Characterization of rutabaga for genetic diversity and as a source of clubroot resistance

by

Zhiyu Yu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Plant Science

Department of Agricultural, Food and Nutritional Science
University of Alberta

Abstract

Rutabaga (*Brassica napus* ssp. *napobrassica* (L.) Hanelt) is widely grown as a vegetable crop and animal fodder, and is reported to be an excellent source of clubroot (*Plasmodiophora brassicae*) resistance genes. In this study, the genetic diversity and clubroot resistance of 124 rutabaga accessions from the Nordic countries (Norway, Sweden, Finland, Denmark, and Iceland) were assessed using a 15K Brassica single nucleotide polymorphism (SNP) array. Filtering was done to remove markers that did not amplify genomic DNA, and monomorphic and low coverage site markers. Allelic frequency statistics calculated with the retained 6861 SNP markers indicated that the rutabaga accessions from Norway, Sweden, Finland and Denmark were not genetically different from each other. In contrast, accessions from these countries were significantly different from the Icelandic accessions ($P < 0.05$). The rutabaga accessions were also evaluated in the greenhouse for their reaction to five single-spore isolates representing *P. brassicae* pathotypes 2F, 3H, 5I, 6M, and 8N and 12 field isolates representing pathotypes 2B, 3A, 3O, 5C, 5G, 5K, 5L, 5X (two isolates, L-G2 and L-G3), 8E, 8J and 8P. The rutabaga accessions exhibited differential reactions to the 17 isolates with 0.8-46.4% resistant (R), 4.0-20.0% moderately resistant (MR) and 32.8-93.6% susceptible (S). Nine accessions with broad-spectrum (R + MR) resistance also were identified. The rutabaga accessions were genotyped with 63 PCR-based primers linked to previously identified clubroot resistance genes. Genome-wide association studies (GWAS) using the genotypic (SNP + PCR-based markers) and phenotypic data identified 45 SNPs (36 on the A-genome and 9 on the C-genome or scaffolds) and 4 PCR-based markers that were associated strongly with resistance to isolates representing 13 pathotypes (2F, 3H, 5I, 6M, 8N, 2B, 3A, 3O, 5C, 5G, 5K, 5L and 8P). The SNPs identified in this study will be important for marker-assisted breeding (MAS) of clubroot resistant cruciferous crops.

Preface

This thesis is an original work by me, Zhiyu Yu, who conducted all of the experiments and wrote the first drafts of all chapters. My Supervisor Dr. Stephen Strelkov reviewed and revised all chapters of this thesis. My Supervisory Committee member Dr. Rudolph Fredua-Agyeman helped with the data analysis and contributed to the revisions of Chapters 2 and 3. Dr. Robert Conner (Agriculture and Agri-Food Canada) also reviewed Chapter 2 and provided suggestions for improvements. Once the draft thesis was complete, Dr. Strelkov and my co-supervisor Dr. Sheau-Fang Hwang approved the dissertation to go to defense.

The *Brassica napus* ssp. *napobrassica* accessions used in this thesis were obtained from Nordic Genetic Resource Center (Nordgen, Sweden) in 2015. The *Plasmodiophora brassicae* isolates in Chapter 3 were provided by Drs. Strelkov and Hwang. Dr. Fredua-Agyeman provided guidance and suggestions on experimental design and statistical analysis for Chapters 2 and 3. Mr. Chris Sehn and Ms. Yixiao Wang (University of Alberta) provided assistance with plant and DNA sampling and PCR genotyping in Chapters 2 and 3. Staff and students from the University of Alberta Plant Pathology Lab assisted with the greenhouse inoculation experiments described in Chapter 3.

Financial support was provided by grants to Dr. Fredua-Agyeman from the Alberta Crop Industry Development Fund (ACIDF) and to Drs. Strelkov and Hwang from the Canadian Agricultural Partnership (Agriculture and Agri-Food Canada and the Canola Council of Canada). I also acknowledge the in-kind support provided by Alberta Agriculture and Forestry and the University of Alberta, including access to greenhouse facilities and related laboratory infrastructure.

Acknowledgements

I would like to acknowledge my supervisor Dr. Stephen Strelkov, co-supervisor Dr. Sheau-Fang Hwang and my committee member Dr. Rudolph Fredua-Agyeman for their kind guidance and support throughout my M.Sc. program. I would also like to thank Dr. Robert Conner for critically reviewing Chapter 2. I gratefully appreciate the technical help I received from Dr. Qixing Zhou and Mr. Heting Fu (Alberta Agriculture and Forestry), along with the help from the summer students at the University of Alberta Plant Pathology Lab. I would like to thank Dr. Boyd Mori for serving as the arm's-length examiner on my defense and providing suggestions for this thesis.

My sincere, heartfelt gratitude goes to Drs. Strelkov and Hwang for inspiring me in plant pathology and leading me to the door of scientific research. I also gratefully thank Dr. Fredua-Agyeman for his kind guidance and help on my projects.

I address my special thanks to my colleague and dearest friend Ms. Yixiao Wang for her support and help during this entire process, who walked me through the darkest moments.

Lastly, I would like to thank my parents Mr. Jiankun Yu and Ms. Qiufang Lin for their encouragement and support over the years and my late grandmother Lanhua Yan who always expected me to step forward and be strong.

I would like to acknowledge the financial support from the Alberta Crop Industry Development Fund (ACIDF) and the Canadian Agricultural Partnership (Agriculture and Agri-Food Canada and the Canola Council of Canada). I also thank Alberta Agriculture and Forestry and University of Alberta for their generous in-kind support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Literature Review

1.1 Clubroot

1.1.1 Introduction to clubroot

Clubroot, caused by *Plasmodiophora brassicae* Woronin, is a soilborne disease of the Brassicaceae or crucifer family. Disease development is characterized by the formation of galls on the roots of infected plants. These galls interfere with normal root function, most notably water and nutrient uptake. When galling is severe, significant yield and quality losses can occur (Wallenhammar 1996; Dixon 2009). *Plasmodiophora brassicae* is an obligate parasite, growing and reproducing only in its living hosts. In the absence of a host, *P. brassicae* persists in the soil as long-lived resting spores, which can survive for more than 15 years (Wallenhammar 1996). This persistence makes clubroot management difficult.

1.1.2 Importance of clubroot

Clubroot is a major constraint to the production of cruciferous crops worldwide, with global yield losses estimated at 10 - 15% (Crête 1981, Dixon 2009). In the early 1980s, Crête (1981) examined the frequency of clubroot infestation in 18 countries or regions and found an average infection of 11%, with *B. oleracea* the most widely affected species. In East Asia, clubroot of Chinese cabbage (*Brassica. rapa* ssp. *pekinensis*) and oilseed rape (*B. napus* L.) is a devastating problem (Dixon 2009; Piao et al. 2009). For example, 3-4 million ha of cropland are infested with *P. brassicae* in China, representing one third of the total *Brassica* vegetable growing area nationwide and resulting in yield losses of 20 - 30% (Wang *et al.* 2008; Shen *et al.* 2009). Wang *et al.* (2008) estimated clubroot-induced yield losses of 10.2% in oilseed rape in China, with an average field infestation of 15%. The disease is also a major issue in Europe, both on Brassica vegetables and oilseeds (Crête 1981, Dixon 2009). In Sweden, yield losses of up to 50% were reported on oilseed rape (Wallenhammar et al. 1999).

Clubroot also occurs in Australia (Donald et al. 2014) and Latin America (Botero et al. 2019), mainly on cruciferous vegetables, as well as in the United States (Chittem et al. 2014).

Historically in Canada, clubroot was a problem for the growers of cruciferous vegetables in British Columbia, Ontario, Quebec and the Maritime Provinces (Howard *et al*. 2010). The disease was not reported on the highly valuable western Canadian canola (*Brassica napus* L.) crop until 2003, when a dozen cases were found in central Alberta (Tewari et al. 2005). Since then, clubroot has become one of the most important canola disease issues, and there were 3,044 confirmed field infestations in Alberta by 2018 (Strelkov et al. 2019). The disease appears to be spreading into Saskatchewan and Manitoba, with 46 and 33 field infestations reported in those provinces, respectively, in the past year (Ziesman et al. 2019; Froese et al. 2019). Yield losses of up to 30-100% have been observed in heavily diseased canola crops in Alberta (Tewari et al. 2005; Strelkov and Hwang 2014).

