, etc.

Besides the initial BSA analysis using the *B. oleracea* pangenome as a reference that was able to locate the Downy Mildew resistance to a small region in C2. Using the *B. oleracea* HDEM broccoli assembly and tools like NLR annotator, structural variation (SV) was observed in the location of the mapped loci (NLR cluster) and nearby regions in the form of deletions on coding and non-coding regions only present in the susceptible bulks and the DB susceptible parental genotype. In addition, copy number variation (CNV) of the number of predicted NLR genes in the pangenome NLR cluster was also observed in comparison to the number of predicted NLR genes

present in the NLR cluster of the HDEM assembly, ranging from 5 to 1 copy respectively. Several studies have reported that this structural and copy number variation of NLR genes associated with plant pathogen resistance is very frequent (Lee & Chae, 2020; Wang et al., 2021). For this reason, the creation of a 1028 draft genome was necessary to fully resolve the NLR cluster structure present in the 1028 Downy Mildew resistant genotype. The 1028 draft genome assembly and the annotation created using RNAseq data from the parental line revealed that the NLR cluster present in the 1028 draft genome contained a total of 12 predicted NLR genes in a 350Kb region, from which 9 are predicted complete NLRs and 3 are predicted to be incomplete NLRs from which 2 are not expressed (Figure 4.12). This reveals a big difference in copy number variations of NLR genes in the NLR cluster associated with the resistance present in the 1028 genotype, ranging from 1 to 9 predicted NLR in all three assemblies used for the analysis.

In addition, up to now, there was still no high-quality, comprehensive assembled kale (*B. oleracea* var. *acephala*) genome, which hinders greatly basic genetics and genomics research, as well as crop improvement. The closest assembly available for kale cultivars is TO1000, a rapid-cycling Chinese kale, but its growth, morphology and molecular levels are very different compared to a kale variety like 1028 or dwarf blue. For this reason, it is important to have available several genome assemblies for different morphotypes of *B. oleracea* and other brassica species. Three of the most used *B. oleracea* reference genomes by researchers are TO1000, 02-12 and the pangenome were created using short read sequencing and all three assemblies are known for errors and gaps, making them difficult to use for many studies (Lee et al., 2016; Liu et al., 2017; Zhang et al., 2015). In the last few years, long-read sequencing like PacBio HiFi or Nanopore PromethION has

become more affordable, therefore, the use of long-read sequencing on present and future assemblies will be able to improve considerably the genome assembly's integrity, particularly on high repeat regions like NLR clusters. Recently, some assemblies have been published for some brassica cultivars using long-read sequencings, like HDEM (broccoli) and cabbage and cauliflower assemblies (Belser et al., 2018; Guo et al., 2021; Lv et al., 2020). Therefore, this draft assembly of a kale-type *B. oleracea* cultivar using PacBio long-read sequences, in combination with other recent long-read assemblies in other cultivars could help future research to improve *B. oleracea* crops.

To promote marker-assisted selection (MAS) and gene pyramiding into desirable *B. oleracea* cultivars it is essential to find DNA markers associated with the traits needed. Compared to traditional breeding, molecular markers allow for a fast and targeted selection to improve and/or develop new and enhanced commercial varieties. For that reason, attention has switched to molecular breeding, which could make it easier to combine various sources of resistance into existing elite varieties making the resistances more durable over time. With this objective in mind, in this study, the use of BFRs to select SNPs and create highly associated KASP markers with the 1028 Downy Mildew resistant haplotype was very effective. From the 26 SNPs tested on both the pangenome (12 KASPs) and the 1028 draft assembly (14 KASPs), 11 were polymorphic and diagnostic of the Downy Mildew resistance observed in the 1028 genotype. Besides their use as a control for the KASP assay, KASP markers 1, 14, 16 and 17 that are outside of the mapped Downy Mildew resistance loci can be used to observe the allele enrichment that happens the closer the SNPs are to where the loci are located in the genome due to recombination events. It is worth mentioning the biggest difference observed between the number of polymorphic and diagnostic

SNPs from all 26 assessed in the pangenome assembly (17%) compared with the 1028 assembly (64%). Most KASP markers designed using the SNPs in the pangenome assembly have a high number of missing points and/or were not polymorphic or diagnostic of the resistance, even though the BSA variant data suggested it could be a reliable diagnostic SNP. This difference compared to the 1028 assembly could be explained due to the structural variation observed in the NLR cluster and nearby regions for that specific location in the *B. oleracea* Chromosome 2.

In this chapter, the use of different layers of data facilitated the identification of a new source of resistance to downy mildew on the *B. oleracea* Chromosome 2. This resistance was defined as an NLR cluster with a high degree of structural variation when compared to different *B. oleracea* reference genomes. The creation of a draft genome assembly using long-read sequencing technology of the resistant parental genotype 1028 helped resolve the NLR genes in the cluster which will help with the cloning of the candidate NLR genes. Finally, the identification of this new source of resistance in combination with the 11 associated KASP markers will be very helpful in molecular breeding in order to incorporate this Downy Mildew resistance to current and new *B. oleracea* varieties.

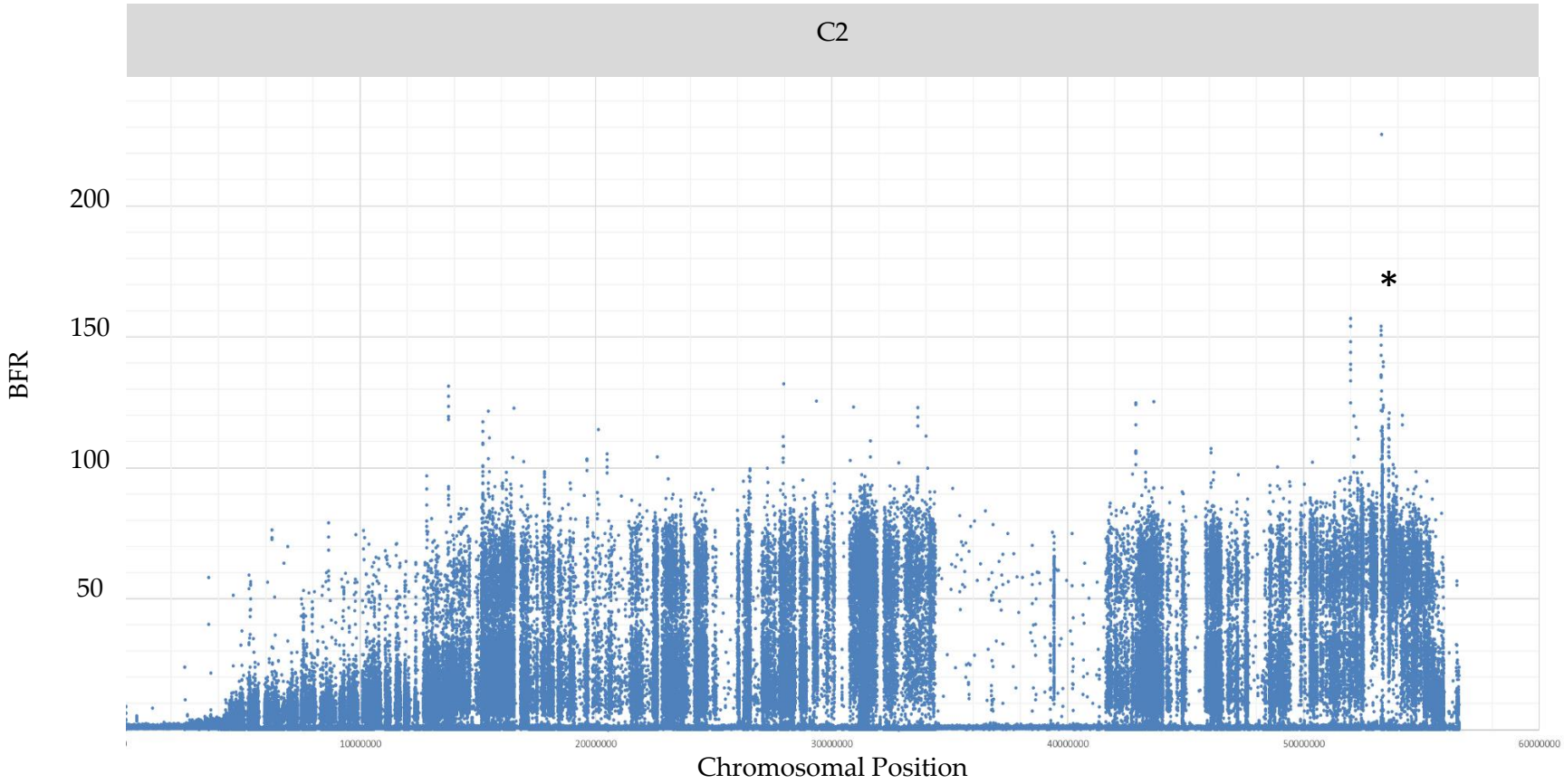


Figure 4.14. Chromosome 2 variants BFRs of only two bulks (RES_BULK1 – 40 F2 recombinants and SUS_BULK1 – 35 F2 recombinants) plotted against theeir position in the *B. oleracea* pangenome assembly. Higher BFR values are indicative of an allele that is highly associated with Downy Mildew resistance. * Location of the peak on 49,000,000-55,000,000 C2 position (Pangenome).

Chapter 5

General discussion

The research presented in this thesis started with the main objective of finding and mapping new sources of resistance to *H. brassicae* in *B. oleracea*. The main steps to achieve this started with the creation of a collection and derivation of 9 *H. brassicae* field isolates into single spore isolates from distinct geographical locations in the UK and USA and the optimization of the *B. oleracea* – downy mildew pathosystem under laboratory conditions (Chapter 2). This collection of isolates was used to identify new sources of *H. brassicae* resistance by assessing 3 different sources of plant material at the cotyledon stage (Chapter 3).

Identified *H. brassicae* resistant individuals were used to create biparental mapping populations. From all the F_2 populations created, a specific cross-segregating for a single dominant gene was selected for genetic mapping using a bulk segregant analysis approach (Chapter 4). This F_2 mapping population was developed by crossing the resistant Downy Mildew breeding line from CN seeds 1028 and the susceptible Downy Mildew commercial variety Dwarf Blue from CN seeds (Figure 4.3). Both parental lines were

also assessed at the adult plant stage and were confirmed to be susceptible (DB) and resistant (1028) at both stages.

The bulk segregant analysis was able to map the downy mildew resistance present in the 1028 parent to a very narrow region (100Kb) in the *B. oleracea* chromosome 2. By using different bioinformatic tools the C2 region was confirmed to be an NLR cluster and by comparing it to other *B. oleracea* genome assemblies a high level of copy number variation was observed.

To fully resolve the NLR cluster sequence, a draft genome assembly of the 1028 resistant parent was created using PacBio HiFi long-read sequencing and expressed genes were annotated using RNAseq data. The resolved NLR cluster in the 1028 genome assembly was confirmed to contain a total of 12 predicted NLR genes in a 350Kb region from which 9 were predicted to be complete and expressed NLR genes (Figure 4.12). Using significantly associated SNPs from the bulk segregant analysis, KASP markers were designed and tested on all 466 F_2 individuals and polymorphic and diagnostic markers for the downy mildew resistance present in 1028 were identified to be used for marker-assisted selection.

Overall, the studies presented in this thesis have identified several pathotype-independent resistant *B. oleracea* genotypes to downy mildew from different sources of plant material (Chapter 3) by using a total of 9 single spored derived *H. brassicae* field isolates collected in 2016, 2018 and 2019 from different *B. oleracea* fields in the UK and USA (Chapter 2). These newly identified sources of downy mildew resistance could provide a solid foundation to improve the management of this disease in the fields of many *B. oleracea* crops, especially vegetable crucifers. Downy incidence has increased over the last few decades in many Brassica growing countries most

likely due to the steady rise of temperatures in some regions of the world caused by climate change, and this will most likely continue, making the disease more widespread. Due to the systemic nature of the disease, various cultural and chemical techniques have been reported to be ineffective against downy mildew in the fields (Barbetti et al., 2011; Vicente et al., 2012) and in some cases the continuous use of certain products has lead to the emergence of new more virulent isolates (Brophy & Laing, 1992; Singh et al., 2013). For this reason, the adoption of genetically resistant plants is one of the most effective, low-cost, and environmentally friendly ways to prevent downy mildew in the fields. To achieve this objective, the sources of resistance identified in this thesis could be incorporated into breeding programs to improve existent commercial *B. oleracea* varieties or to be deployed immediately if the resistant genotypes have adequate agronomical characteristics. Although some of the most resistant genotypes identified in this thesis were also assessed at the adult plant stage, cotyledon assessment was the primary methodology to identify resistant individuals. Some authors have reported that seedling and adult plant stage downy mildew resistance is controlled by different loci (Coelho et al., 1998; Coelho & Monteiro, 2003a; Dickson & Petzoldt, 1993), but based on the results from the genotypes tested in this thesis the resistance observed at the seedling stage was also present in later stages of the plant development. So even though in some genotypes downy mildew resistance at the seedling stage may not be indicative of resistance at later stages, using cotyledons to assess downy mildew resistance is a good indicator of overall resistance at the adult plant stage under laboratory or field conditions (Jensen et al., 1999; B. D. Jensen et al., 1999; Monteiro & Williams, 1989; Wang et al., 2000).

As I reported in Chapter 3, downy mildew physiological specialization was

observed for the isolates used for this thesis. First, by identifying 8 of the 28 resistant genotypes in the 390 *B. oleracea* accessions from Warwick to be *B. rapa* (5) and *B. napus* (3) using a mPCR assay for *Brassica* identification; and second by assessing for downy mildew resistance a *Brassica* diversity panel consisting of 98 different genotypes of all the *Brassica* species in the triangle of U and close relatives (16 *B. oleracea*, 16 *B. rapa*, 15 *B. nigra*, 15 *B. napus*, 15 *B. juncea* and 15 *B. carinata*; 2 *D. erucoides*, 3 *E. sativa* and 1 *A. thaliana*), and third a set of 134 *B. napus* genotypes against an inoculum mix of all 9 single spore isolates from the UK and USA in which only 9 showed susceptibility. These experiments highlight the level of host specificity present in the *H. brassicae* isolates used in this thesis. This specialization has long been documented (Gaumann, 1926; Hiura & Kanegae, 1934; Thung, 1926; Wang, 1944). Since these initial reports, other authors have encountered the same host specialization of downy mildew isolates collected from many brassica species and cultivars (Chang, 1964; Dickinson & Greenhalgh, 1977; Felton & Walker, 1946; Lucas et al., 1994; Masheva et al., 1996; Mehta & Saharan, 1994; Morris & Knox-Davies, 1980; Moss et al., 1988; Natti, 1958; Semb, 1969; Sherriff & Lucas, 1990; Tham et al., 1994) but the methodologies and genotypes they used were not standardised and/or reproducible. For this reason, over the years some authors have tried to create a standardised set of host differentials of isogenic genotypes for *H. brassicae* isolates (Coelho et al., 2012; Mehta & Saharan, 1994). However, these standardised sets haven't been fully adopted and there is still a need to identify and establish a universal set of host differential genotypes for *H. brassicae* to categorize the pathotypes of all reported isolates. Having this information, we could know whether the new and previously reported *H. brassicae* pathotypes including the ones used in this thesis, are the same or different. Therefore, knowing

this will be crucial if the objective is to develop new durable downy mildew resistant Brassicaceae varieties, especially for gene pyramiding strategies.