1.1.3 History of clubroot

Clubroot is well-known throughout the history of *Brassica* cultivation. A 4th century description of 'spongy roots' on vegetables by Pallidus is believed to be the earliest record of this disease (Watson and Baker 1969; Howard et al. 2010). Multiple records or illustrations of clubroot symptoms also exist from 16th and 17th century Europe (Dixon 2009; Howard et al. 2010). Indeed, clubroot was widely reported across Europe in the 17th and 18th centuries, before Michael Woronin identified *P. brassicae* as the causal agent of this disease in 1878 (Dixon 2009). The occurrence of clubroot increased in the 19th century during the industrial revolution, as cruciferous crops were grown widely (reviewed by Dixon 2009 and Howard et al. 2010). Clubroot first appeared on the coasts of the North American and Australasian colonies later in the 19th century. It was known to occur in Nova Scotia, Canada, by the 1910s (Estey 1994), and on the opposite side of the country (Victoria, British Columbia) by 1920 (Rankin and Fraser 1920). The earliest reports of clubroot in Japan are from the 1930s on

Chinese cabbage (*B. rapa* ssp. *pekinensis*) (reviewed by Wang et al. 2011). Clubroot was first reported in Taiwan province, off the coast of southeast China, in the 1940s during the Japanese occupation. The disease then appears to have spread across the Taiwan Strait to Fujian Province (Z.H. Wang, Fujian Agriculture and Forestry University, personal communication). A journal article regarding the occurrence of clubroot on the Chinese mainland (Yuan 1995) states that the disease had spread to Sichuan Province in southwest China by the 1970s. Currently, clubroot occurs in more than 60 countries and on all continents except Antarctica (Dixon 2009).

1.1.4 Taxonomy

Historically, *P. brassicae* was placed in the genus *Plasmodiophora* in the order *Plasmodiophorales*, which included seven other genera (Karling 1968). In addition to *P. brassicae*, the genus *Plasmodiophora* includes four other species: *P. diplantherae, P. halophilae, P. fici-repentis and P. bicaudate*. More recently, Neuhauser et al. (2010) classified *P. brassicae* along with the other plasmodiophorids in the order *Plasmodiophorida* in the *Phytomyxea*, in a new hierarchical system based on the sorting of protein sequence data. Currently, *P. brassicae* is grouped in the *Rhizaria* supergroup, phylum *Cercozoa*, class *Phytomyxea*, order *Plasmodiophorida* (Neuhauser et al. 2011; Burki and Keeling 2014).

1.1.5 Life cycle

The life cycle of *P. brassicae* (Fig.1) has been studied extensively over the past few decades (reviewed by Kageyama and Asano 2009). It consists of four main stages: existence in the soil, infection of the host root hairs (primary infection), infection of the root cortex (secondary infection), and dispersion of resting spores. As a soil-borne pathogen, *P. brassicae* survives as resting spores, which originate from old, decaying root galls (Gibbs 1931; Wallenhammar 1996). The subspherical or spherical resting spores serve as the

primary inoculum, and germinate under good soil moisture conditions to release pyriform or spindle-shaped biflagellate zoospores (Yano et al. 1991; Kageyama and Asano 2009). Germination appears to be enhanced by the presence of root exudates (Friberg et al. 2005, 2006). The zoospores encyst on and penetrate the root hairs, forming primary plasmodia (Katsura et al. 1970; Ingram and Tommerup 1972; Kageyama and Asano, 2009).

The primary plasmodia divide to form zoosporangia, which extend as clusters in the root hairs (Suzuki et al. 1992). Secondary zoospores emerge from these zoosporangia and are released back into the soil; subsequently, they re-infect the host, initiating the cortical infection stage (Dobson and Gabrielson 1983; Naiki et al. 1984). Once in the cortical tissue, the pathogen develops into intracellular secondary plasmodia (Asano and Kageyama 2006; Kageyama and Asano 2009). At this stage, hormonal regulation of the host tissues is affected, resulting in hyperplasia and hypertrophy, and the development of the typical root gall symptoms (Siemens et al. 2009; Ludwig-Müller et al. 2009). The plasmodia eventually cleave into huge numbers of new resting spores. As the host plant matures and senesces, the root galls begin to decompose, releasing the resting spores back into the soil (Ingram and Tommerup 1972; Ikegayami et al. 1982; Schwelm et al. 2015). It has been estimated that as many as 16 billion resting spores can be produced from one large gall (Hwang 2018).

1.1.6 Epidemiology and host range

The movement of infested soil represents the most common way in which clubroot spreads from field to field (Cao et al. 2009; Dixon 2009). Cao et al. (2009) noted that the incidence of clubroot in Alberta canola crops is highest at the field entrances, suggesting the introduction of *P. brassicae* on infested soil carried on farm and other machinery. There is also evidence that the pathogen can spread by wind and/or water erosion (Dixon 2009; Rennie et al. 2015). The presence of *P. brassicae* resting spores also was confirmed on potato tubers and the

seeds of various crops, although infestation levels were generally low (Rennie *et al.* 2011). The manure of animals fed on diseased plants or root tissue may also contribute to the spread of clubroot (Creelman 1965).

All species in the family Brassicaceae are potential hosts of *P. brassicae* (Dixon 2009), and the susceptibility of 89 species in eight genera including *Brassica*, *Raphanus* and *Arabidopsis* was confirmed by Karling (1968). Many cruciferous weeds, including wild mustard (*Sinapis arvensis* L.), shepherd's purse (*Capsella bursa-pastoris* (L.) Medik.) and stinkweed (*Thlaspi arvense* L.), also are susceptible to clubroot; this has important management implications, since *P. brassicae* can maintain its population on these weedy hosts in the absence of cruciferous crops (Colhoun 1958; Karling 1968; Buczacki and Ockendon 1979; Ludwig-Müller et al. 1999). This can reduce the effectiveness of crop rotation as a strategy for clubroot control.

1.1.7 Physiologic specialization

Physiologic specialization occurs in the clubroot pathosystem, and many pathotypes or races are known to exist. Pathotypes refer to populations or isolates of *P. brassicae* collected from infested soil or plant material that show differential virulence on different hosts (Buczacki et al. 1975). Numerous host differential sets have been proposed over the past 60 years to identify pathotypes of *P. brassicae*. Among the most widely used have been the systems of Williams (1966), Somé et al. (1996) and the European Clubroot Differential (ECD) (Buczacki et al. 1975). The differential hosts of Williams include two rutabagas (*B. napus* ssp. *napobrassica*) 'Laurentian' and 'Wilhelmsburger', and two cabbages (*B. oleracea* var. *capitata*) 'Badger Shipper' and 'Jersey Queen'. These four hosts can distinguish a theoretical maximum of 16 pathotypes, but may not be ideally suited to identify pathotypes of *P. brassicae* recovered from canola production systems (Strelkov and Hwang 2014). In the 1970s, a new differential system, the ECD set, was proposed, which consisted of 15 hosts

(five hosts each of B. *rapa*, *B. napus*, and *B. oleracea*) (Buczacki *et al.* 1975). While it has a greater differentiating capacity than the hosts of Williams, the ECD system also has a complicated pathotype nomenclature and includes several redundant hosts; as such, it has not been widely utilized outside of Europe. Another system, developed in France by Somé *et al.* (1996), includes only three *B. napus* genotypes, but also appears to lack differentiating capacity, at least with respect to Canadian populations of *P. brassicae* (Strelkov et al. 2018).