In Chapter 4, using a biparental F_2 mapping population derived from a cross between the highly downy mildew resistant genotype 1028 and the susceptible commercial variety dwarf blue, previously phenotyped in Chapter 3. I reported how I was able to map a new pathotype-independent *H. brassicae* resistance to a 100Kb region in *B. oleracea* Chromosome 2 by bulk segregant analysis. The resistance observed at seedling stage on the resistant parental genotype 1028 was also confirmed to be present at adult plant stage. In *A. thaliana*, 20 sources of downy mildew resistance have been mapped (Crute et al., 1997), but this attention has not been translated to *Brassica* crops. To date, only 4 Downy mildew resistances have been mapped in *B. oleracea* and only 3 of them have a known chromosomal location, BoDowny Mildew1 in chromosome 5 (Gao et al., 2003; Giovannelli et al., 2002), Pp523 in chromosome 8 (Carlier et al., 2012; Coelho & Monteiro, 2003b; Farinho et al., 2004) and Ppa207 in chromosome 2 (Saha et al., 2020), and to this point, no R gene has been cloned in *B. oleracea* or any other *Brassica* species.

After mapping the *H. brassicae* resistance to a 100Kb region in Chromosome 2, I used different bioinformatic tools and discovered that most genes in the mapping region were NLRs (Chapter 4). By comparing this NLR cluster structure with two different genome assemblies (Pangenome and HDEM) I realised that the copy number of the NLR genes present in the cluster varied from 5 predicted complete NLRs in the pangenome assembly to only 1 in the HDEM assembly. On top of this, checking the bulk read alignments on both assemblies, I observed a read enrichment in some NLR motifs only present in the resistant bulks and big deletions in the NLR cluster and neighbouring regions only present in the susceptible bulks. These two

observations lead me to conclude that the NLR cluster region in the resistant parental line 1028 could also have different copy number variation compared to the other two assemblies and due to the enrichment and the high sequence similarity between NLR genes, the NLR cluster could have more NLR genes.

By using different bioinformatic tools the C2 region was confirmed to be an NLR cluster and by comparing it to other *B. oleracea* genome assemblies a high level of copy number variation was observed. Therefore, to fully resolve the NLR genes included in the cluster region present in the 1028 genotype, a long read sequencing assembly was performed using PacBio HiFi and genes were annotated using RNAseq data. The resulting 1028 draft assembly confirmed the observations and mapped the resistance to a single contig in the assembly. A total of 12 NLR genes were predicted and mapped to a 350Kb region in the contig, from which 9 were predicted to be complete and were expressed NLR genes (Figure 4.12). This copy number variation present in the NLR cluster compared with different assemblies suggests that using a single reference will hardly capture all gene paralogs in a cluster accurately and representatively and that characterizing these additional structural variations requires a reference-agnostic approach (Lee & Chae, 2020). In addition to this, the genome complexity caused by the triplication of the ancestral *Brassica* genome (O'neill & Bancroft, 2000; Parkin et al., 2005) has resulted in the clustering and various structural variations of R genes, making their identification and possible cloning difficult due to the high sequence similarity between this genes which makes them difficult to assemble properly if short reads are used. In addition, even though the draft assembly is not continuous, it is the first long-read sequencing assembly in a *B. oleracea* kale variety to date. In combination with other recent long read

assemblies in other cultivars, this long read sequence could help future research to improve *B. oleracea* crops, especially for research involving high repeat regions like NLR clusters. Finally, the identified candidate R genes in the NLR cluster in combination with the gDNA sequence from the 1028 draft genome assembly and the cDNA from the annotation could be used to synthesize the candidate genes and identify the gene/s responsible for the downy mildew resistance observed in the 1028 genotype.

After mapping the downy mildew resistance by bulk segregant analysis and identifying potential candidate genes it is critical to identify DNA markers linked to the trait in order to facilitate marker-assisted selection and future R gene pyramiding strategies. Even though traditional breeding methods are very effective they are very time-consuming. Marker-assisted selection is a good alternative and could in many cases improve the productivity and accuracy of breeding programs greatly. For this reason, KASP markers were designed and tested on all 466 F_2 individuals and polymorphic and diagnostic markers for the downy mildew resistance present in 1028 were identified in order to be used for marker-assisted selection. Even though, Polymarker (Ricardo H. Ramirez-Gonzalez et al., 2015) was used to design all KASP markers due to the recent whole genome duplication and the ancestral Brassica genome triplication, no closely related gene copies were found. Finally, although this KASP markers were very specific in diagnosing the resistance present in the 1028 genotype, they need to be further validated including more *B. oleracea* breeding varieties.

For decades, breeding for pathogen resistance has solely focused on deploying commercial varieties with a single source of resistance, impacting the stability of the identified resistance genes due to selection pressure between the host and the pathogen (McDonald & Linde, 2002). Therefore, in order to

deploy more durable resistance to downy mildew and other plant pathogens, breeders should avoid the deployment of single dominant and race-specific genes, and focus on breeding several sources of qualitative and quantitative resistances into a single commercial variety. To achieve this, new sources of resistance to *H. brassicae* should be mapped and ideally cloned in *B. olereacea* and other *Brassica* species and wild relatives to increase the number of new sources of resistance available for breeders.

To conclude, this thesis I have identified several sources of resistance in different genotypes to *H. brassicae* isolates collected from *B. oleracea* hosts. By using a bulk segregant analysis approach one of the identified resistances present in the 1028 genotype was mapped to an NLR cluster in the *B. oleracea* Chromosome 2. This investigation including the KASP markers for the new C2 downy mildew resistance and the identified genotypes could provide *Brassica* breeders with new sources of resistance to improve existing commercial varieties or to create new ones. Finally, as discussed earlier, relying on a single assembly for all brassica species could hinder the identification and possible characterization of new resistance genes due to the high level of structural variation usually present in NLR clusters. For this reason, the creation of the first long read Kale draft assembly in combination with other brassicas long read assemblies will help *B. oleracea* future research.

Appendix A

DNA preparation from *Brassica* for Illumina Sequencing

This protocol was developed by Lorelei Bilham at the JIC

Reagents used:

Nuclei Extraction Buffer (NEB):
10 mM TRIS-HCl (pH:9.5)
10 mM EDTA (pH:8.0)
100 mM KCl
500 mM Sucrose
4 mM Spermidine
1 mM Spermine
0.1% ◎-mercaptoethanol

Lysis Buffer:
10% Triton-x
90% Nuclei Extraction Buffer

CTAB Extraction Buffer:
100 mM TRIS-HCl (pH:7.5)
0.7 M NaCl
10 mM EDTA (pH:8.0)
1% CTAB (adjusted in this study, originally in SOP; 0.1%)
1% ◎-mercaptoethanol

Others:
3M NaAc
Chloroform/Isoamyl Alcohol (24:1)
Phenol/Chloroform/Isoamyl Alcohol (24:23:1)
Proteinase K (10 μg/μl)
RNase A (10 μg/μl) RNase T1 (1000u/μl)

<u>Preparation of leaf tissue:</u>

1. Grow plants to the 2-4 true leaf stage before starving by placing a cardboard box over the plants to exclude light for 48hours. This should reducethe amount of sugars/carbohydrates in the leaves.
2. Harvest leaf tissue into foils and snap freeze in liquid nitrogen.
3. Turn on the centrifuge to cool to 4°C and water bath to 60°C

4. Grind 3g of leaf tissue to a fine powder in liquid N2

Nuclei extraction:

5. Transfer to 50ml falcon tube containing 40ml ice-cold nuclei extraction buffer
6. Vortex until ground disuse is evenly distributed through buffer
7. Filter the homogenised tissue through two layers of Miracloth (Calbiochem) to remove leaf tissue material

Lysis:

8. Add 8ml lysis buffer to filtered homogenate and keep on ice for 2min
9. Centrifuge at 1,000g for 20min at 4°C. This should pellet the nuclei at the bottom of the tube.
10. Remove the supernatant
11. Resuspend the pellet in 3ml of CTAB extraction buffer
12. Incubate in the waterbath for 30min at 60°C
13. Drop waterbath temperature to 37°C

Chloroform/isoamyl alcohol (24:1) extraction:

14. Perform x1 chloroform/isoamyl alcohol (24:1) clean by adding equal volumes of chloroform/isoamyl alcohol
15. Rotate for 10min in cold room
16. Centrifuge for 10min at 1,000g

RNase/Protinase treatment:

17. Remove the aqueous phase into a clean 15ml falcon tube and add RNase T1 (0.6µl stock) and RNaseA (15µl stock) to the sample so the final concentrations are 50units/ml and 50µg/ml respectively
18. Mix the sample by inverting the tube 10-20 times
19. Incubate in the waterbath at 37°C for 45min
20. Make the solution 150µg/ml Protinase K (1/100 of vol if stock solution is at 10µg/µl)
21. Incubate in waterbath at 37°C for 45min

Phenol/chloroform/isoamyl alcohol (24:23:1) extractions:

22. Add an equal volume of phenol/chloroform/isoamyl alcohol (24:23:1)
23. Invert x20 to mix
24. Centrifuge for 10min at 1,000g
25. Retain the supernatant and repeat steps 22-24

Chloroform/isoamyl alcohol (24:1) extraction:

26. Perform x1 chloroform/isoamyl alcohol (24:1) clean by adding equal volumes of chloroform/isoamyl alcohol
27. Rotate for 10min in cold room
28. Centrifuge for 10min at 1,000g. Retain top phase for precipitation by removing into a clean 15ml falcon tube

DNA precipitation:

29. Add 10% 3M NaAc and 2.5-3 x volume of 100% ethanol
30. If DNA spindles out of solution then centrifuge at 1,000g for 1min to gently pellet and then wash with 3-5ml 75%
31. Gently centrifuge and air dry DNA
32. Gently resuspend DNA so as not to shear, in 50µl dH_2O

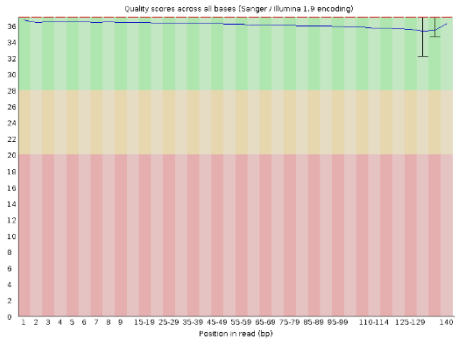
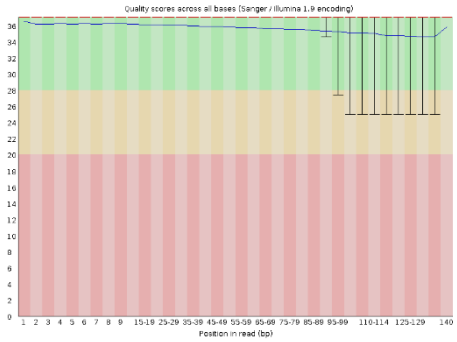
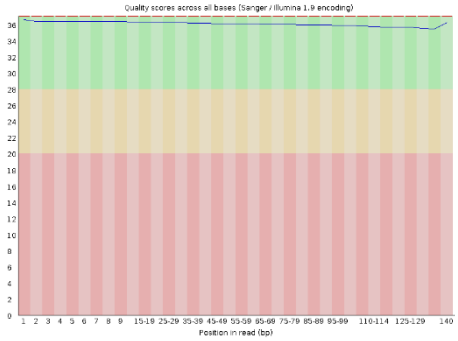
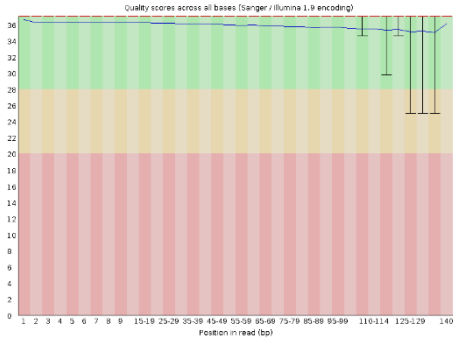
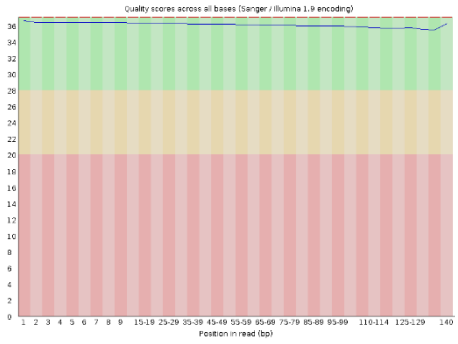
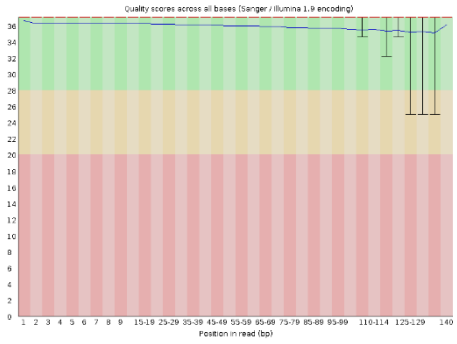
QA:

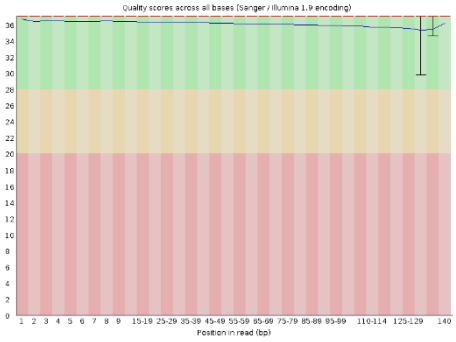
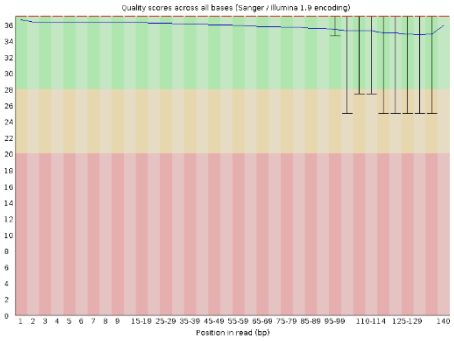
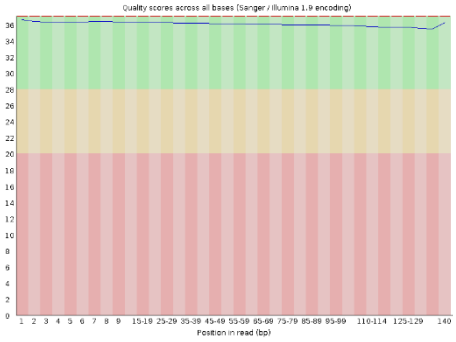
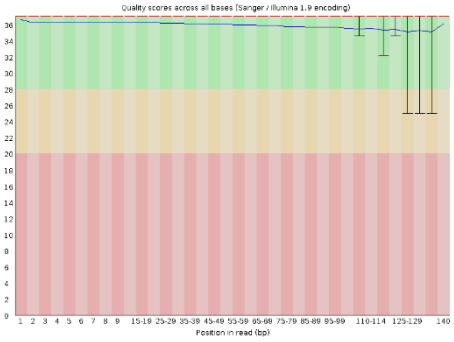
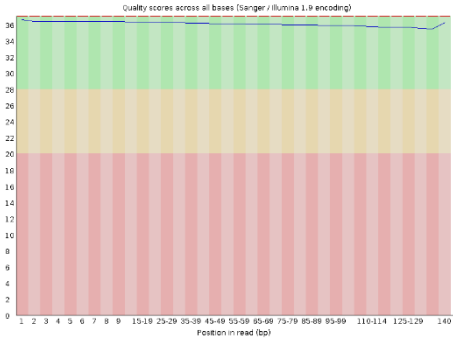
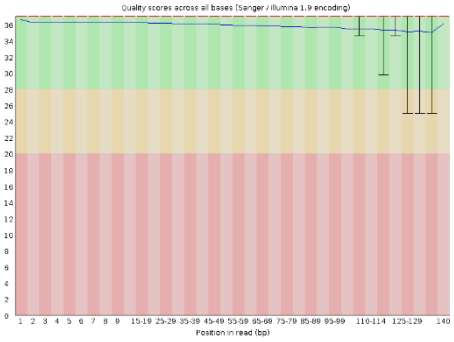
33. Quantify by Qubit BR and run 100-250ng on 0.8-1% gel

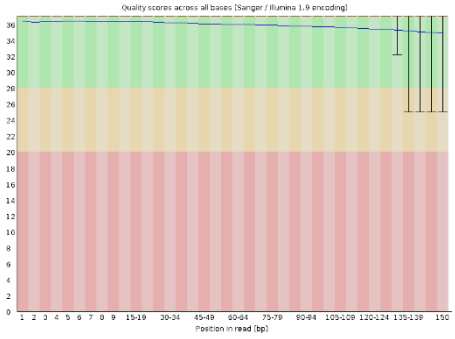
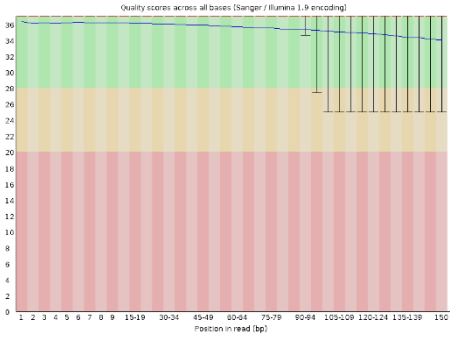
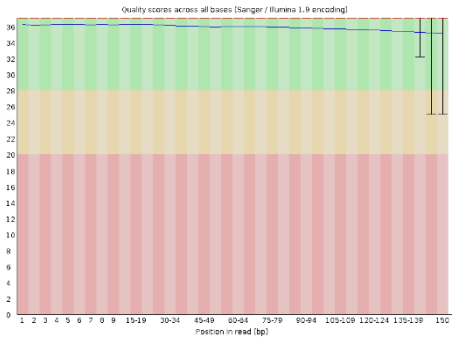
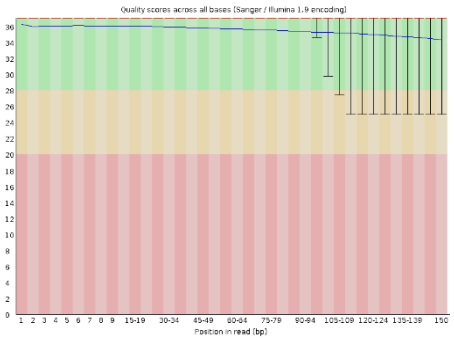
Appendix B

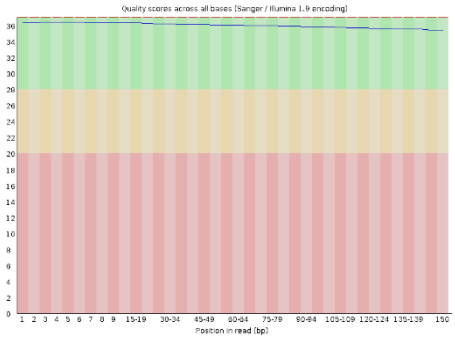
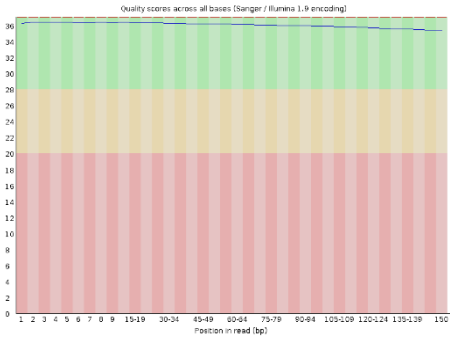
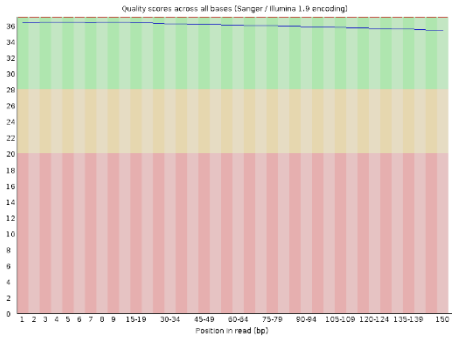
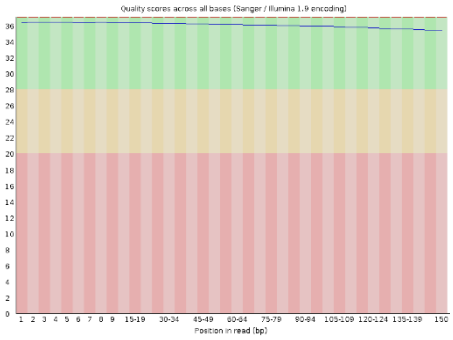
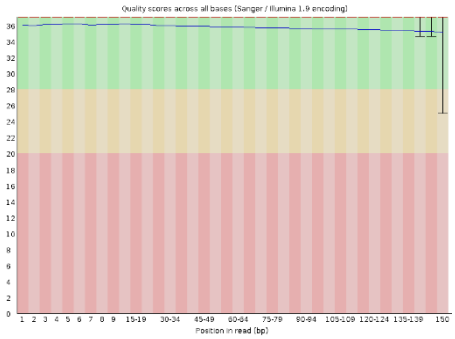
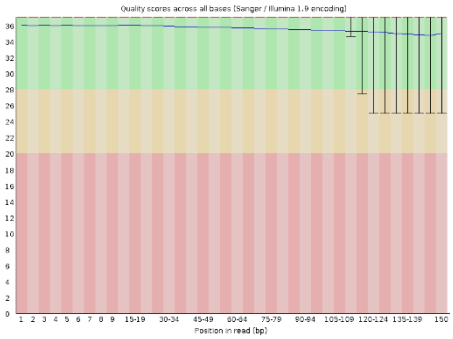
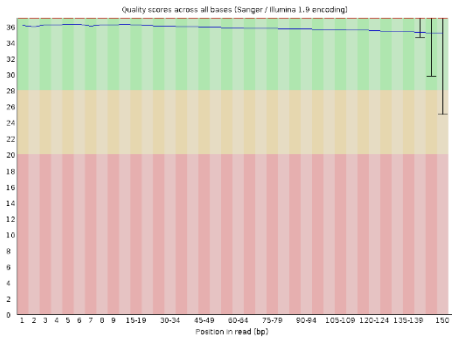
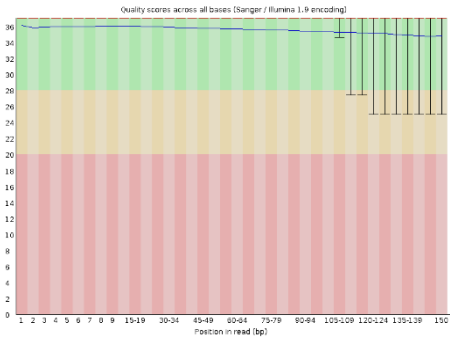
Quality Control Illumina DNA and RNA Reads

B1. Per base quality scores

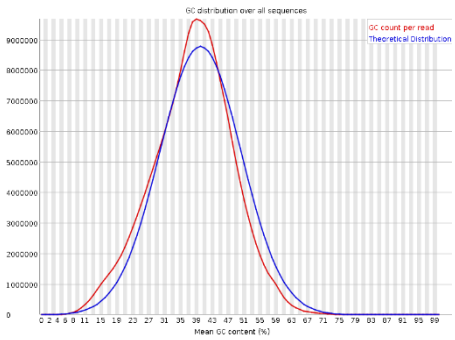
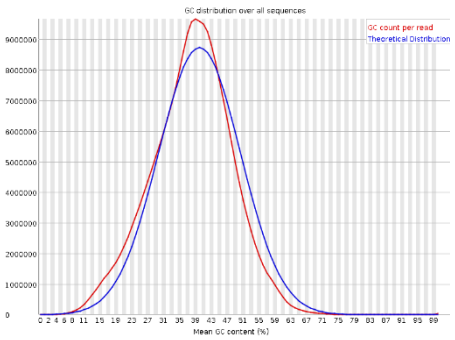
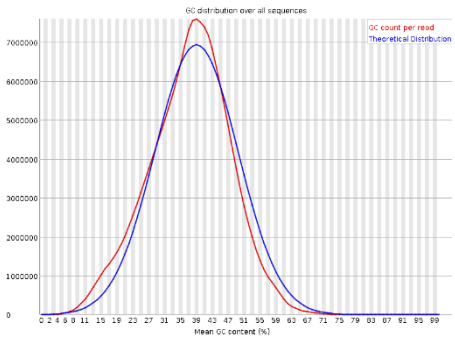
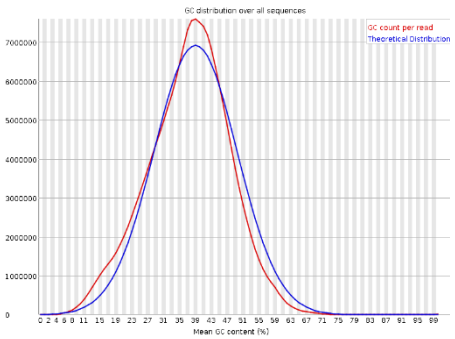
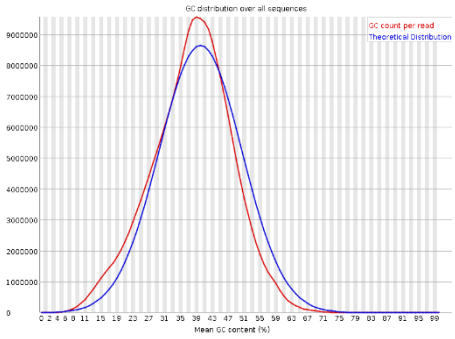
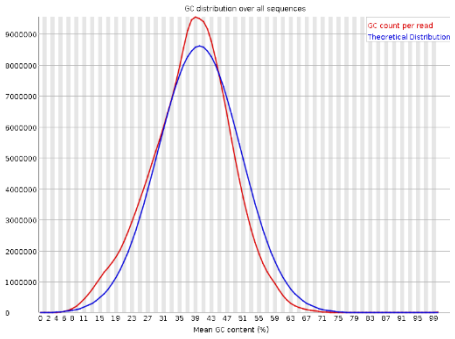
Sample	Read 1	Read 2
RES_BULK1		
RES_BULK2		
RES_BULK3		

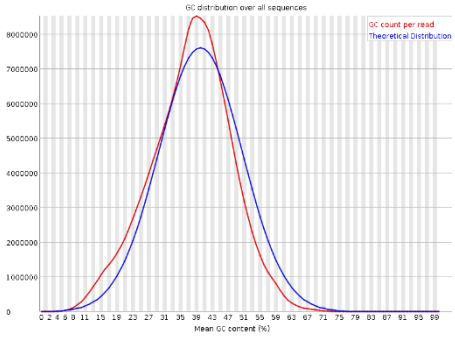
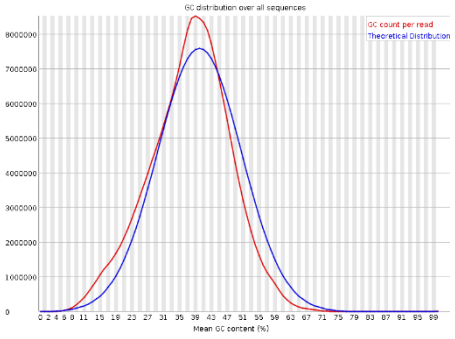
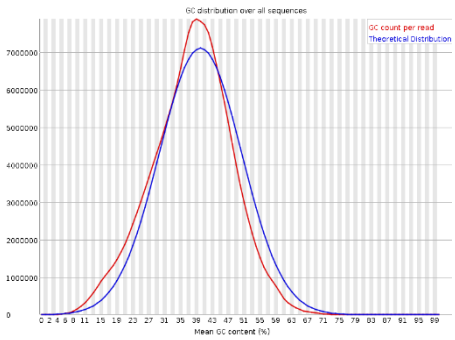
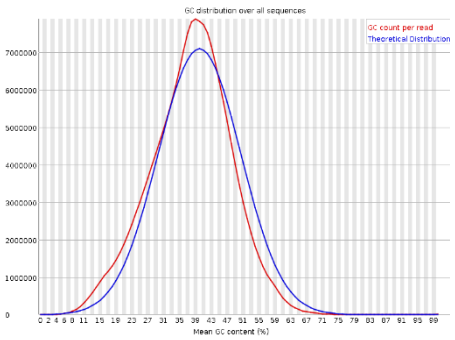
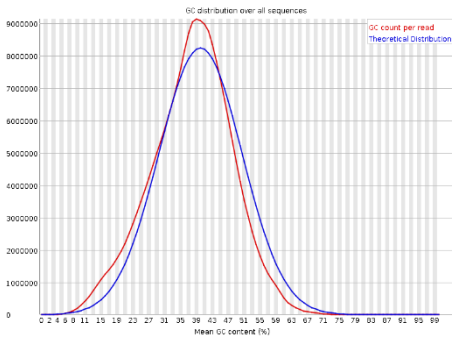
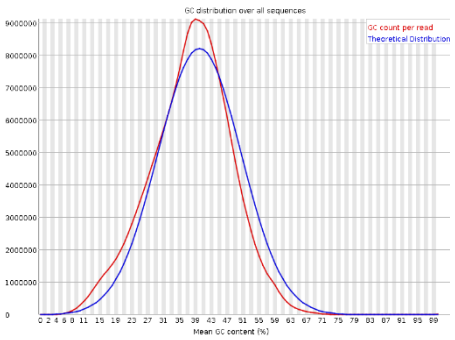
Sample	Read 1	Read 2
SUS_BULK1		
SUS_BULK2		
SUS_BULK3		