The recent identification of new virulence phenotypes in *P. brassicae* populations recovered from previously clubroot resistant (CR) canola in western Canada (Strelkov et al. 2016) highlighted the limitations of some of these differential sets. These 'new' populations of the clubroot pathogen, although differing in their virulence on CR canola, could not be distinguished from the 'old' populations based on the reactions of the differentials of Williams (1966), the ECD set (Buczacki et al. 1975), or Somé et al. (1996). As such, a new Canadian Clubroot Differential (CCD) set was developed by Canadian researchers, which includes 13 genotypes including all of the differentials of Williams and Somé et al., selected hosts of the ECD set, as well as the canola/rapeseed cultivars 'Mendel', 'Westar' and '45H29' (Strelkov et al. 2018). Pathotypes are distinguished by their virulence patterns on these differential hosts, with each unique pathotype designated with an uppercase letter. Since the CCD set includes all of the hosts of Williams and Somé et al., it is also possible to assign pathotype designations according to those systems. For example, pathotypes A and H on the CCD Set share similar virulence patterns but differ in their ability to cause disease on the CR canola '45H29' (pathotype A is virulent on '45H29', while H is avirulent on these hosts) (Strelkov et al. 2018). Both pathotypes, however, are classified as pathotype 3 on the differentials of Williams (1966). As such, they are widely referred to as pathotypes 3A and 3H.

1.2 Management of Clubroot

1.2.1 Background

The management of clubroot is challenging, since *P. brassicae* has a wide host range and produces very large numbers of long-lasting resting spores that are difficult to eradicate (Buczacki and Ockendon 1979; Wallenhammar 1996). Nonetheless, cultural and chemical control, the deployment of resistant cultivars, and biocontrol have all been suggested for the management of this disease (Krupinsky *et al.* 2002; Khoury and Makkouk 2010). Strategies for the cultural management of clubroot can include rotation away from susceptible crops, planting of bait crops, and the application of soil amendments to make conditions less favorable for *P. brassicae* (Donald and Porter 2009). The application of fungicides and soil fumigants to infested soil also has been evaluated for the control of clubroot (Donald and Porter 2009; Hwang et al. 2012). Cultural and chemical management strategies have produced mixed results, however, and the planting of CR hosts is often regarded as the most convenient and effective way to manage this disease. While resistant varieties of different Brassica crops are available, sources of resistance appear to be limited and there is a risk of resistance breakdown (Rahman et al. 2014; Strelkov et al. 2016, 2018). An integrated approach, which combines multiple control strategies, is required for the effective and sustainable management of clubroot.

1.2.2 Cultural control

Rotation

Crop rotations with non-host crops are effective in reducing clubroot incidence and severity, especially in regions with recent outbreaks (Donald and Porter 2009). While resting spores have been estimated to persist for up to 17 years in the soil, (Wallenhammar 1996), a ≥ 2 year rotation away from susceptible hosts seems to be effective for significantly reducing *P.*

brassicae spore concentrations in the soil (Peng et al. 2015; Ernst et al. 2019). Nevertheless, the efficacy of crop rotation varies depending on the non-host species included in the rotation (Ikegami 1985) and the prevalence of weedy hosts in a field (Colhoun 1958). In theory at least, long rotations can reduce the concentration of soil inoculum in heavily infested fields to a controllable level (Donald and Porter 2009).

Rouging and bait crops

Rouging refers to the removal of infected plants, and was recommended in early studies from the 1940s and 1950s (reviewed by Donald and Porter 2009). The removal of diseased root tissues can help to prevent the release of additional inoculum. The planting of bait plants can also be used to reduce soil inoculum levels (Yamagishi et al. 1986; Murakami et al. 2000). Susceptible hosts stimulate the germination of *P. brassicae* resting spores, so the planting and early removal of these bait crops, before the pathogen can form a new generation of spores, can help to deplete the soil inoculum level (Murakami *et al.* 2004). Non-host crops such as leek (*Allium ampeloprasum* var. *porrum* (L.) Gay), rye (*Secale cereale* L.), perennial ryegrass (*Lolium perenne* L.) and red clover (*Trifolium pratense* L.) may also be effective bait crops, since their root exudates stimulate resting spore germination (Friberg *et al.* 2005, 2006). However, the planting of bait crops may not be economical due to the extra costs and labour involved, and does not appear to be effective in heavily infested fields (Ahmed et al. 2011; Harling and Kennedy 1991; Donald and Porter 2009).

Soil amendments

Soil amendments that increase soil pH and/or calcium content have been effective in reducing clubroot incidence and severity on vegetable Brassica crops (Donald et al. 2004; Tremblay et al. 2005). For example, Myers and Campbell (1985) observed fewer zoosporangia on root hairs and less galling symptoms on the hosts when 1.5 mM calcium was included in the soil

and the pH was > 7.1. Webster and Dixon (1991) also suggested that an alkaline environment restricts the development of plasmodia and zoosporangia, consequently limiting secondary infection. However, under field conditions in central Alberta, only the highest rates (7.5t/ha or 8.5t/ha) of lime evaluated could reduce clubroot severity (Hwang et al. 2011). A recent study by Fox (in-progress) found that hydrated lime provides better clubroot control than limestone.

Calcium cyanamide also has been used in Europe as a soil amendment in clubroot-infested soil (reviewed by Dixon 2017). It reacts with water and forms hydrated lime and cyanamide in the soil; while the calcium salt contributes to control by the mechanisms described above, the cyanamide anion is fungitoxic, providing additional control (Conforth 1971; Dixon 2017). The chemical is also an effective fertilizer. Studies show that calcium cyanamide considerably reduces clubroot severity on cruciferous vegetables and oilseed rape in Europe (Dixon and Brokenshire 1981; Dixon and Wilson 1983; Naiki and Dixon 1987). However, in field experiments conducted in Canada, mixed results were observed. While calcium cyanamide had no or little effect on clubroot development on canola, it did reduce the severity of the disease in cauliflower (Tremblay et al., 2005; Hwang et al. 2011).

1.2.3 Chemical control

Fungicides

Systemic and contact fungicides are differentiated by their modes of action (Dixon 2001; Tilman et al. 2002). Contact fungicides have an effect on multiple metabolic processes of the target pathogen, but may also be toxic to different species and microflora. Thus, most newer fungicides are less toxic systemic compounds that have a highly specific biochemical target, minimizing the effects on the environment. A variety of products have been evaluated for the

control of clubroot, both as soil drenches (to reduce soilborne inoculum) and as seed treatments (to provide protection from the inoculum present in the soil).

The fungicides azoxystrobin, flusulfamide, and carbathiin and thiram in combination (Vitavax RS) were tested as seed treatments in field experiments conducted in western Canada, but they did not appear to significantly suppress clubroot development (Hwang et al. 2011, 2014). The fungicides thiophanate methyl, PCNB (pentachloronitrobenzene), flusulfamide, fluazinam and cyazofamid have been used to manage clubroot in vegetable systems, and are usually applied as soil drenches (Buczacki et al. 1976; Colhoun 1958; Mitani et al. 2003; Donald et al. 2001; Adhikari 2010). Some of these products are now off the market due to environmental concerns. Studies with fungicide soil drenches in western Canada indicated that the level of control they afforded was limited and not sufficient to justify the costs associated with their application (Hwang et al. 2011, 2014).

Soil fumigation

Soil fumigants are toxic chemicals that form a gas following application, thereby helping to control pests. Unfortunately, these chemicals affect not just the target pest species, but can also the entire soil microbial community (White and Buczacki 1977; Donald and Porter 2009). Fumigants may also have phytotoxic effects, particularly if the interval from fumigation to sowing of the crop is not sufficiently long (Noling 2008). Many chemical fumigants have been prohibited in recent years due to ecological and safety concerns (Donald and Porter 2009). For example, White and Buczacki (1977) suggested that chloropicrin, dazomet and methyl bromide were effective for clubroot control, but methyl bromide has been withdrawn from the market worldwide (Donald and Porter 2009). Other fumigants, such as Dazomet, are still available but their cost may be prohibitive for large-scale applications (Ahmad 1994; Buczacki and White 1979; Donald and Porter 2009). Despite these limitations, there has been interest in soil fumigation to control new or localized infestations by P.

brassicae. Currently, metham sodium (Metham or Vapam) is a preferred fumigant for clubroot management (Donald and Porter 2009; Hwang et al. 2014). While the consistency of this product was uncertain in some earlier studies (Wiggell et al. 1961; White and Buczacki 1977), its application did result in significant reductions in clubroot on canola in field experiments conducted in western Canada (Hwang et al. 2014). As such, it may have some potential in the containment of new introductions of *P. brassicae* where the pathogen is not otherwise present.