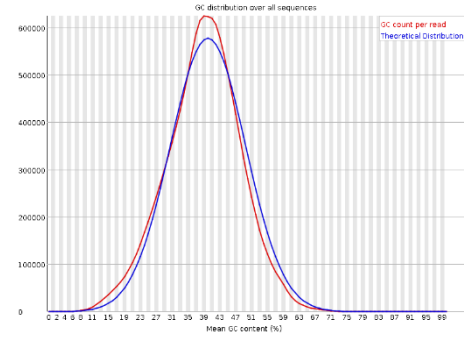
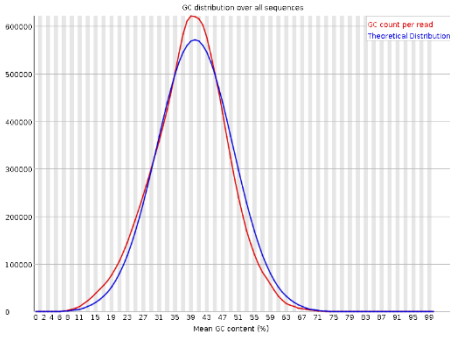
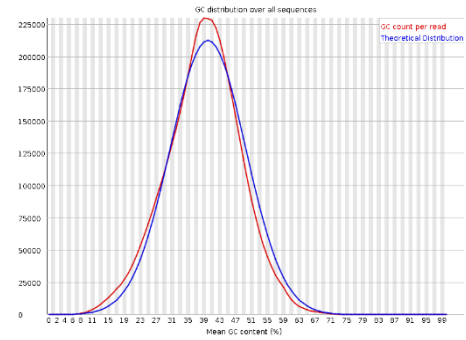
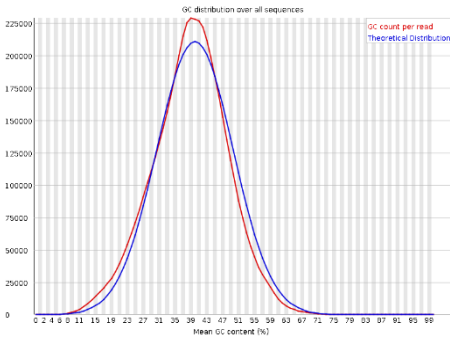
Sample	Read 1	Read 2
DB_SUS		
		

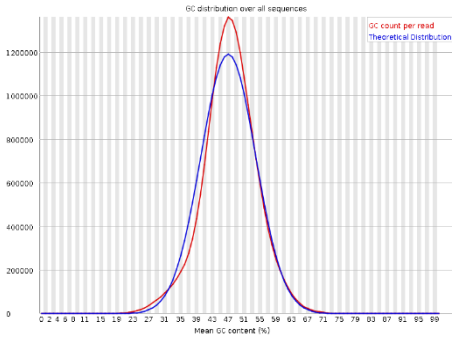
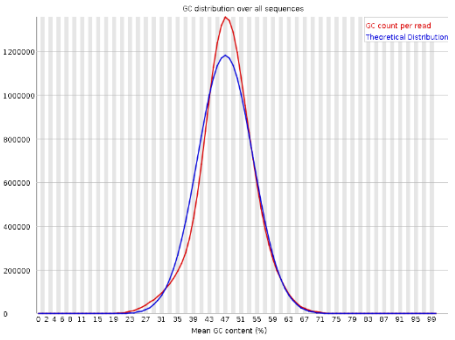
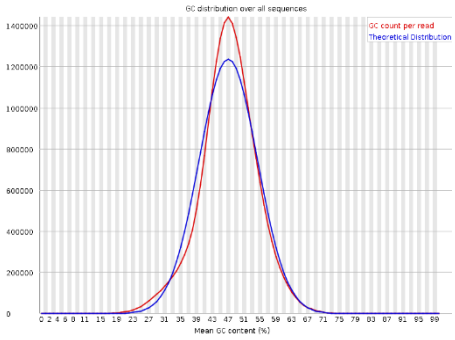
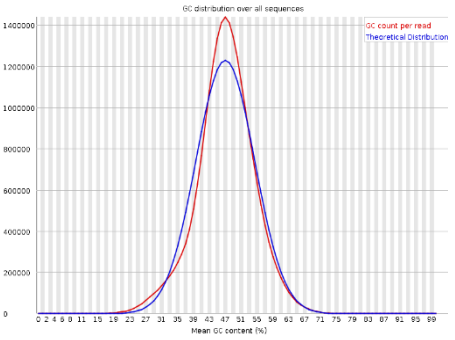
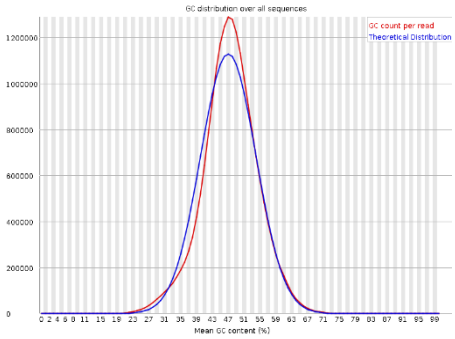
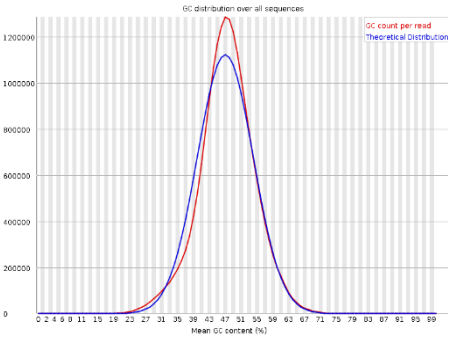
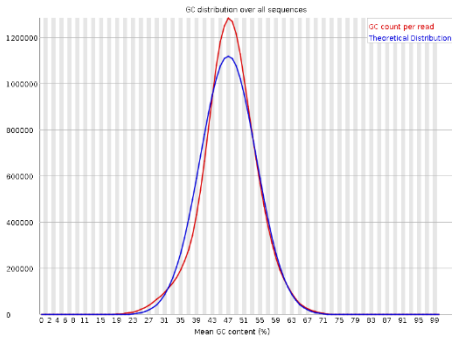
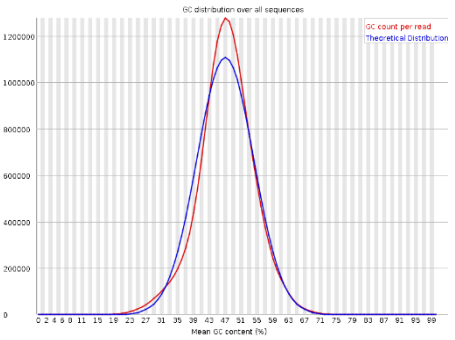
Sample	Read 1	Read 2
1028CON1		
1028INO2		
DB03CON_1		
DB03CON_2		

B2. Sequence GC content

Sample	Read 1	Read 2
RES_BULK1		
RES_BULK2		
RES_BULK3		

Sample	Read 1	Read 2
SUS_BULK1		
SUS_BULK2		
SUS_BULK3		

Sample	Read 1	Read 2
DB_SUS		
		

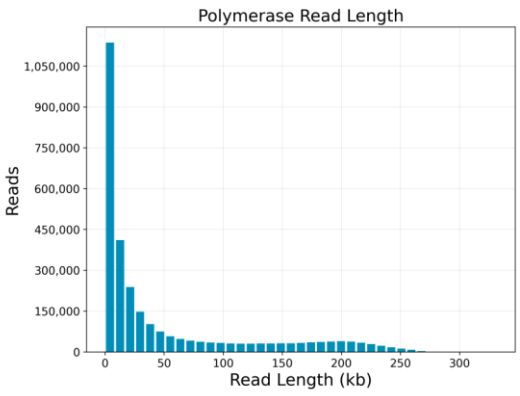
Sample	Read 1	Read 2
1028CON1		
1028INO2		
DB03CON_1		
DB03CON_2		

Appendix C

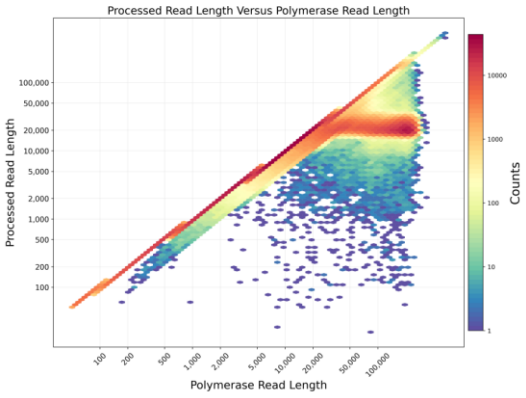
PacBio HiFi SMRT Cell Run QC

C1. Raw Data Report

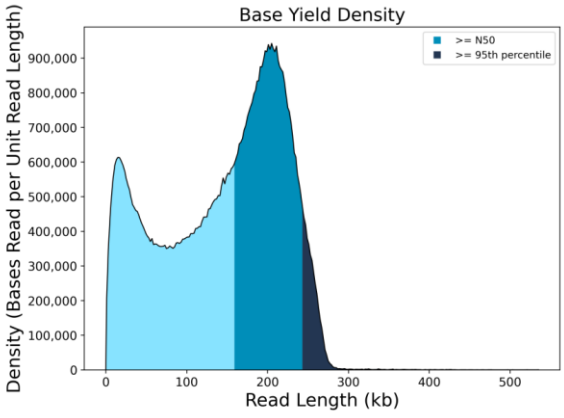
Analysis Metric	Value
Polymerase Read Bases	143,244,712,207
Polymerase Reads	2,931,042
Polymerase Read Length (mean)	48,872
Polymerase Read N50	160,750
Longest Subread Length (mean)	16,735
Longest Subread N50	24,750
Unique Molecular Yield	47,063,584,768



(a) Polymerase Read Length



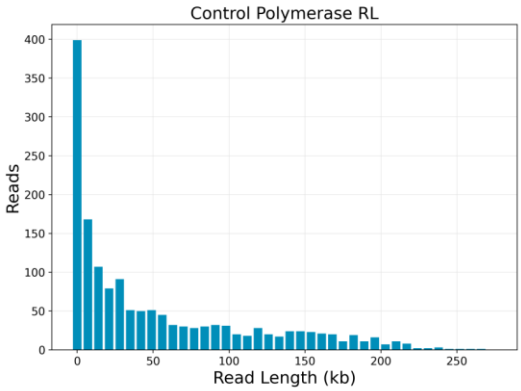
(b) Longest Subread Length Versus Polymerase Read Length



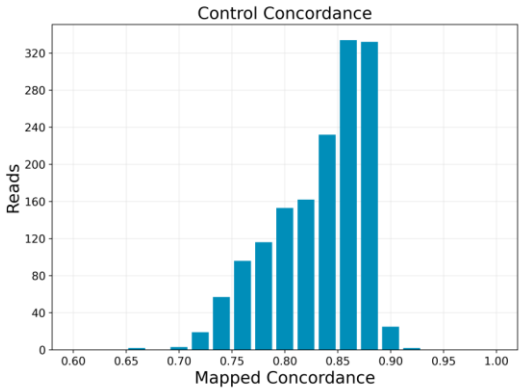
(c) Base Yield Density

C2. Control Report

Analysis Metric	Value
Number of of Control Reads	1,533
Control Read Length Mean	54,487
Control Read Concordance Mean	0.84
Control Read Concordance Mode	0.87



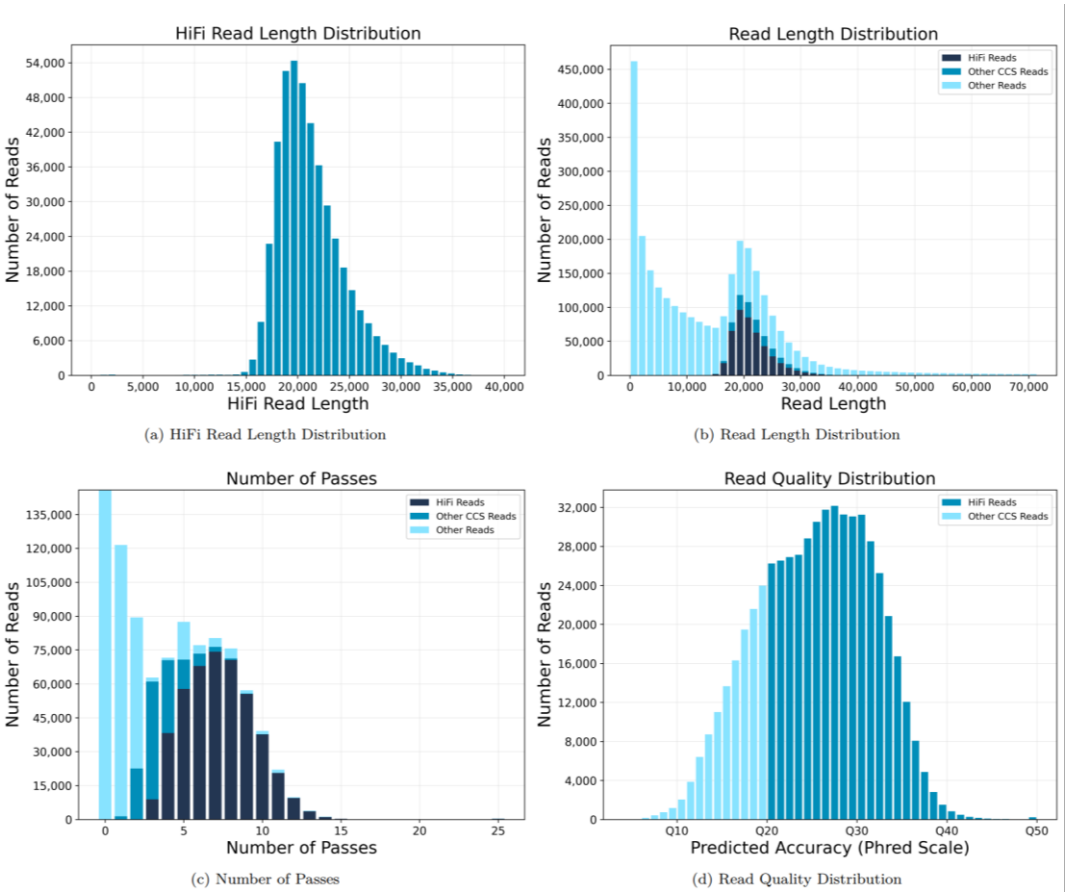
(a) Control Read Length



(b) Control Read Quality

C3. CCS Analysis Report

Analysis Metric	Value
HiFi Reads	446,709
HiFi Yield (bp)	9,507,944,809
HiFi Read Length (mean, bp)	21,284
HiFi Read Quality (median)	Q27
HiFi Number of Passes (mean)	7
<Q20 Reads	129,616
<Q20 Yield (bp)	2,918,013,187
<Q20 Read Length (mean, bp)	22,512
<Q20 Read Quality (median)	Q17