1.2.4 Biocontrol

The control of *P. brassicae* with other microorganisms is regarded as one of the most environmentally friendly approaches to clubroot management (Donald and Porter 2009; Peng et al. 2011). However, biocontrol is limited by the availability and the consistency of the biocontrol agents. Kim et al. (2004) examined hundreds of isolates of Streptomyces sp. from Korea and found that one isolate KACC91027 could significantly supress P. brassicae by releasing an unknown active compound. Unfortunately, no follow-up studies appear to have been published. Arie et al. (1998, 1999) identified and tested an isolate of *Phoma glomerate* that generates the mycotoxin epoxydon, which appeared to supress the development of *P. brassicae* by inducing antiauxin activity. Narisawa *et al.* (1998, 2005) observed that the endophytic fungus *Heteroconium chaetospira* reduces clubroot on Chinese cabbage, especially at lower inoculum levels. In a follow-up study, Usuki and Narisawa (2007) found that *H. chaetospira* colonizes the roots of Chinese cabbage and feeds nitrogen back to this host.

In Canada, Peng et al. (2011) evaluated various registered commercial biocontrol products, including *Bacillus subtilis* (Serenade), *Gliocladium catenulatum* (Prestop), *Streptomyces griseoviridis* (Mycostop), *S. lydicus* (Actinovate) and *Trichoderma harzianum* (Root Shield), for their efficacy against clubroot. These researchers found that

while *B. subtilis*, *G. catenulaum* and *S. griseoviridis* reduced clubroot severity by 61-91% in experiments conducted under controlled conditions, they did not have a significant effect in field trials (Peng et al. 2011). At present, biocontrol is not used as a tool in the management of clubroot, at least in Canada. Nevertheless, biocontrol may have potential for the long-term management of this disease, and additional research may be warranted.

1.2.5 Resistance and breeding

Brassica genomes

The evolutionary and genetic relationships between the six major Brassica species are illustrated in the "Triangle of U" (U 1935; Fig.2), and have largely constituted the foundation of modern Brassica breeding. Most of the globally important Brassica crops are varieties of *B. rapa*, *B. napus* and *B. oleracea*. *Brassica rapa* (A genome) includes turnips (*B. rapa* var. *rapa*), Chinese cabbage (*B. rapa* ssp. *pekinensis*) and bok choy (*B. rapa* ssp. *Chinensis*); *B. oleracea* (C genome) includes cabbage (*B. oleracea* var. *capitata*), kale (*B. oleracea* var. *viridis*), broccoli (*B. oleracea* var. *italica*) and cauliflower (*B. oleracea* var. *botrytis*). *Brassica napus* (AC genome) was derived from the hybridization of *B. rapa* and *B. oleracea*, and includes rapeseed or canola (*B. napus* var. *napus*) and rutabaga (*B. napus* ssp. *napobrassica*).

Clubroot resistance in *B. rapa*

Breeding of CR Chinese cabbage (*B. rapa* var. *pekinensis*) was initiated in Japan in the 1960s, since most cultivars of this species are highly susceptible to clubroot (Yoshikawa 1983; Hirai 2006). European fodder turnip (*B. rapa* var. *rapifera*), including the cultivars 'Siloga', 'Gelria', 'Milan White' and 'Debra' with reported clubroot resistance, have been used as major donors in breeding programs in both Japan and China (Kuginuki et al. 1997; Suwabe et al. 2003, 2006; Piao et al. 2004; Hirai et al. 2004). The resistance in European

fodder turnips is believed to be controlled by several genes independently (Piao et al. 2009). Multiple A-genome clubroot resistance genes from *B. rapa* have been identified and mapped. The genes *CRa, CRb, CRb^{Kato}, CRk* were mapped to chromosome A03 (Matsumoto et al. 1998; Piao et al. 2004; Sakamoto et al. 2008; Kato et al. 2013), while *CRc* and *Crr1* were mapped to chromosomes A02 and A08, respectively (Suwabe et al. 2003; Matsumoto et al. 2012).

Clubroot resistance in *B. oleracea*

Crisp et al. (1989) assessed 1047 *B. oleracea* accessions for resistance to two *P. brassicae* isolates from the UK, and found that most resistant accessions were kale (*B. oleracea* var. *viridis*), Brussels sprouts (*B. oleracea* var. *gemmifera*) or cabbages (*B. oleracea* var. *capitata*). Hasan et al. (2012) identified four cabbage accessions (of 48 tested) that were resistant to five *P. brassicae* isolates from Canada. Most recently, Fredua-agyeman et al. (2019) reported that nine of 65 *B. oleracea* accessions evaluated were resistant or moderately resistant to a collection of 22 *P. brassicae* isolates from Canada, including seven kale and two Brussels sprouts. Clubroot resistance in *B. oleracea* appears to be mainly quantitative and reflects the cumulative effect of multiple genes. For example, the resistance in the cabbage cultivars 'Bohmerwaldkohl' and 'Badger Shipper' was shown to be controlled by additive and/or recessive genes (Crute and Pink 1989; Laurens and Thomas 1993; Voorrips and Visser, 1993; Yoshikawa, 1983)

Clubroot resistance in *B. napus*

A few lines of rapeseed or canola (*B. napus* var. *napus*) have been reported with clubroot resistance, including the European winter-type 'Mendel' and 'Tosca', New Zealand resistant rape selections, and canola hybrid '45H29' (Buczacki et al. 1975; Diederichsen et al. 2003, 2006; Strelkov et al. 2018). Resistance in many commercial CR rapeseed or canola varieties

has been introduced via intra- or inter-specific crosses from *B. rapa*, *B. oleracea* and rutabaga (*B. napus* spp. *napobrassica*) (Fredua-Agyeman and Rahman 2016; Fredua-Agyeman et al. 2018; Rahman et al. 2014). The clubroot resistance in most Canadian canola cultivars appears to have been derived from 'Mendel' (Fredua-Agyeman et al. 2018).

Rutabaga is known as a good source of clubroot resistance, especially compared with rapeseed or canola in the same genus (Fredua-Agyeman et al. 2019; Hasan et al. 2012; Rahman et al. 2014). Rutabaga cultivars such as 'Wilhemsburger', 'Laurentian' and 'Danish Giant' are used commonly in breeding programs and possess resistance to different *P. brassicae* pathotypes (Shattuck and Proudfoot 1990). Nonetheless, the resistance in some rutabaga lines may come from other species (Fredua-Agyeman et al. 2019). The European CR rutabaga 'Invitation' was derived from the hybridization of *B. rapa* ssp. *rapifera* line AABBCC and the cabbage 'Bohmerwaldkohl' (Bradshaw et al. 1997). Shattuck and Proudfoot (1990) also noted that early breeding programs in Canada used European fodder turnip lines to develop CR rutabaga.

Clubroot-resistance within the *Brassicaceae*

Several recent large-scale screening projects have identified resistance in other Brassicas (Hasan et al. 2012; Peng et al., 2014; Ren et al. 2016; Fredua-Agyeman et al. 2019). Accessions of *B. nigra* in particular appear to be valuable sources of resistance. In independent resistance screening of Brassicas belonging to various species, Hasan et al. (2012) and Peng et al. (2014) found that almost all of the *B. nigra* accessions tested possessed resistance to 'old' strains of the clubroot pathogen (collected prior to the loss of resistance in CR canola). Similarly, 37 of 63 *B. nigra* accessions screened by Fredua-Agyeman et al. (2019) were resistant or moderately resistant to 22 *P. brassica* isolates originally recovered from CR and non-CR hosts. In contrast, none of the *B. juncea* or *B. carinata* accessions tested

in these studies appeared to be resistant, although one *B. juncea* accession from China did have good resistance to pathotype 4 (Ren et al. 2016).