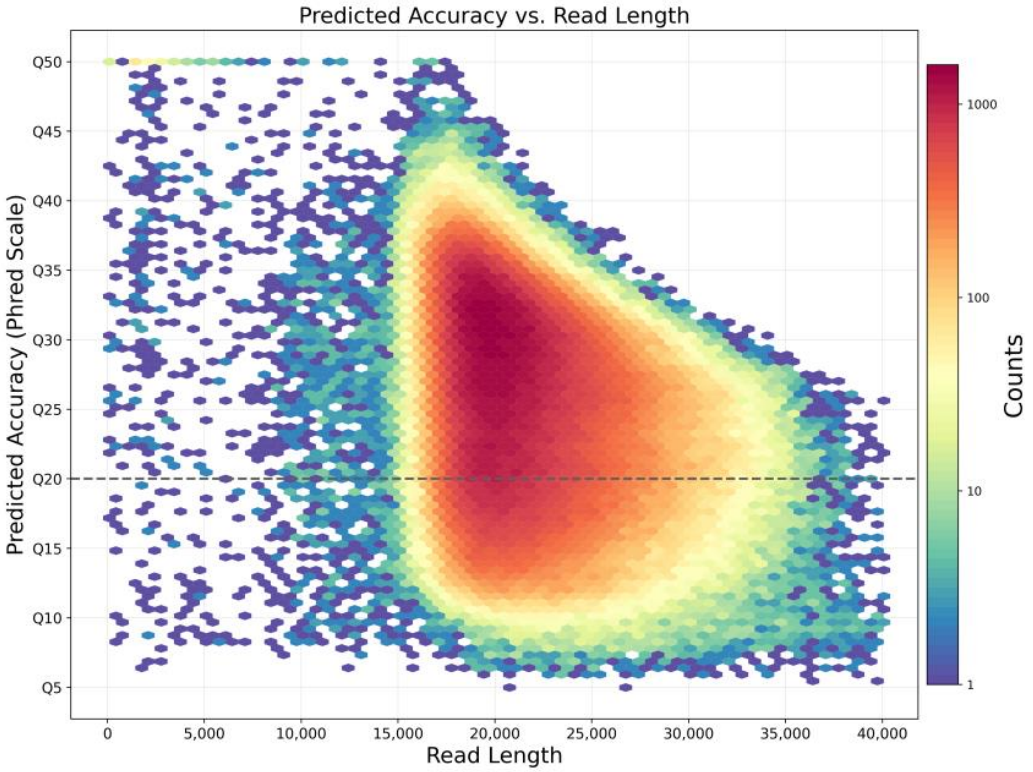


(a) HiFi Read Length Distribution

(b) Read Length Distribution

(c) Number of Passes

(d) Read Quality Distribution



(e) Predicted Accuracy vs. Read Length

Bibliography

Achar, P. (1992). Dose-response relationship of Peronospora parasitica in Raphanus sativus. *Phyton*.

Achar, P. N. (1995). Tissue Culture Technique to Determine the Viability of Peronospora parasitica in *Brassica oleracea*. *Journal of Phytopathology*, *143*(11-12), 647-649. https://doi.org/https://doi.org/10.1111/j.1439-0434.1995.tb00217.x

Al-Shehbaz, I. A. (2012). A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon*, *61*(5), 931-954.

Arias, T., Beilstein, M. A., Tang, M., McKain, M. R., & Pires, J. C. (2014). Diversification times among *Brassica* (Brassicaceae) crops suggest hybrid formation after 20 million years of divergence. *American journal of botany*, *101*(1), 86-91.

Babraham-Bioinformatics. (2012). *FastQC A Quality Control tool for High Throughput Sequence Data.* In (Version 0.11.9) http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barbetti, M. J., Li, C. X., Garg, H., Li, H., Banga, S. K., Banga, S. S., Sandhu, P. S., Singh, R., Singh, D., Liu, S. Y., Gurung, A. M., & Salisbury, P. A. (2011). Host resistance in oilseed Brassicas against Sclerotinia-renewed hope for managing a recalcitrant pathogen. . pp. 713-715, In: Proceedings of the 13th International Rapeseed Congress. Prague, Czech Republic.,

Barnes, W. (1968). Development of downy mildew resistant broccoli, cabbage, and collards. *HortScience*, *3*, 78.

Baxter, L., Tripathy, S., Ishaque, N., Boot, N., Cabral, A., Kemen, E., Thines, M., Ah-Fong, A., Anderson, R., & Badejoko, W. (2010). Signatures of adaptation to obligate biotrophy in the Hyaloperonospora arabidopsidis genome. *Science*, *330*(6010), 1549-1551.

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*(43), 18724-18728.

Belbahri, L., Calmin, G., Pawlowski, J., & Lefort, F. (2005). Phylogenetic analysis and real time PCR detection of a presumbably undescribed Peronospora species on sweet basil and sage. *Mycological Research*, *109*(11), 1276-1287.

Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F.-C., Falentin, C., Genete, M., Berrabah, W., Chèvre, A.-M., Delourme, R., Deniot, G.,

Denoeud, F., Duffé, P., Engelen, S., Lemainque, A., Manzanares-Dauleux, M., Martin, G., Morice, J., Noel, B., . . . Aury, J.-M. (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat Plants*, *4*(11), 879-887. https://doi.org/10.1038/s41477-018-0289-4

Bevan, M. W., Uauy, C., Wulff, B. B., Zhou, J., Krasileva, K., & Clark, M. D. (2017). Genomic innovation for crop improvement. *Nature*, *543*(7645), 346-354.

Bigeard, J., Colcombet, J., & Hirt, H. (2015). Signaling mechanisms in pattern-triggered immunity (PTI). *Mol Plant*, *8*(4), 521-539. https://doi.org/10.1016/j.molp.2014.12.022

Bittner-Eddy, P. D., Crute, I. R., Holub, E. B., & Beynon, J. L. (2000). RPP13 is a simple locus in Arabidopsis thaliana for alleles that specify downy mildew resistance to different avirulence determinants in Peronospora parasitica. *Plant J*, *21*(2), 177-188. https://doi.org/10.1046/j.1365-313x.2000.00664.x

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120. https://doi.org/10.1093/bioinformatics/btu170

Boller, T., & Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu Rev Plant Biol*, *60*, 379-406. https://doi.org/10.1146/annurev.arplant.57.032905.105346

Bonnet, A., & Blancard, D. (1987). Resistance of radish (Raphanus sativus L.) to downy mildew, Peronospora parasitica. *Cruciferae Newsletter*(12), 98-99. https://hal.inrae.fr/hal-02728694

Botella, M. A., Parker, J. E., Frost, L. N., Bittner-Eddy, P. D., Beynon, J. L., Daniels, M. J., Holub, E. B., & Jones, J. D. (1998). Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *Plant Cell*, *10*(11), 1847-1860.

Branca, F., Bahcevandziev, K., Perticone, V., & Monteiro, A. (2005). Sources of resistance to downy mildew (Peronospora parasitica (Pers. (ex Fr.) Fr.) in Sicilian germplasm of cauliflower and broccoli. *Biodiversity & Conservation*, *14*(4), 841-848. https://doi.org/10.1007/s10531-004-0652-9

Brophy, T., & Laing, M. (1992). Screening of fungicides for the control of downy mildew on container-grown cabbage seedlings. *Crop Protection*, *11*(2), 160-164.

Butler, E. (1918). Fungi and diseases in plant. *Thaker Spink & Co., Calcutta, pp 297–300*.

Byford, W. (1967). Host specialization of Peronospora farinosa on Beta, Spinacia and Chenopodium. *Transactions of the British Mycological Society*, *50*(4), 603-607.

Cai, C., Wang, X., Liu, B., Wu, J., Liang, J., Cui, Y., Cheng, F., & Wang, X. (2017). *Brassica* rapa genome 2.0: a reference upgrade through sequence re-assembly and gene re-annotation. *Molecular plant*, *10*(4), 649-651.

Cai, X., Chang, L., Zhang, T., Chen, H., Zhang, L., Lin, R., Liang, J., Wu, J., Freeling, M., & Wang, X. (2021). Impacts of allopolyploidization and structural variation on intraspecific diversification in *Brassica* rapa. *Genome Biology*, *22*(1), 1-24.

Carlier, J. D., Alabaça, C. A., Coelho, P. S., Monteiro, A. A., & Leitão, J. M. (2012). The downy mildew resistance locus Pp523 is located on chromosome C8 of *Brassica oleracea* L. *Plant Breeding*, *131*(1), 170-175.

Carlsson, M., Von Bothmer, R., & Merker, A. (2004). Screening and evaluation of resistance to downy mildew (Peronospora parasitica) and clubroot (Plasmodiophora brassicae) in genetic resources of *Brassica oleracea*. *Hereditas*, *141*(3), 293-300. https://doi.org/10.1111/j.1601-5223.2004.01818.x

Carvalho, T., & Monteiro, A. (1996). Preliminary study on the inheritance of resistance to downy mildew (Peronospora parasitica (Pers. ex. Fr.) Fr.) at cotyledon stage in Tronchuda cabbage'Algarvia'. *Cruciferae Newsletter*(18), 104-105.

Chang, I. (1964). A preliminary study on the physiological differentiation of the downy mildews (Peronospora parasitica (Pers.) Fr.) of Chinese cabbage and other cruciferous vegetables in teh vicinity of Peking and Tientsin. *Acta Phytopathol. Sin.*, *7*, 33-44.

Chang, I., Xu, R., & Chiu, W. (1963). On the primary sources of infection of the downy mildew of Chinese cabbage caused by Peronospora parasitica (Pers.) Fr. and the limited systemic infection of seedlings. . *Acta Phytopath Sin 6:153–162*.

Channon, A. G. (1981). Downy mildew of Brassicas. In: Spencer Downy Mildew (ed) The downy mildews. *Academic Press, London. pp 321–339.*

Chisholm, S. T., Coaker, G., Day, B., & Staskawicz, B. J. (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*, *124*(4), 803-814. https://doi.org/10.1016/j.cell.2006.02.008

Choi, Y.-J., Danielsen, S., Lübeck, M., Hong, S.-B., Delhey, R., & Shin, H.-D. (2010). Morphological and molecular characterization of the causal agent of downy mildew on quinoa (Chenopodium quinoa). *Mycopathologia*, *169*(5), 403-412.

Choi, Y.-J., Denchev, C. M., & Shin, H.-D. (2008). Morphological and molecular analyses support the existence of host-specific Peronospora species infecting Chenopodium. *Mycopathologia*, *165*(3), 155-164.

Choi, Y.-J., Hong, S.-B., & Shin, H.-D. (2007). Extreme size and sequence variation in the ITS rDNA of Bremia lactucae. *Mycopathologia*, *163*(2), 91-95.

Choi, Y.-J., Klosterman, S. J., Kummer, V., Voglmayr, H., Shin, H.-D., & Thines, M. (2015). Multi-locus tree and species tree approaches toward resolving a complex clade of downy mildews (Straminipila, Oomycota), including pathogens of beet and spinach. *Molecular phylogenetics and evolution*, *86*, 24-34.

Choi, Y.-J., Shin, H.-D., & Thines, M. (2009). Two novel Peronospora species are associated with recent reports of downy mildew on sages. *Mycological Research*, *113*(12), 1340-1350.

Chorin, M. (1946). Downy mildew on cauliflower curds. *Palestine J Bot Jerusalem Series-I*, *2*, 258-259.

Chou, C. (1970). An electron-microscope study of host penetration and early stages of haustorium formation of Peronospora parasitica (Fr.) Tul. on cabbage cotyledons. *Ann Bot*, *34*(1), 189-204.

Coates, M. E., & Beynon, J. L. (2010). Hyaloperonospora arabidopsidis as a pathogen model. *Annual Review of Phytopathology*, *48*, 329-345.

Coelho, P., Bahcevandziev, K., ValÃ©rio, L., Monteiro, A., Leckie, D., Astley, D., Crute, I. R., & Boukema, I. (1998). The relationship between cotyledon and adult plant resistance to downy mildew (peronospora parasitica) in brassica oleracea.

Coelho, P. S., & Monteiro, A. A. (2003a). Expression of resistance to downy mildew at cotyledon and adult plant stages in *Brassica oleracea* L [journal article]. *Euphytica*, *133*(3), 279-284. https://doi.org/10.1023/a:1025787608919

Coelho, P. S., & Monteiro, A. A. (2003b). Inheritance of downy mildew resistance in mature broccoli plants [journal article]. *Euphytica*, *131*(1), 65-69. https://doi.org/10.1023/a:1023008619400

Coelho, P. S., Valério, L., & Monteiro, A. A. (2009). Leaf position, leaf age and plant age affect the expression of downy mildew resistance in *Brassica oleracea*. *European Journal of Plant Pathology*, *125*(2), 179-188.

Coelho, P. S., Vicente, J. G., Monteiro, A. A., & Holub, E. B. (2012). Pathotypic diversity of Hyaloperonospora brassicae collected from *Brassica oleracea* [journal article]. *European Journal of Plant Pathology*, *134*(4), 763-771. https://doi.org/10.1007/s10658-012-0052-z

Constantinescu, O. (1991). *An annotated list of 'Peronospora' names*. University of Uppsala.

Constantinescu, O., & Fatehi, J. (2002). Peronospora-like fungi (Chromista, Peronosporales) parasitic on Brassicaceae and related hosts. *Nova Hedwigia*, *74*, 291-338.

Corda, A. (1837). Icones fungorum hucusuque cognitorum 1:20.

Couto, D., & Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. *Nat Rev Immunol*, *16*(9), 537-552. https://doi.org/10.1038/nri.2016.77

Crute, I., Norwood, J., & Gordon, P. (1985). Resistance to phenylamide fungicides in lettuce and brassica downy mildew. Proceedings of the Mixture Centenary Meeting, Bordeaux, British Crop Protection Council Monograph,

Crute, I. R., Holub, E. B., & Burdon, J. J. (1997). The gene-for-gene relationship in plant-parasite interactions.

Cui, H., Tsuda, K., & Parker, J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annu Rev Plant Biol*, *66*, 487-511. https://doi.org/10.1146/annurev-arplant-050213-040012

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

Davison, E. M. (1968). Development of sporangiophores of Peronospora parasitica (Per. ex Fr.) *Fr. Annals of Botany 32: 623-631*.

De Smet, I., Voss, U., Jürgens, G., & Beeckman, T. (2009). Receptor-like kinases shape the plant. *Nat Cell Biol*, *11*(10), 1166-1173. https://doi.org/10.1038/ncb1009-1166

Dick, M. (2001). The peronosporomycetes. In *Systematics and evolution* (pp. 39-72). Springer.

Dick, M. (2002). Towards and Understanding of the Evolution of the Downy Mildews. *Advances in downy mildew research*, 1-57.

Dickinson, C., & Greenhalgh, J. (1977). Host range and taxonomy of Peronospora on crucifers. *Transactions of the British Mycological Society*, *69*(1), 111-116.

Dickson, M., & Petzoldt, R. (1996). Seedling downy mildew resistance in *B. oleracea*. *Cruciferae Newsletter*.

Dickson, M. H., & Petzoldt, R. (1993). Plant Age and Isolate Source Affect Expression of Downy Mildew Resistance in Broccoli. *HortScience*, *28*, 730–731.

Edwards, D., Batley, J., Parkin, I., & Kole, C. (2011). *Genetics, genomics and breeding of oilseed Brassicas*. CRC Press.

Erik Garrison, G. M. (2012). *Haplotype-based variant detection from short-read sequencing*. In https://arxiv.org/abs/1207.3907

Eshraghi, L., Barbetti, M., Li, H., Danehloueipour, N., & Sivasithamparam, K. (2007). Resistance in oilseed rape (*Brassica* napus) and Indian mustard (*Brassica* juncea) to a mixture of Pseudocercosporella capsellae isolates from Western Australia. *Field Crops Research*, *101*(1), 37-43.

Farinho, M., Coelho, P., Carlier, J., Svetleva, D., Monteiro, A., & Leitao, J. (2004). Mapping of a locus for adult plant resistance to downy mildew in broccoli (*Brassica oleracea* convar. italica). *Theor Appl Genet*, *109*(7), 1392-1398. https://doi.org/10.1007/s00122-004-1747-0

Farinhó, M., Coelho, P., Monteiro, A., & Leitão, J. (2007). SCAR and CAPS markers flanking the *Brassica oleracea* L. Pp523 downy mildew resistance locus demarcate a genomic region syntenic to the top arm end of Arabidopsis thaliana L. chromosome 1. *Euphytica*, *157*(1), 215-221.

Farnham, M. W., Wang, M., & Thomas, C. E. (2002). A single dominant gene for downy mildew resistance in broccoli [journal article]. *Euphytica*, *128*(3), 405-407. https://doi.org/10.1023/a:1021279426050

Felton, M. W., & Walker, J. (1946). Environal factors affecting downy mildew of cabbage.

Gao, M., Li, G., Yang, B., Farnham, M., & Quiros, C. (2003). Search for candidate genes for downy mildew resistance in broccoli. Plant and Animal Genomes XI Conf. Abstract,

Gao, M., Li, G., Yang, B., Qiu, D., Farnham, M., & Quiros, C. (2007). High-density *Brassica oleracea* linkage map: identification of useful new linkages. *Theoretical and Applied Genetics*, *115*(2), 277-287.

García-Blázquez, G., Göker, M., Voglmayr, H., Martín, M. P., Tellería, M. T., & Oberwinkler, F. (2008). Phylogeny of Peronospora, parasitic on Fabaceae, based on ITS sequences. *Mycological Research*, *112*(5), 502-512.

Gaumann, E. (1918). Uber die Formen der Peronospora parasitica Ein Beitrag Zur Spezies frage bei parasitischen Pilzen. . *Beih Bot Zentralblatt 35:395–533*.

Gaumann, E. (1923). Contributions toward a monograph of the genus Peronospora corda. *Beitrage zur kryptogamenflora de schweiz*, *5*, 360.

Gaumann, E. (1926). On the specialization of downy mildew (Peronospora brassicae Gaum.) on cabbage and related species. *Landwirtschaftliches Jahrbuch der Schweiz*, *40*, 463-468.

Ge, X., Li, H., Han, S.-e., Sivasithamparam, K., & Barbetti, M. (2008). Evaluation of Australian *Brassica* napus genotypes for resistance to

the downy mildew pathogen, Hyaloperonospora parasitica. *Australian journal of agricultural research*, *59*(11), 1030-1034.

Giovannelli, J. L., Farnham, M. W., Wang, M., & Strand, A. E. (2002). Development of sequence characterized amplified region markers linked to downy mildew resistance in broccoli. *Journal of the American Society for Horticultural Science*, *127*(4), 597-601.

Giovannoni, J. J., Wing, R. A., Ganal, M. W., & Tanksley, S. D. (1991). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic acids research*, *19*(23), 6553-6568.

Gisi, U. (2002). Chemical control of downy mildews. *Advances in downy mildew research*, 119-159.

Göker, M. (2006). Anmerkungen zur aktuellen Taxonomie der Falschen Mehltaupilze und einiger ihrer Verwandter. *Mitteilungen aus der Biologischen Bundesanstalt für Land-und Forstwirtschaft Berlin-Dahlem*, *405*, 142-149.

Goker, M., Garcia-Blazquez, G., Voglmayr, H., Telleria, M. T., & Martin, M. P. (2009). Molecular taxonomy of phytopathogenic fungi: a case study in Peronospora. *PLoS One*, *4*(7), e6319. https://doi.org/10.1371/journal.pone.0006319

Göker, M., Riethmüller, A., Voglmayr, H., Weiss, M., & Oberwinkler, F. (2004). Phylogeny of Hyaloperonospora based on nuclear ribosomal internal transcribed spacer sequences. *Mycological Progress*, *3*(2), 83-94.

Göker, M., Voglmayr, H., Blázquez, G. G., & Oberwinkler, F. (2009). Species delimitation in downy mildews: the case of Hyaloperonospora in the light of nuclear ribosomal ITS and LSU sequences. *Mycological Research*, *113*(3), 308-325.

Göker, M., Voglmayr, H., Riethmüller, A., & Oberwinkler, F. (2007). How do obligate parasites evolve? A multi-gene phylogenetic analysis of downy mildews. *Fungal Genetics and Biology*, *44*(2), 105-122.

Göker, M., Voglmayr, H., Riethmüller, A., Weiß, M., & Oberwinkler, F. (2003). Taxonomic aspects of Peronosporaceae inferred from Bayesian molecular phylogenetics. *Canadian Journal of Botany*, *81*(7), 672-683. https://doi.org/10.1139/b03-066

Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., Chan, C. K. K., Severn-Ellis, A., McCombie, W. R., Parkin, I. A. P., Paterson, A. H., Pires, J. C., Sharpe, A. G., Tang, H., Teakle, G. R., Town, C. D., Batley, J., & Edwards, D. (2016). The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nature Communications*, *7*(1), 13390. https://doi.org/10.1038/ncomms13390

Guo, N., Wang, S., Gao, L., Liu, Y., Wang, X., Lai, E., Duan, M., Wang, G., Li, J., Yang, M., Zong, M., Han, S., Pei, Y., Borm, T., Sun, H., Miao, L., Liu, D., Yu, F., Zhang, W., . . . Liu, F. (2021). Genome sequencing sheds light on the contribution of structural variants to *Brassica oleracea* diversification. *BMC Biology*, *19*(1), 93. https://doi.org/10.1186/s12915-021-01031-2

Gust, A. A., & Felix, G. (2014). Receptor like proteins associate with SOBIR1-type of adaptors to form bimolecular receptor kinases. *Curr Opin Plant Biol*, *21*, 104-111. https://doi.org/10.1016/j.pbi.2014.07.007

Guttenberg, H. V., & Schmoller, H. (1958). [Attempted culture of Peronospora brassicae Gäum]. *Arch Mikrobiol*, *30*(3), 268-279. (Kulturversuche mit Peronospora brassicae Gäum.)

Hiura, M., & Kanegae, H. (1934). STUDIES ON THE Downy MildewS OF CRUCIFEROUS VEGETABLES IN JAPAN I. *札幌博物学会会報*, *13*(3), 125-133.

Holub, E. B. (2001). The arms race is ancient history in Arabidopsis, the wildflower. *Nature Reviews Genetics*, *2*(7), 516-527.

Hoser-Krauze, J., Lakowska-Ryk, E., & Antosik, J. (1984). Resistance of cauliflower and broccoli (*B. oleracea* L. botrytis L.) seedlings to downy mildew, Peronospora parasitica. *Cruciferae Newsletter*, *9*, 92-93.

Hoser-Krauze, J., Lakowska-Ryk, E., & Antosik, J. (1995). The inheritance of resistance of some *Brassica oleracea* L. cultivars and lines to downy mildew-Peronospora parasitica [Pers.] ex Fr. *Journal of Applied Genetics*, *36*(1).

Ingram, D. S. (1969). The Susceptibility of *Brassica* Callus to Infection by Peronospora Parasitica. *Microbiology*, *58*(3), 391-401. https://doi.org/https://doi.org/10.1099/00221287-58-3-391

Jang, P., & Safeeulla, K. (1990). Production and viability of Peronospora parasitica in radish. *Proc Indian Acad Sci Plant Sci 100:117–122*.

Jeger, M. J., & Spencer-Phillips, P. T. (2004). *Advances in downy mildew research*. Kluwer Academic Publishers.

Jensen, Hockenhull, & Munk. (1999). Seedling and adult plant resistance to downy mildew (Peronospora parasitica) in cauliflower (*Brassica oleracea* convar. botrytis var. botrytis). *Plant Pathology*, *48*(5), 604-612. https://doi.org/doi:10.1046/j.1365-3059.1999.00388.x

Jensen, B., Hockenhull, J., & Munk, L. (1999). Seedling and adult plant resistance to downy mildew (Peronospora parasitica) in cauliflower (*Brassica oleracea* convar. botrytis var. botrytis). *Plant Pathology*.

Jensen, B. D., Værbak, S., Munk, L., & Andersen, S. B. (1999). Characterization and inheritance of partial resistance to downy

mildew, Peronospora parasitica, in breeding material of broccoli, *Brassica oleracea* convar. botrytis var. italica. *Plant Breeding*, *118*(6), 549-554. https://doi.org/doi:10.1046/j.1439-0523.1999.00409.x

Jones, J. D., & Dangl, J. L. (2006). The plant immune system. *Nature*, *444*(7117), 323-329. https://doi.org/10.1038/nature05286

Jones, J. D. G., Vance, R. E., & Dangl, J. L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science*, *354*(6316), aaf6395-aaf6395. https://doi.org/10.1126/science.aaf6395

Kenneth, R. (1981). Downy mildews of graminaceous crops [Taxonomy, host range, control].

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357-360. https://doi.org/10.1038/nmeth.3317

Kim, S., Song, Y. H., Lee, J.-Y., Choi, S. R., Dhandapani, V., Jang, C. S., Lim, Y. P., & Han, T. (2011). Identification of the BrRHP1 locus that confers resistance to downy mildew in Chinese cabbage (*Brassica* rapa ssp. pekinensis) and development of linked molecular markers. *Theoretical and Applied Genetics*, *123*(7), 1183-1192.

Kluczewski, S., & Lucas, J. (1982). Development and physiology of infection by the downy mildew fungus Peronospora parasitica (Pers. ex Fr.) Fr. in susceptible and resistant *Brassica* species. *Plant Pathology*, *31*(4), 373-379.

Kluczewski, S. M., & Lucas, J. A. (1982). Development and physiology of infection by the downy mildew fungus Peronospora parasitica (Pers. ex Fr.) Fr. in susceptible and resistant *Brassica* species. *Plant Pathology*, *31*, 373-379.

Kluczewski, S. M., & Lucas, J. A. (1983). Host infection and oospore formation by Peronospora parasitica in agricultural and horticultural *Brassica* species. *Transactions of the British Mycological Society*, *81*(3), 591-596. https://doi.org/https://doi.org/10.1016/S0007-1536(83)80130-2

Koh, J. C. O., Barbulescu, D. M., Norton, S., Redden, B., Salisbury, P. A., Kaur, S., Cogan, N., & Slater, A. T. (2017). A multiplex PCR for rapid identification of *Brassica* species in the triangle of U. *Plant Methods*, *13*, 49. https://doi.org/10.1186/s13007-017-0200-8

Koike, S. T., Gladders, P., & Paulus, A. O. (2009). *Vegetable diseases : a color handbook*. Academic Press.

Kolte, S. (1985). *Diseases of annual edible oilseed crops, vol. 2 Rapeseed mustard and sesame diseases.* .

Kontaxis, D., Mayberry, K., & Rubatzky, V. (1979). Reaction of cauliflower cultivars to downy mildew in Imperial Valley. *California Agric 33:19*.

Krober, H. (1969). Infection of Kohlrabi by Peronospora parasitica and of tobacco by
Peronospora parasitica in relation to sporel concentration. *Phytopathology Z 66:180-187.*
Krober, H. (1970). Longevity of spore of several Peronospora species under different environmental conditions. . *Phytopathol Z 69:64–70*.
Krober, H. (1981). Germination and longevity of spore of some Peronosporaceae kept under different conditions. . *Zeit Pflanzenk Pflanzenschutz 88:510–517*.
LeBeau, F. (1945). Systemic invasion of cabbage seedlings by the downy mildew fungus. *Aust J Agric Res 71:453–463*.
Lebeda, A., & Syrovátko, P. (1988). Specificity of Bremia lactucae isolates from Lactuca sativa and some Asteraceae plants. *Acta Phytopathologica et Entomologica Hungarica*, *23*(1), 39-48.
Lee, J., Izzah, N. K., Choi, B.-S., Joh, H. J., Lee, S.-C., Perumal, S., Seo, J., Ahn, K., Jo, E. J., & Choi, G. J. (2016). Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA research*, *23*(1), 29-41.
Lee, J. M., & Sonnhammer, E. L. (2003). Genomic gene clustering analysis of pathways in eukaryotes. *Genome Res*, *13*(5), 875-882.
Lee, J. S., Lee, H. B., Shin, H. D., & Choi, Y. J. (2017). Diversity, Phylogeny, and Host-Specialization of Hyaloperonospora Species in Korea. *Mycobiology*, *45*(3), 139-149. https://doi.org/10.5941/myco.2017.45.3.139
Lee, R. R. Q., & Chae, E. (2020). Variation Patterns of NLR Clusters in Arabidopsis thaliana Genomes. *Plant Communications*, *1*(4), 100089. https://doi.org/https://doi.org/10.1016/j.xplc.2020.100089
Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754-1760. https://doi.org/10.1093/bioinformatics/btp324
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079. https://doi.org/10.1093/bioinformatics/btp352
Li, H., Yu, S. C., Zhang, F. L., Yu, Y. J., Zhao, X. Y., Zhang, D. S., & Zhao, X. (2011). [Development of molecular markers linked to the resistant QTL for downy mildew in *Brassica* rapa L. ssp. pekinensis]. *Yi Chuan*, *33*(11), 1271-1278. https://doi.org/10.3724/sp.j.1005.2011.01271
Li, P., Su, T., Zhao, X., Wang, W., Zhang, D., Yu, Y., Bayer, P. E., Edwards, D., Yu, S., & Zhang, F. (2021). Assembly of the non-heading pak choi

genome and comparison with the genomes of heading Chinese cabbage and the oilseed yellow sarson. *Plant Biotechnol J*, *19*(5), 966-976.