1.2.6 Marker-based techniques for the identification of clubroot-resistance genes

Molecular markers, including amplified fragment length polymorphisms (AFLP), cleaved amplified polymorphic sequences (CAPS), random amplification of polymorphic DNA (RAPD), restriction fragment length polymorphisms (RFLP), sequence characterized amplified regions (SCAR), sequence tagged sites (STS) and simple sequence repeats (SSR), have been used widely for linkage-based identification and mapping of both qualitative and quantitative clubroot-resistance gene loci (Hasan and Rahman 2016; Landry et al. 1992; Voorrips et al. 1997; Matsumoto et al. 1998; Piao et al., 2004; Saito et al. 2006; Suwabe et al., 2006; Ueno et al. 2012; Zhang et al. 2014). These markers are used to amplify products by PCR, with the amplicon sizes obtained from different individuals determined by electrophoresis (Vignal et al. 2002). More recently, however, the use of the novel single nucleotide polymorphism (SNP) arrays has emerged as the preferred method for the identification of clubroot-resistance genes; the advantages of SNP arrays include high coverage and marker density for linkage map construction or genotyping.

The biochip-based SNP array requires less time and can provide more consistent results than PCR-based approaches (Gupta et al. 2014; Vignal et al. 2002). The recent advent of fairly inexpensive high throughput sequencing technologies has enabled whole genome sequencing of the Chinese cabbage ‘Chiifu-401-42’ (Wang et al. 2011), cabbage line 02–12 (Liu et al. 2014), and the winter rapeseed ‘Darmor-bzh’ (Chalhoub et al. 2014), as well as *B. juncea* and *B. nigra* accessions (Yang et al. 2016). Highly abundant SNPs and insertions and deletions (InDels) were found throughout these genomes, and used to design array systems for various *Brassica* species (Clarke et al. 2016; Gupta et al. 2014; Liu et al. 2013). The *B.*

napus SNP array achieved a high marker density of one marker per 0.08 cM (Clarke et al. 2016), compared with a recent linkage map of the same species generated by PCR-based markers (SSR and AFLP) of an average of 2.77 cM per marker (Nurhasanah and Ecke 2016). The high-resolution detection of gene-level polymorphisms with SNP arrays also made possible the identification of clubroot-resistance gene loci via genome-wide association studies (GWAS), since gene mapping by conventional PCR-based markers usually required pedigree information and near-isogenic lines (Gupta et al. 2014; Neik et al. 2017).

1.3 Rutabaga as a source of clubroot resistance for canola

1.3.1 Background

The first CR canola cultivars in Canada were released in 2009-2010 and carried strong resistance to pathotypes 3 and 5 of *P. brassicae*, as defined on the differentials of Williams (1966) (Strelkov and Hwang 2014). Most of these cultivars appear to carry the *CRa* and/or CRb^{Kato} genes from *B. rapa* spp. *rapifera* line AABBCC or from the rapeseed 'Mendel' (Fredua-Agyeman et al. 2018). Given their efficacy against the prevalent pathotypes of *P. brassicae* in western Canada, these CR canola cultivars soon began to be grown over large acreages in this region, often in short rotations. This resulted in significant selection pressure on *P. brassicae* populations, leading to the emergence (starting in 2013) of 'new' pathotypes able to overcome this resistance (Strelkov et al. 2016). The number of fields with resistance issues has continued to increase since 2013 (Strelkov et al. 2019), and pathotypes that can overcome resistance had been recovered from nearly 200 fields by 2018 (S.E. Strelkov, personal communication). Testing on the differentials of the CCD set has identified at least 11 distinct pathotypes in Alberta that can overcome resistance, with pathotype 3A being predominant (Strelkov et al. 2018). It is clear that new and effective sources of

resistance must be identified to complement other clubroot management methods, and as noted earlier, rutabaga may represent an important such source.

Taxonomically, rutabaga (also known as swede) is grouped with rapeseed or canola; both are *B. napus* derived from the hybridization of *B. rapa* and *B. oleracea*. Rutabaga is grown in Europe, North America, New Zealand, Australia and China (Gowers 2010) and is known as a source of clubroot resistance genes (Ayers and Lelachur 1972, Shattuck and Proudfoot 1990; Bradshaw et al. 1997; Hasan et al. 2012, Fredua-Agyeman et al. 2019). Indeed, the rutabagas 'Wilhemsburger' and 'Laurentian' were included in multiple clubroot differential sets because of their resistance to different isolates of *P. brassicae* (Williams 1966, Buczacki et al. 1975; Somé et al. 1996; Strelkov et al. 2018). Since this crop is fully fertile with canola, it is particularly amenable for clubroot resistance breeding in canola (Gowers 2010, Rahman et al. 2014).

1.3.2 Rutabaga

Rutabaga is cultivated as a table vegetable and as fodder for animals (Gowers 2010). It is rich in sugar content and vitamins A and C. It is also a source of folate, potassium, dietary fiber, calcium, iron, and niacin for human consumption (Bradshaw and Griffiths 1990; Gowers 2010). Rutabaga requires vernalization for at least eight-weeks prior to flowering and propagation, but the bulbs may be harvested for consumption without this process (Gowers 2010). The bulb's appearance, chemical compounds, disease resistance and yield are of breeding interest (Bradshaw and Griffiths 1990; Gowers 2010). The main selection criterion for the yield of rutabaga is the dry matter content and quality parameters including sugar content (Bradshaw and Griffiths 1990). Although rutabaga has a relatively low dry matter content compared with other animal feeds, its digestibility and the metabolizable energy content of the leaves and bulbs are of high standards for animal consumption (Griffiths et al.

1991; Bradshaw and Griffiths 1990). Rutabaga cultivars may vary in both bulb and leaf shape and colour, foliage growth habits, quality parameters such as diameter and fresh weight, and disease resistance (Grant et al. 1982; Gemmell et al. 1990; Gowers 2010).

History of rutabaga

Although the Swiss botanist Caspar Bauhin noted the cultivation of rutabaga in Sweden in 1620 (reviewed by Gowers 2010), there is a lack of information to confirm that Sweden is the sole center of origin of this crop. The Swedish literature indicates that rutabaga originated from Gotland Island in Sweden (A.C. Wallenhammer, personal communication), while Ahokas (2004, reviewed by Gowers 2010) suggested that this crop may have originated in Ingria Province, Finland. Since Norway, Sweden and Finland are not isolated geographically, a well-accepted hypothesis is that rutabaga originated from the Scandinavian region. Regardless of its exact origin in Scandinavia, rutabaga was introduced to the United Kingdom via Germany in the18th century (Harvey 1949), and then brought to North America and Australia by European immigrants in the early 19th century (Sturtevant 1919; reviewed by Gowers 2010).

1.3.3 Clubroot resistance in rutabaga

Clubroot resistance in rutabaga has been widely reported and studied (reviewed by Piao et al. 2009 and Rahman et al. 2014). Early studies noted rutabaga lines showing clubroot resistance under controlled environmental or field conditions (Karling 1968; Lammerink 1967; Ayers and Lelachur 1972). Ayers and Lelachur (1972) reported that the rutabaga cultivars 'York' and 'Wilhemsburger' were resistant to pathotypes 2 and 3 of *P. brassicae* (as defined on the differentials of Williams 1966), while the cultivar 'Ditmars S2' was resistant to pathotype 3. There are one or two major dominant genes in these cultivars. The gene in 'Wilhemsburger' that confers resistance to pathotype 2 is a single dominant gene designated *CR2a* by Landry

et al. (1992). Broad-spectrum resistance also was reported in 'Wilhemsburger' to the 17 *P. brassicae* pathotypes identified from Canada on the CCD set, while 'Laurentian' was resistant to seven of these pathotypes (Strelkov et al. 2018). Hasan et al. (2012) found that three rutabaga accessions were resistant to the *P. brassicae* pathotypes 2F, 3H, 5I, 6M and 8N, as defined on the CCD set (Strelkov et al. 2018). Another study, by Hasan and Rahman (2016), mapped the resistance gene of one accession by simple sequence repeat (SSR) markers to a genomic region on the A08 chromosome of *B. rapa.* Nonetheless, no in-depth molecular studies have been conducted on the differential reactions of rutabaga accessions to *P. brassicae.*

1.4 Research objectives and hypotheses

Given the importance of clubroot as a disease of *B. napus* and other crucifers in Canada and worldwide, and given the importance of rutabaga as a source of clubroot resistance, this thesis had two main objectives: (1) to evaluate the genetic diversity and population structure of 124 rutabaga accessions from five Nordic countries (Norway, Sweden, Finland, Denmark and Iceland); and (2) to screen this collection of rutabagas for resistance to important *P. brassicae* pathotypes from Canada, characterizing any resistance identified using a suite of molecular genetics tools. I hypothesize that (1) high genetic variability will be found among the rutabaga accessions as a result of domestication and selection, and (2) sources of resistance effective against the 'new' pathotypes of *P. brassicae* from Canada will be identified in the collection,

1.5 Figures

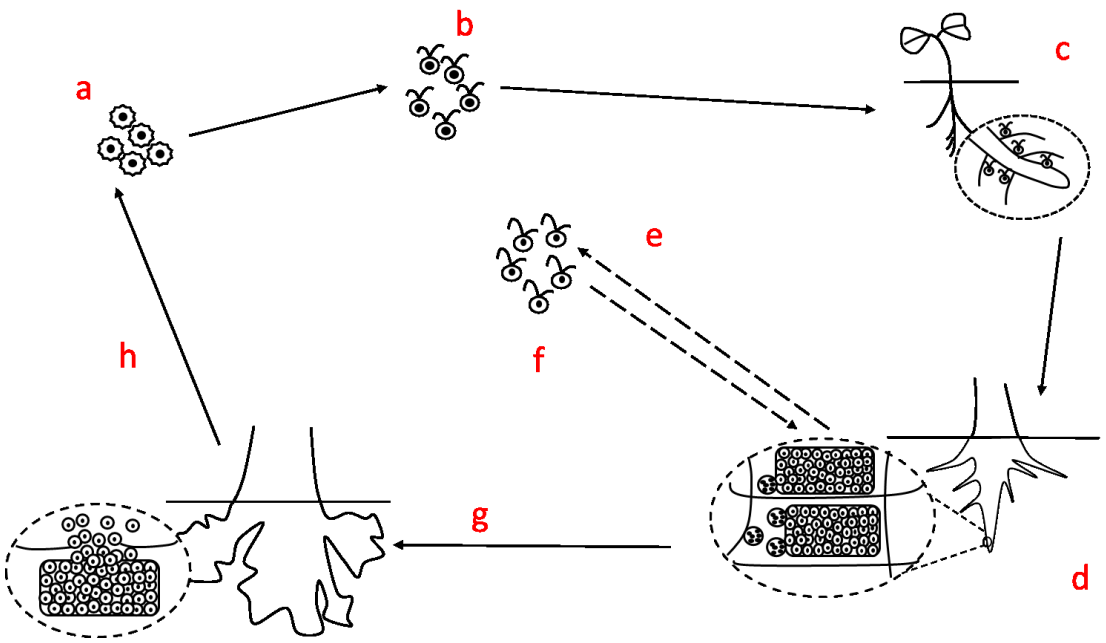

Figure 1.1 Life cycle of *Plasmodiophora brassicae* (adapted from Kageyama and Asano 2009; Schwelm et al. 2015). a. Resting spores in the soil; b. Zoospores germinated from the resting spores; c. Zoospores penetrate the host root hairs; d. Primary plasmodia form within the root hairs; e. Secondary zoospores released from zoosporangia formed from the primary plasmodia; f. Secondary zoospores penetrate the cortical tissues and form secondary plasmodia; g. The plasmodia develop into resting spores as the host plant matures; h. Release of resting spores into surrounding soil.

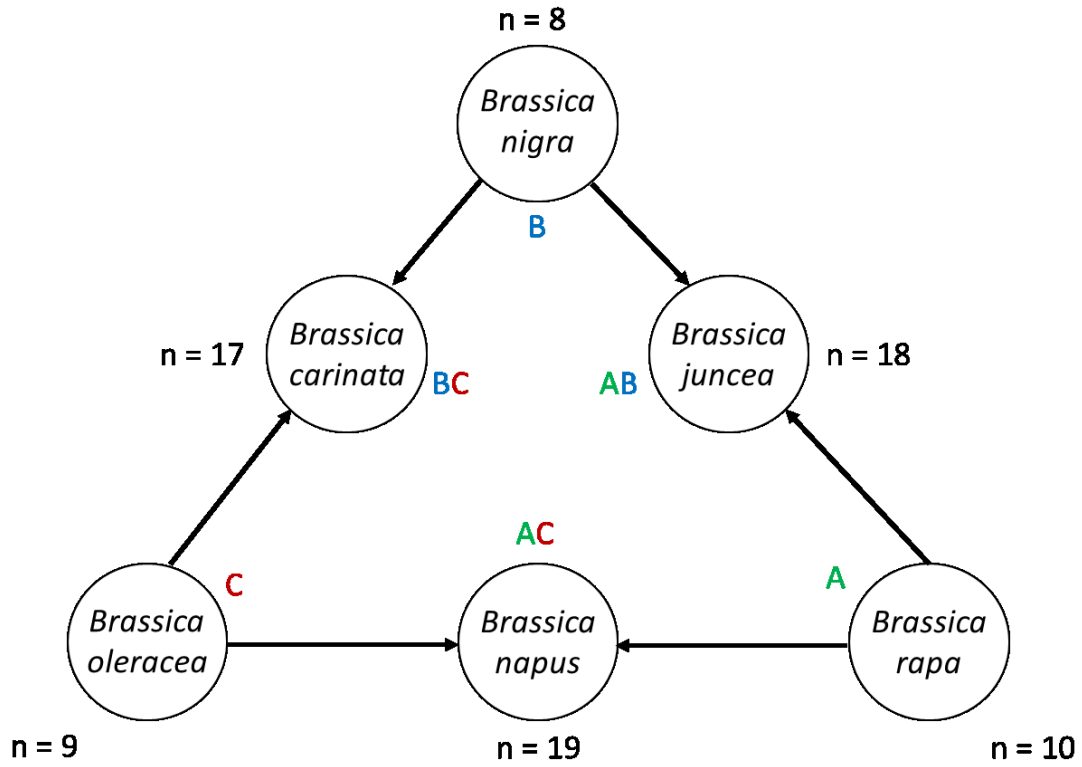

Figure 1.2 The "Triangle of U". A, B, C = genome type; n = number chromosome sets (Based on U 1935).

Chapter 2: Genetic diversity and population structure of rutabaga accessions from Nordic countries

2.1 Introduction

Brassica napus ssp. *napobrassica* (L.) Hanelt, called 'rutabagge' in Sweden, 'rutabaga' in the USA and Canada, and 'swede' in the UK, New Zealand and Australia, is a cool-weather root crop thought to have been derived from the natural or spontaneous hybridization between *B. rapa* (turnip) and *B. oleracea* (cabbage or kale) (Iñiguez and Federico 2011). Rutabaga is often assumed to have originated in Sweden, but may have come from Finland (Ahokas 2004 as cited in Gowers 2010). Nevertheless, it was distributed from Sweden (where it grew in the wild before 1400) to England, Germany and other European countries at about the end of the 18th century (Harvey 1949) and was introduced to North America by European immigrants in the early 19th century (Sturtevant 1919). Therefore, the Nordic countries are considered the center of rutabaga domestication and diversity.