Li, Y., Liu, G.-F., Ma, L.-M., Liu, T.-K., Zhang, C.-W., Xiao, D., Zheng, H.-K., Chen, F., & Hou, X.-L. (2020). A chromosome-level reference genome of non-heading Chinese cabbage [*Brassica* campestris (syn. *Brassica* rapa) ssp. chinensis]. *Hortic Res*, *7*.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., Zhao, M., Ma, J., Yu, J., & Huang, S. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, *5*(1), 1-11.

Liu, X.-p., Gao, B.-z., Han, F.-q., Fang, Z.-y., Yang, L.-m., Zhuang, M., Lv, H.-h., Liu, Y.-m., Li, Z.-s., & Cai, C.-c. (2017). Genetics and fine mapping of a purple leaf gene, BoPr, in ornamental kale (*Brassica oleracea* L. var. acephala). *BMC Genomics*, *18*(1), 1-9.

Lucas, J., Crute, I., SHERRIFF, C., & GORDON, P. L. (1988). The identification of a gene for race-specific resistance to Peronospora parasitica (downy mildew) in *Brassica* napus var. oleifera (oilseed rape). *Plant Pathology*, *37*(4), 538-545.

Lucas, J. A., Hayter, J. B. R., & Crute, I. R. (1994). The downy mildews: Host specificity and pathogenesis. *Pathogenicity and Host Specificity in Plant Diseases*, *II-Eukaryotes*, pp 217-238. (Pergamon Press, London)

Lukens, L. N., Quijada, P. A., Udall, J., Pires, J. C., Schranz, M. E., & Osborn, T. C. (2004). Genome redundancy and plasticity within ancient and recent *Brassica* crop species. *Biological Journal of the Linnean Society*, *82*(4), 665-674.

Lv, H., Wang, Y., Han, F., Ji, J., Fang, Z., Zhuang, M., Li, Z., Zhang, Y., & Yang, L. (2020). A high-quality reference genome for cabbage obtained with SMRT reveals novel genomic features and evolutionary characteristics. *Sci Rep*, *10*(1), 12394. https://doi.org/10.1038/s41598-020-69389-x

Macho, A. P., & Zipfel, C. (2014). Plant PRRs and the activation of innate immune signaling. *Mol Cell*, *54*(2), 263-272. https://doi.org/10.1016/j.molcel.2014.03.028

Mahajan, V., Gill, H., & Singh, R. (1991). Screening of cauliflower germplasm lines against downy mildew (Peronospora parasitica). *Cruciferae Newsl*, *14*(15), 148-149.

Mahajan, V., Gill, H. S., & More, T. A. (1995). Inheritance of downy mildew resistance in Indian cauliflower (group III) [journal article]. *Euphytica*, *86*(1), 1-3. https://doi.org/10.1007/bf00035932

Martinez, S. A., Shorinola, O., Conselman, S., See, D., Skinner, D. Z., Uauy, C., & Steber, C. M. (2020). Exome sequencing of bulked segregants identified a novel TaMKK3-A allele linked to the wheat ERA8 ABA-hypersensitive germination phenotype. *Theoretical and Applied Genetics*, *133*(3), 719-736. https://doi.org/10.1007/s00122-019-03503-0

Masheva, S., Antonova, G., & Bahariev, D. (1996). Pathogenicity of two isolates Peronospora parasitica with different district origin. *Cruciferae Newsletter*(18), 118-119.

McDonald, B. A., & Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol*, *40*, 349-379. https://doi.org/10.1146/annurev.phyto.40.120501.101443

McDowell, J. M., Dhandaydham, M., Long, T. A., Aarts, M. G., Goff, S., Holub, E. B., & Dangl, J. L. (1998). Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell*, *10*(11), 1861-1874.

McMeekin, D. (1960). The role of the oospores of Peronospora parasitica from cabbage and radish.

. *Phytopathology, 50:93–97.*

McMeekin, D. (1981). The Response of Peronospora parasitica to a Liver Medium. *Mycologia*, *73*(3), 551-554. https://doi.org/10.2307/3759608

Mehta, N., & Saharan, G. (1994). Morphological and pathological variations in Peronospora parasitica infecting *Brassica* species. *Indian Phytopathology*, *47*(2), 153-158.

Ménard, R., Larue, J.-P., Silué, D., & Thouvenot, D. (1999). Glucosinolates in cauliflower as biochemical markers for resistance against downy mildew. *Phytochemistry*, *52*(1), 29-35. https://doi.org/https://doi.org/10.1016/S0031-9422(99)00165-X

Meyers, B. C., Kozik, A., Griego, A., Kuang, H., & Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell*, *15*(4), 809-834. https://doi.org/10.1105/tpc.009308

Michelmore, R. W., & Ingram, D. S. (1980). Heterothallism in Bremia lactucae. *Transactions of the British Mycological Society*, *75*(1), 47-56. https://doi.org/https://doi.org/10.1016/S0007-1536(80)80193-8

Michelmore, R. W., Paran, I., & Kesseli, R. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the national academy of sciences*, *88*(21), 9828-9832.

Mohammed, A. E., You, M. P., Banga, S. S., & Barbetti, M. J. (2019). Resistances to downy mildew (Hyaloperonospora brassicae) in

diverse Brassicaceae offer new disease management opportunities for oilseed and vegetable crucifer industries. *European Journal of Plant Pathology*, *153*(3), 915-929.

Monot, C., Penguilly, D., & Silue, D. (2010). First confirmed report of downy mildew caused by Hyaloperonospora parasitica on broccoli, cauliflower and Romanesco-type cauliflower heads in France. *Plant Pathology*, *59*(6), 1165-1165. https://eurekamag.com/research/066/237/066237313.php

Monteiro, A. A., Coelho, P. S., Bahcevandziev, K., & Valério, L. (2005). Inheritance of downy mildew resistance at cotyledon and adult-plant stages in 'Couve Algarvia' (*Brassica oleracea* var. tronchuda) [journal article]. *Euphytica*, *141*(1), 85-92. https://doi.org/10.1007/s10681-005-5696-8

Monteiro, A. A., & Williams, P. H. (1989). The exploration of genetic resources of Portuguese cabbage and kale for resistance to several *Brassica* diseases. *Euphytica*, *41*(3), 215-225. https://doi.org/10.1007/BF00021588

Morris, M., & Knox-Davies, P. (1980). Raphanis raphanistrum as a weed host of pathogens of cultivated Cruciferae in the western Cape Province of South Africa. *Phytophylactica*, *12*(2), 53-56.

Moss, N. A., Crute, I. R., Lucas, J. A., & Gordon, P. L. (1988). Requirements for analysis of host-species specificity in Peronospora parasitica (downy mildew). *Cruciferae Newsletter*, *13*, 114-116.

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

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant Journal*, *60*(2), 218-226.

Nashaat, N., & Awasthi, R. P. (1995). Evidence for Differential Resistance to Peronospora parasitica (downy mildew) in Accessions of *Brassica* juncea (mustard) at the Cotyledon Stage. *Journal of Phytopathology*, *143*, 157-159.

Nashaat, N., Heran, A., Awasthi, R., & Kolte, S. (2004). Differential response and genes for resistance to Peronospora parasitica (downy mildew) in *Brassica* juncea (mustard). *Plant Breeding*, *123*(6), 512-515.

Nashaat, N., Heran, A., Mitchell, S., & Awasthi, R. (1997). New genes for resistance to downy mildew (Peronospora parasitica) in oilseed rape (*Brassica* napus ssp. oleifera). *Plant Pathology*, *46*(6), 964-968.

Nashaat, N. I., & Rawlinson, C. J. (1994). The response of oilseed rape (*Brassica* napus ssp. oleifera) accessions with different glucosinolate and erucic acid contents to four isolates of Peronospora parasitica (downy mildew) and the identification of new sources of resistance. *Plant Pathology*, *43*(2), 278-285. https://doi.org/https://doi.org/10.1111/j.1365-3059.1994.tb02686.x

Natti, J. (1958). Resistance of broccoli and other crucifers to downy mildew. *Plant Disease Reporter 42, 656-662*.

Natti, J., & Atkin, J. (1960). Inheritance of downy mildew resistance in broccoli. . *Phytopathology 50: 241.*

Natti, J. J., Dickson, M. H., & Atkin, J. D. (1967). Resistance of *Brassica* varieties to downy mildew. . *Phytopathology 57: 144-147*.

Neik, T. X., Barbetti, M. J., & Batley, J. (2017). Current Status and Challenges in Identifying Disease Resistance Genes in *Brassica* napus. *Frontiers in Plant Science*, *8*, 1788. https://doi.org/10.3389/fpls.2017.01788

O'neill, C. M., & Bancroft, I. (2000). Comparative physical mapping of segments of the genome of *Brassica oleracea* var. alboglabra that are homoeologous to sequenced regions of chromosomes 4 and 5 of Arabidopsis thaliana. *The Plant Journal*, *23*(2), 233-243.

OECD. (2012). Consensus document on the biology of the *Brassica* crops (*Brassica* spp.). *Series on Harmonisation of Regulatory oversight of Biotechnology, No 54, OECD*, 142.

Pandey, K., Pandey, P., Singh, B., Kalloo, G., & Kapoor, K. (2001). *Sources of resistance to downy mildew (Peronospora parasitica) disesae in Asiatic group of cauliflower*.

Paritosh, K., Pradhan, A. K., & Pental, D. (2020). A highly contiguous genome assembly of *Brassica* nigra (BB) and revised nomenclature for the pseudochromosomes. *BMC Genomics*, *21*(1), 887. https://doi.org/10.1186/s12864-020-07271-w

Paritosh, K., Yadava, S. K., Singh, P., Bhayana, L., Mukhopadhyay, A., Gupta, V., Bisht, N. C., Zhang, J., Kudrna, D. A., Copetti, D., Wing, R. A., Reddy Lachagari, V. B., Pradhan, A. K., & Pental, D. (2021). A chromosome-scale assembly of allotetraploid *Brassica* juncea (AABB) elucidates comparative architecture of the A and B genomes. *Plant Biotechnol J*, *19*(3), 602-614. https://doi.org/https://doi.org/10.1111/pbi.13492

Parker, J. E., Coleman, M. J., Szabò, V., Frost, L. N., Schmidt, R., van der Biezen, E. A., Moores, T., Dean, C., Daniels, M. J., & Jones, J. (1997). The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6. *Plant Cell*, *9*(6), 879-894.

Parkin, I. A., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., & Lydiate, D. J. (2005). Segmental structure of the *Brassica* napus genome based on comparative analysis with Arabidopsis thaliana. *Genetics*, *171*(2), 765-781. https://doi.org/10.1534/genetics.105.042093

Parkin, I. A. P., 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, J. C., . . . Sharpe, A. G. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biology*, *15*(6), R77. https://doi.org/10.1186/gb-2014-15-6-r77

Paul, V., Klodt-Bussmann, E., Dapprich, P., Capelli, C., & Tewari, J. (1998). Results on preservation, epidemiology, and aggressiveness of Peronospora parasitica and results with regard to the disease resistance of the pathogen on *Brassica* napus.

Persoon, C. (1796). Observations mycologicae sur descriptions tan novorum quam notibilium fungorum exhibitae. . *Part 1: 115p*.

Pinckard, J. A. (1942). The mechanism of spore dispersal in Peronospora tabacina and certain other downy mildew fungi. *Phytopathology 32: 505-511*.

The Plant List. Version 1. (2010).

Ramirez-Gonzalez, R. H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M., & Uauy, C. (2015). RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnol J*, *13*(5), 613-624. https://doi.org/10.1111/pbi.12281

Ramirez-Gonzalez, R. H., Uauy, C., & Caccamo, M. (2015). PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics*, *31*(12), 2038-2039. https://doi.org/10.1093/bioinformatics/btv069

Ramsay, L. D., Jennings, D. E., Kearsey, M. J., Marshall, D. F., Bohuon, E. J. R., Arthur, A. E., & Lydiate, D. J. (1996). The construction of a substitution library of recombinant backcross lines in *Brassica oleracea* for the precision mapping of quantitative trait loci. *Genome*, *39*(3), 558-567. https://doi.org/10.1139/g96-071 %M 18469917

Rooster, L., Spiessens, K., Reycke, L., & Janssens, M. (1999). Late autumn cultivation of cauliflowers. Difference in resistance to downy mildew. . *Proeftuinnieuws 9:12–13*.

Runge, F., Choi, Y.-J., & Thines, M. (2011). Phylogenetic investigations in the genus Pseudoperonospora reveal overlooked species and cryptic diversity in the P. cubensis species cluster. *European Journal of Plant Pathology*, *129*(2), 135-146.

Sadowski, J., & Kole, C. (2016). Genetics, Genomics and Breeding of Vegetable Brassicas. https://nls.ldls.org.uk/welcome.html?ark:/81055/vdc_100045257829.0x000001

Saha, P., Ghoshal, C., Ray, S., Saha, N. D., Srivastava, M., Kalia, P., & Tomar, B. (2020). Genetic analysis of downy mildew resistance and identification of molecular markers linked to resistance gene Ppa207 on chromosome 2 in cauliflower. *Euphytica*, *216*(11), 1-13.

Satou, M., & Fukumoto, F. (1993). Preservation of Spore of Broccoli Downy Mildew Fungus with Cryogenic Protectants by Freezing at-80°C. *Japanese Journal of Phytopathology*, *59*(5), 492-499. https://doi.org/10.3186/jjphytopath.59.492

Séguin-Swartz, C. C. a. G. (2005). BREEDING FOR DISEASE RESISTANCE IN OILSEED CROPS IN INDIA. *Annual Review of Plant Pathology (Vol. 3)*, *3*, 101-142.

Semb, L. (1969). Cabbage downy mildew. *Jord Avling*, *12*, 32-35.

Sharma, B. R., Dhiman, J. S., Thakur, J. C., Singh, A., & Bajaj, K. L. (1991). Multiple disease resistance in cauliflower. *Advances in Horticultural Science*, *5*(1), 30-34. http://www.jstor.org/stable/42881574

Sharma, S., Kapoor, K., & Gill, H. (1995). Screening against sclerotinia rot (Sclerotinia sclerotiorum), downy mildew (Peronospora parasitica) and black rot (Xanthomonas campestris) in cauliflower (*Brassica oleracea* var botrytis subvar cauliflora). *Indian Journal of Agricultural Sciences (India)*.

Sharma, S. K., McLean, K., Colgan, R. J., Rees, D., Young, S., Sønderkær, M., Terry, L. A., Turnbull, C., Taylor, M. A., & Bryan, G. J. (2021). Combining conventional QTL analysis and whole-exome capture-based bulk-segregant analysis provides new genetic insights into tuber sprout elongation and dormancy release in a diploid potato population. *Heredity*, *127*(3), 253-265. https://doi.org/10.1038/s41437-021-00459-0

Shaw, R. K., Shen, Y., Zhao, Z., Sheng, X., Wang, J., Yu, H., & Gu, H. (2021). Molecular Breeding Strategy and Challenges Towards Improvement of Downy Mildew Resistance in Cauliflower (*Brassica oleracea* var. botrytis L.) [Review]. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.667757

Sherf, A. F., & MacNab, A. A. (1986). *Vegetable diseases and their control*. John Wiley & Sons.