Rutabagas are grown for use as a table vegetable and as fodder for animals (Gowers 2010). The roots are rich in vitamins A, C and fibre; are low in calories and have trace amounts of vitamin B_1, B_2, potassium, calcium, magnesium and iron (Bradshaw and Griffiths 1990; Gowers 2010). Like most cruciferous vegetables, they have antioxidant and anti-cancer properties (Pasko et al. 2013). The leaves have much higher levels of protein (17-18%) than the roots (0.6-2.0%) (Pivovarova, 1979; Jung et al., 1986). However, most of the components are non-protein nitrogen (urea and ammonia), which can be converted into protein by microbes in the stomach of ruminants, but not in pigs (Livingstone et al. 1977).

Rutabagas vary considerably in morphology, disease resistance, seed yield and quality parameters such as erucic acid and glucosinolate content (Gemmell et al. 1990; Gowers, 2010). Breeding efforts have targeted root appearance and flesh colour, earliness, drought

tolerance, improvement in resistance to diseases, broadening genetic diversity and quality traits associated with the seeds (Ayers and Lelacheur 1972; Bradshaw and Griffiths 1990; Gowers 2010; Hasan et al. 2012; Hassan and Rahman 2016). Quantitative traits such as root length, diameter and fresh weight are also of interest for crop improvement (Grant et al. 1982).

Genetic variation in plants is a key pillar of biodiversity and provides the resources for the development of new and improved cultivars with desirable characteristics (Govindaraj et al. 2015). In addition, studying diversity in natural plant populations makes it possible to understand genetic exchange or gene flow within and between populations (Schaal et al. 1998). Bus et al. (2011) used 89 simple sequence repeats (SSR) markers to estimate genetic diversity in 509 *B. napus* inbred lines of which 73 were swedes or rutabagas. Similarly, Diers and Osborn (1994) used 43 restriction fragment length polymorphism (RFLP) markers to group 83 *B. napus* lines of which two were rutabagas. Mailer et al. (1994) reported that a set of 100 randomly amplified polymorphic DNAs (RAPD) could identify four rutabaga accessions among 23 cultivars of *B. napus*. These molecular studies indicated that spring oilseed rape, winter oilseed rape, fodder and vegetable types, and rutabagas formed separate clusters of *B. napus*.

The previous studies, however, did not examine the genetic differences or similarities that existed within the rutabaga accessions and the genetic variability among populations from different countries. In addition, the number of SSR, RFLP and RAPD markers used in those studies also were relatively very small. Therefore, the aim of the present study was to use high-throughput genotyping with *Brassica* SNP markers to estimate genetic diversity in rutabaga accessions from five Nordic countries (Norway, Sweden, Finland, Denmark and Iceland).

2.2 Materials and Methods

2.2.1 Plant Material

Seeds of 124 rutabaga accessions obtained from the Nordic Genetic Resource Center (NordGen, Sweden) were used in this study, including 23 accessions from Denmark, 12 from Finland, 11 from Iceland, 28 from Norway and 50 accessions from Sweden (Figure 2.1). Hereafter, the rutabaga accessions from the five countries will be referred to as the DNK-, FIN-, ISL-, NOR- and SWE-subpopulations, respectively. In addition, seeds of three commercial rutabaga cultivars 'Laurentian' from Canada, 'Wilhemsburger' from Germany and 'Krasnoselskaya' from Russia were included as the out-group. Details on the accessions are presented in Table 2.6. Two to four seeds of each accession were grown in 13 ×13 ×15 cm pots filled with Sunshine Mix #4 Aggregate Plus Growing Mix (Sungro Horticulture Canada Ltd) and kept in a growth chamber with a 16h/8h (22ºC) day/night cycle for 4 weeks. Leaf tissue (~0.25 g) was collected from two plants of each accession in 1.5 mL microcentrifuge tubes on ice. The samples were stored at -20°C and shipped on dry ice for SNP genotyping.

2.2.2 SNP genotyping and filtering

SNP genotyping of the 124 rutabaga accessions and three commercial cultivars was performed with a 15K SNP *Brassica* array at TraitGenetics GmbH, Gatersleben, Germany. Filtering was done to remove monomorphic and low coverage site SNP markers, those with minor allele frequency (MAF) < 0.05, and SNPs with missing data for > 5% of the accessions. Six thousand eight hundred sixty-one SNP markers were retained for the calculation of the genetic diversity indices and the population structure analyses.

2.2.3 Allele frequency-based population structure analyses

The proportion of polymorphic loci (%*P*), the mean number of alleles per locus (*Na*), the mean number of effective alleles per locus (*Ne*), the mean expected heterozygosity ($\overline{H}_e$), the mean unbiased expected heterozygosity ($U\overline{H}_e$), the mean number of alleles with a frequency ≥ 5% (*Na Freq* ≥ 5%), mean number of common alleles found in ≤ 25% and ≤ 50% of the subpopulations (*Na comm* ≤ 25% and *Na comm* ≤ 50%; respectively) and Shannon's information index (*I*), within and among subpopulations, as well as Wright's (1965) genetic differentiation F-statistics (F_{ST}) between the populations were determined with GenAlEx 6.5 (Peakall and Smouse 2006, 2012). The F_{ST} values were assessed at 1000 random permutations across the 6861 loci.

In addition, the polymorphism information content (PIC), minor allele frequency (MAF) and the expected heterozygosity at any given locus (*He*) also called gene diversity (*D*) (Weir and Cockerham 1984) were evaluated for the DNK-, FIN-, ISL-, NOR- and SWE-subpopulations and the entire population using POWERMAKER v3.25 (Liu and Muse 2005).

2.2.4 Distance-based population structure analyses

The genetic and similarity distance matrices within and among the subpopulations were calculated for the 6861 SNP markers and the 124 accessions using both GenAlEx 6.5 (Peakall and Smouse 2006, 2012) and TASSEL v5.2.2.5 (Bradbury et al. 2007).

The matrices were used to test the hierarchical partitioning of the analysis of molecular variance (AMOVA) among regions, populations and within accessions and their level of statistical significance was assessed based on 10,000 permutations (Excoffier et al.1992). In addition, patterns in the population were inferred or visualized by Principal Coordinates Analysis (PCoA) (Patterson et al. 2006). The AMOVA and PCoA were conducted with GenAlEx 6.5. The unweighted pair group method with arithmetic mean (UPGMA) (Sokal and Michener 1958) and neighbour-joining (NJ) (Saitu and Nei 1987) clustering methods

implemented in TASSEL v5.2.2.5 (Bradbury et al. 2007) were used to generate phylogenetic trees.

2.2.5 Bayesian population structure analyses

A Bayesian clustering approach, applying a Markov Chain Monte Carlo (MCMC) algorithm implemented in the population-genetic software *STRUCTURE* v2.3.4 (Pritchard et al. 2000), was used to assign the 124 rutabaga accessions from the various countries into a number of genetically homogeneous clusters (*K*) based on the 6861 SNP markers. In addition, *STRUCTURE* was used to assign the rutabagas 'Laurentian' (Canada), 'Wilhemsburger' (Germany) and 'Krasnoselskaya' (Russia) into the Nordic subpopulations showing similar variation patterns.

STRUCTURE was run at a number of burn-in periods from 5000 to 20000 iterations and MCMC analyses from 5000 to 50000 permutations and with the accessions unassigned to any population or country. Runs for each K=1-10 were replicated 10 times. Overall, nine runs of burn-in period and MCMC values of 5000 × 5000, 5000 × 10000, 5000 × 20000, 5000 × 50000, 10000 × 10000, 10000 × 20000, 10000 × 50000, 20000 × 20000 and 20000 × 50000, respectively, were conducted. Each run was repeated three times to identify the burn-in period and MCMC value which gave consistent clusters. The replicate runs for the different K values for the parameters which gave consistent clusters were aligned and compared with CLUMPP (Jacobsson and Rosenberg 2007), and the results displayed graphically as bar plots using DISTRUCT (Rosenberg 2004). The most likely number of clusters (the ad hoc ΔK test) and average log-likelihood plots were determined following Evanno et al. (2005) with STRUCTURE HARVESTER v0.6.94 (Earl and vanHoldt 2012). Accessions were assigned to a specific cluster (K) if the probability of membership was ≥ 0.70, with those that did not meet this threshold considered as an admixture.