Sheriff, C., & Lucas, J. (1987). Variation in host specificity in the brassica population of Peronospora parasitica *Blackwell Scientific Publication, Oxford., In Day, P.R. and Jellis, GJ. (eds.)*, pp 333-335.

Sheriff, C., & Lucas, J. A. (1989a). Cytogenetic study of heterothallic and homothallic isolates of Peronospora parasitica. *Mycological Research*, *92(3): 302-310*.

Sheriff, C., & Lucas, J. A. (1989b). Heterothallism and homothallism in Peronosporaparasitica. *Mycological Research*, *92(3): 311-316*.

Sherriff, C., & Lucas, J. A. (1990). The host range of isolates of downy mildew, Peronospora parasitica, from *Brassica* crop species. *Plant Pathology*, *39*, 77-91.

Shiraishi, M., Sakamoto, K., Asada, Y., Nagatani, T., & Hidaka, H. (1975). A Scanning Electron Microscopic Observation on the Surface of Japanese Radish Leaves Infected by Peronospora parasitica (Fr.) Fr. *Japanese Journal of Phytopathology*, *41*(1), 24-32.

Shiu, S. H. a., & Bleecker, A. (2001). Shiu, S. H. & Bleecker, A. B. Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proc. Natl Acad. Sci. USA 98, 10763-10768. *Proc Natl Acad Sci U S A*, *98*, 10763-10768. https://doi.org/10.1073/pnas.181141598

Silué, D., Laviec, C., & Tirilly, Y. (1995). New sources of resistance in cauliflower (*Brassica oleracea* var. botrytis) to crucifer downy mildew caused by Peronospora parasitica. *Journal of Phytopathology*, *143*(11-12), 659-661.

Silue, D., Nashaat, N. I., & Tirilly, Y. (1996). Differential responses of *Brassica oleracea* and B. rapa accessions to seven isolates of Peronospora parasitica at the cotyledon stage. *Plant Disease*, *80 (2)*, pp. 142-144.

Silue, D., Nashaat, N. I., & Tirilly, Y. (1996). *Differential Responses of Brassica oleracea and B. rapa Accessions to Seven Isolates of Peronospora parasitica at the Cotyledon Stage* (Vol. 80). https://doi.org/10.1094/PD-80-0142

Singh, R., Trivedi, B., Gill, H., & Sen, B. (1987). Breeding for resistance for black rot, downy mildew and curd blight in Indian cauliflower. *Cruciferae Newsletter*, *12*, 96-97.

Singh, S., Sharma, S., Kalia, P., Deshmukh, R., Katara, J., Sharma, P., & Sharma, T. (2015). Identification of putative DNA markers for disease resistance breeding in Indian cauliflower (*Brassica oleracea* var. botrytis L.).

Singh, S., Sharma, S., Kalia, P., Deshmukh, R., Kumar, V., Sharma, P., & Sharma, T. (2012). Molecular mapping of the downy mildew resistance gene Ppa3 in cauliflower (*Brassica oleracea* var. botrytis L.). *The Journal of Horticultural Science and Biotechnology*, *87*(2), 137-143.

Singh, S., Sharma, S., Kalia, P., Sharma, P., Kumar, V., Kumar, R., Meena, B., Kumar, S., & Sharma, T. (2013). Screening of cauliflower (*Brassica oleracea* L. var. botrytis L.) germplasm for resistance to downy

mildew [Hyaloperonospora parasitica Constant (Pers.: Fr) Fr.] and designing appropriate multiple resistance breeding strategies. *The Journal of Horticultural Science and Biotechnology*, *88*(1), 103-109.

Singh, S., Singh, D., & Bains, B. (1980). In vivo cellulase and pectinase production by Albugo candida and Peronospora parasitica. . *Indian Phytopath 33:370–371*.

Steuernagel, B., Witek, K., Krattinger, S. G., Ramirez-Gonzalez, R. H., Schoonbeek, H.-j., Yu, G., Baggs, E., Witek, A. I., Yadav, I., Krasileva, K. V., Jones, J. D. G., Uauy, C., Keller, B., Ridout, C. J., & Wulff, B. B. H. (2018). Physical and transcriptional organisation of the bread wheat intracellular immune receptor repertoire. *bioRxiv*, 339424. https://doi.org/10.1101/339424

Steuernagel, B., Witek, K., Krattinger, S. G., Ramirez-Gonzalez, R. H., Schoonbeek, H.-j., Yu, G., Baggs, E., Witek, A. I., Yadav, I., Krasileva, K. V., Jones, J. D. G., Uauy, C., Keller, B., Ridout, C. J., & Wulff, B. B. H. (2020). The NLR-Annotator Tool Enables Annotation of the Intracellular Immune Receptor Repertoire. *Plant Physiol*, *183*(2), 468-482. https://doi.org/10.1104/pp.19.01273

Sun, D., Wang, C., Zhang, X., Zhang, W., Jiang, H., Yao, X., Liu, L., Wen, Z., Niu, G., & Shan, X. (2019). Draft genome sequence of cauliflower (*Brassica oleracea* L. var. botrytis) provides new insights into the C genome in *Brassica* species. *Hortic Res*, *6*(1), 82. https://doi.org/10.1038/s41438-019-0164-0

Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A., Mitsuoka, C., Utsushi, H., Natsume, S., Kanzaki, H., Matsumura, H., Saitoh, H., Yoshida, K., Cano, L. M., Kamoun, S., & Terauchi, R. (2013). MutMap-Gap: whole-genome resequencing of mutant F_2 progeny bulk combined with *de novo* assembly of gap regions identifies the rice blast resistance gene Pii. *New Phytol*, *200*(1), 276-283. https://doi.org/10.1111/nph.12369

Takken, F. L., & Goverse, A. (2012). How to build a pathogen detector: structural basis of NB-LRR function. *Curr Opin Plant Biol*, *15*(4), 375-384. https://doi.org/10.1016/j.pbi.2012.05.001

Tham, F. Y., Lucas, J. A., & Wilson, Z. A. (1994). DNA fingerprinting of Peronospora parasitica, a biotrophic fungal pathogen of crucifers. *Theor Appl Genet*, *88*(3-4), 490-496. https://doi.org/10.1007/bf00223666

Thines, M., & Choi, Y.-J. (2016). Evolution, diversity, and taxonomy of the Peronosporaceae, with focus on the genus Peronospora. *Phytopathology*, *106*(1), 6-18.

Thines, M., Telle, S., Ploch, S., & Runge, F. (2009). Identity of the downy mildew pathogens of basil, coleus, and sage with implications for quarantine measures. *Mycological Research*, *113*(5), 532-540.

Thomas, C. E., & Jourdain, E. (1990). Evaluation of broccoli and cauliflower germplasm for resistance to race 2 of Peronospora parasitica. *HortScience*, *25*(11), 1429-1431.

Thomas, C. E., & Jourdain, E. (1992). Resistance to race 2 of Peronospora parasitica in US plant introductions of *Brassica oleracea* var. capitata. *HortScience*, *27*(10), 1120-1122.

Thomas, C. E., & Jourdain, E. L. (1990). Evaluation of Broccoli and Cauliflower Germplasm for Resistance to Race 2 of Peronospora parasitica. *HortScience HortSci*, *25*(11), 1429-1431. https://doi.org/10.21273/hortsci.25.11.1429

Thung, T. (1926). Peronospora parasitica (Pers.) De by attacking cabbage heads. *Phytopathology*, *16*, 365-366.

Townsend, G. R., & Heuberger, J. W. (1943). Methods for estimating losses caused by diseases in fungicidal experiments. *Plant Dis Reporter*, *27*(17), 340-343. https://eurekamag.com/research/025/008/025008582.php

TransDecoder. https://github.com/TransDecoder/TransDecoder

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*, *28*(5), 511-515. https://doi.org/10.1038/nbt.1621

Trick, M., Adamski, N. M., Mugford, S. G., Jiang, C. C., Febrer, M., & Uauy, C. (2012). Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biol*, *12*, 14. https://doi.org/10.1186/1471-2229-12-14

Tyler, B. M. (2007). Phytophthora sojae: root rot pathogen of soybean and model oomycete. *Mol Plant Pathol*, *8*(1), 1-8.

Van Damme, M., Andel, A., Huibers, R. P., Panstruga, R., Weisbeek, P. J., & Van den Ackerveken, G. (2005). Identification of Arabidopsis loci required for susceptibility to the downy mildew pathogen Hyaloperonospora parasitica. *Molecular Plant-Microbe Interactions*, *18*(6), 583-592.

van Wersch, S., & Li, X. (2019). Stronger when together: clustering of plant NLR disease resistance genes. *Trends Plant Sci*, *24*(8), 688-699.

Verma, A., & Singh, Y. (2018). Inheritance of downy mildew resistance and its relationship with biochemical traits in cauliflower (*Brassica oleracea* L. var. botrytis). *Crop Protection*, *106*, 132-138.

Verma, P., Saharan, G., & Goyal, B. (1994). Peronospora parasitica (Pers. ex Fr.) Fr. (downy mildew) on crucifers: Introduction, bibliography and subject index. . *Saskatoon Res Stn Tech Bull. 1994–02: 51*.

VH, P., & E, K.-B. (1993). First results from simple methods of preservation of Peronospora parasitica. *IOBC/WPRS Bull 16:25–28*.

Vicente, J., Gunn, N., Bailey, L., Pink, D., & Holub, E. (2012). Genetics of resistance to downy mildew in *Brassica oleracea* and breeding towards durable disease control for UK vegetable production. *Plant Pathology*, *61*(3), 600-609.

Voglmayr, H., Choi, Y.-J., & Shin, H.-D. (2014). Multigene phylogeny, taxonomy and reclassification of Hyaloperonospora on Cardamine. *Mycological Progress*, *13*(1), 131-144.

Walley, P. G., Teakle, G. R., Moore, J. D., Allender, C. J., Pink, D. A. C., 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. The Korean Society for Plant Biotechnology*. https://doi.org/10.5010/jpb.2012.39.1.062

Wang, C. (1944). Physiological specialization in Peronospora parasitica and reaction of hosts. *Chinese Journal of Scientific Agriculture*, *1*, 249-257.

Wang, G., Fiers, M., Ellendorff, U., Wang, Z.-z., de Wit, P. J. G. M., Angenent, G. C., & Thomma, B. P. H. J. (2010). The Diverse Roles of Extracellular Leucine-rich Repeat-containing Receptor-like Proteins in Plants. *Critical Reviews in Plant Sciences*, *29*, 285 - 299.

Wang, M., Farnham, M. W., & Thomas, C. E. (2000). Phenotypic Variation for Downy Mildew Resistance among Inbred Broccoli. *HortScience HortSci*, *35*(5), 925-929. https://doi.org/10.21273/hortsci.35.5.925

Wang, W., Chen, L., Fengler, K., Bolar, J., Llaca, V., Wang, X., Clark, C. B., Fleury, T. J., Myrvold, J., Oneal, D., van Dyk, M. M., Hudson, A., Munkvold, J., Baumgarten, A., Thompson, J., Cai, G., Crasta, O., Aggarwal, R., & Ma, J. (2021). A giant NLR gene confers broad-spectrum resistance to Phytophthora sojae in soybean. *Nature Communications*, *12*(1), 6263. https://doi.org/10.1038/s41467-021-26554-8

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*(10), 1035-1039.

Wu, C. H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., & Kamoun, S. (2017). NLR network mediates immunity

to diverse plant pathogens. *Proc Natl Acad Sci U S A*, *114*(30), 8113-8118. https://doi.org/10.1073/pnas.1702041114

Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., Hu, Z., Chen, S., Pental, D., Ju, Y., Yao, P., Li, X., Xie, K., Zhang, J., Wang, J., Liu, F., Ma, W., Shopan, J., Zheng, H., . . . Zhang, M. (2016). The genome sequence of allopolyploid *Brassica* juncea and analysis of differential homoeolog gene expression influencing selection. *Nature genetics*, *48*(10), 1225-1232. https://doi.org/10.1038/ng.3657

Young-Joon, C., Seung-Beom, H., & Hyeon-Dong, S. (2003). Diversity of the Hyaloperonospora parasitica complex from core brassicaceous hosts based on ITS rDNA sequences. *Mycological Research*, *107*(11), 1314-1322.

Yu, S., Su, T., Zhi, S., Zhang, F., Wang, W., Zhang, D., Zhao, X., & Yu, Y. (2016). Construction of a sequence-based bin map and mapping of QTLs for downy mildew resistance at four developmental stages in Chinese cabbage (*Brassica* rapa L. ssp. pekinensis). *Molecular Breeding*, *36*(4), 1-12.

Yu, S., Zhang, F., Yu, R., Zou, Y., Qi, J., Zhao, X., Yu, Y., Zhang, D., & Li, L. (2009). Genetic mapping and localization of a major QTL for seedling resistance to downy mildew in Chinese cabbage (*Brassica* rapa ssp. pekinensis). *Molecular Breeding*, *23*(4), 573-590.

Yu, S., Zhang, F., Zhao, X., Yu, Y., & Zhang, D. (2011). Sequence-characterized amplified region and simple sequence repeat markers for identifying the major quantitative trait locus responsible for seedling resistance to downy mildew in Chinese cabbage (*Brassica* rapa ssp. pekinensis). *Plant Breeding*, *130*(5), 580-583.

Zeilmaker, T., Ludwig, N. R., Elberse, J., Seidl, M. F., Berke, L., Van Doorn, A., Schuurink, R. C., Snel, B., & Van den Ackerveken, G. (2015). Downy Mildew RESISTANT 6 and Downy MildewR 6-LIKE OXYGENASE 1 are partially redundant but distinct suppressors of immunity in Arabidopsis. *The Plant Journal*, *81*(2), 210-222.

Zhang, B., Li, P., Su, T., Li, P., Xin, X., Wang, W., Zhao, X., Yu, Y., Zhang, D., & Yu, S. (2018). BrRLP48, encoding a receptor-like protein, involved in downy mildew resistance in *Brassica* rapa. *Frontiers in Plant Science*, 1708.

Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., King, G. J., & Liu, K. (2015). Disruption of a CAROTENOID CLEAVAGE DIOXYGENASE 4 gene converts flower colour from white to yellow in *Brassica* species. *New Phytologist*, *206*(4), 1513-1526.

Zhang, L., Cai, X., Wu, J., Liu, M., Grob, S., Cheng, F., Liang, J., Cai, C., Liu, Z., & Liu, B. (2018). Improved *Brassica* rapa reference genome by

single-molecule sequencing and chromosome conformation capture technologies. *Hortic Res*, *5*.

Zhi, S., Su, T., Yu, S., Zhang, F., Yu, Y., Zhang, D., Zhao, X., Wang, W., Lu, G., & Zhu, Y. (2016). Genetic characteristics of A01-located resistant loci to downy mildew in Chinese cabbage by genome-wide association studies. *Plant Physiology Journal*, *52*(5), 693-702.