2.2.6 Statistical analysis

Statistical significance between means of the parameters (pairwise and overall) was established by Fisher's protected least significant difference (LSD) test ($P \leq 0.05$) using SAS v. 9.4 (SAS Institute, Inc., Cary, NC, USA).

2.3 Results

2.3.1 SNP marker characteristics

Thirteen thousand seven hundred four SNP markers on the 15K SNP *Brassica* chip were used to screen the 124 rutabaga accessions and three rutabaga cultivars. Among these, 31% (4213 SNPs) were monomorphic, 5% (701 SNPs) were low coverage site markers, and 14% (1929 SNPs) were missing data points for > 5% of the accessions. Thus, filtering removed ≈ 50% of the SNP markers while the remaining ≈ 50% (6861 SNPs) were retained for the diversity analysis. This comprised 4390 A-genome and 2471 C-genome SNP markers.

2.3.2 Allelic patterns and genetic diversity indices among and within populations

Allelic patterns and genetic diversity summary statistics at any given locus or averaged across the 6,861 SNP loci for the rutabaga accessions separately for each country and for the whole collection are presented in Table 2.1 and Figure 2.2 (A to D).

The %*P* detected separately for the NOR-, SWE-, FIN- and DNK- subpopulations were significantly higher (range 88.5-99.6%) than for the ISL-subpopulation (67.9%) ($P < 0.05$) (Table 2.1). The mean *Na* was highest in the SWE-subpopulation (2.236 ± 0.005) and lowest in the ISL-subpopulation (1.707 ± 0.006) (Table 2.1). Similarly, the mean *Ne* and *I* were significantly higher in the SWE-subpopulation (1.590 ± 0.004 and 0.535 ± 0.002, respectively) compared with the ISL-subpopulation (1.299 ± 0.004 and 0.305 ± 0.003, respectively) (Table 2.1). In addition, the *Na Freq* ≥ 5% and *Na common* ≤ 50% were lowest

for the ISL-subpopulation (Figure 2A). Thus, most of the genetic diversity indices for the NOR-, SWE-, FIN- and DNK-subpopulations were not significantly different from each other. However, they were all significantly different from the ISL-subpopulation ($P < 0.05$).

The diversity of the SNP markers expressed as the PIC is presented in Figure 2B. The number of markers with PIC > 0.2 was highest for the SWE-subpopulation (5725 ≈ 83%) and DNK-subpopulation (5170 ≈ 75%), intermediate for the FIN- and NOR-subpopulations (4701-4726 ≈ 69%), and lowest among for the ISL-subpopulation (2742 ≈ 40%). The PIC averaged across the 6861 SNP separately for each population followed similar patterns as the allelic and genetic diversity, with the highest PIC occurring in the SWE-subpopulation (0.35) and the lowest in the ISL-subpopulation (0.18).

The number of SNP markers with MAF ≤ 0.1 was of the order ISL- (4106 ≈ 60%) > FIN- (2115 ≈ 31%) > DNK- (1690 ≈ 25%) > NOR- (1518 ≈ 22%) > SWE-subpopulations (933 ≈ 14%). Thus, the frequency of minor alleles was highest for the ISL-subpopulation, intermediate for the FIN-, DEN- and NOR-subpopulations, and lowest for the SWE-subpopulation (Figure 2C).

The expected heterozygosity per locus (H_e or D) followed similar patterns as the rest of the parameters measured with the exception of the MAF (Figure 2D). Analyses of the gene pool structure ($\overline{H}_e$, expected heterozygosity averaged over all 6861 loci) of the rutabaga accessions from each country suggested that there was no significant difference in the genetic variability of the rutabaga accessions from Sweden (0.345 ± 0.002), Denmark (0.301 ± 0.002), Norway (0.292 ± 0.002), and Finland (0.288 ± 0.002). These accessions were, however, genetically different from the accessions from Iceland (0.191 ± 0.002) (Table 2.1).

2.3.3 Genetic differentiation among regions, populations and within accessions

Pairwise comparisons of population differentiation using the fixation index F_{ST} are presented in Table 2.2. The F_{ST} values for all 10 pairwise combinations of all five subpopulations ranged from 0.032 to 0.133. Pairwise F_{ST} values for NOR/SWE, NOR/FIN and SWE/FIN ranged from 0.032 to 0.067 (lowest); the values for NOR/DNK, SWE/DNK and FIN/DNK ranged from 0.050 to 0.88 (intermediate); whereas the F_{ST} values for the ISL/NOR, ISL/SWE, ISL/DNK and ISL/FIN ranged from 0.103 to 0.133 (highest). Overall, the lowest F_{ST} value was found between the SWE- and FIN-subpopulations and the highest between the ISL- and DNK-subpopulations (Table 2.2).

The AMOVA of the distance matrices obtained with Tassel and GenAlEx for the rutabaga accessions were highly correlated (Tables 2.5A and 2.5B). The AMOVA among and within the five populations partitioned the overall genetic variance into three parts: ≈ 94% attributable to within population differences, whereas ≈ 5% and ≈ 1% of the variation occurred among populations and among regions, respectively ($P = 0.108$) (Figure 2.3A). This suggested only minor differences in the entire rutabaga populations from the different countries.

However, pairwise comparison of the AMOVA (Φ_{PT}) between the populations revealed a higher genetic variance (18 to 27%) between the ISL-subpopulation and the NOR-, SWE-, FIN- and DNK-subpopulations (Table 2.3). Furthermore, the rutabaga accessions from Iceland and Denmark were the most genetically diverse ($\Phi_{PT} = 27\%$) followed by accessions from Iceland and Finland ($\Phi_{PT} = 24\%$). In contrast, rutabaga accessions from Sweden and Finland were the most similar ($\Phi_{PT} = 2\%$) followed by accessions from Norway and Sweden ($\Phi_{PT} = 7\%$).

Thus, the vast majority of the genetic variability could be attributed to within population differences. Nevertheless, the pairwise comparison of the subpopulations suggested that considerable variation existed between the rutabagas from the different countries.

2.3.4 Cluster analyses

The principal coordinate analysis (PCoA) based on the 6861 SNP markers clustered the 124 rutabaga accessions into six heterogeneous subgroups (Figure 2.3B) using the first (PCoA1 ≈ 14.7% of genetic variance) and second (PCoA2 ≈ 11.4% of genetic variance) principal components. Clearly, the rutabaga accessions from Sweden, Norway and Finland were distributed across almost all of the subgroups (P1 to P6 in Figure 2.3B). In contrast, the accessions from Iceland and Denmark were concentrated in subgroup P3 and subgroups P1 and P2, respectively (Figure 2.3B).

The neighbour-joining (NJ) based on the 6861 SNP markers clustered the 124 rutabaga accessions into four major branches (Figure 2.3C). The unrooted phylogenetic trees indicated that the accessions from Sweden were distributed into three of the branches (N1, N2 and N3), those from Norway, Finland and Denmark were segregated into two of the branches (N2 and N4, N2 and N3 and N1 and N2, respectively), whereas accessions from Iceland was concentrated in one branch (N2) (Figure 2.3C).

The unweighted pair group method with arithmetic mean (UPGMA) based on the 6861 SNP markers indicated that the trees for the 124 rutabaga accessions were clustered into five major branches (Figure 2.3D). The accessions from Sweden, Norway, and Finland were widely distributed across at least four of the major branches (Figure 2.3D). Similar to the branching patterns in the NJ analysis, the rutabaga accessions from Denmark and Iceland clustered into two branches (U1 and U2) or one branch (U3), respectively.

Overall, the three multivariate analyses (PCoA + NJ + UPGMA) suggested the existence of four to six groups in the rutabaga accessions. However, correlations with their geographic origin were very low, except for the accessions from Iceland.

The unrooted trees used to depict the NJ and UPGMA do not imply a known ancestral root of the three out-groups (which are coloured orange in Figure 2.3C and 2.3D). However, the results suggested that the rutabaga 'Wilhemsburger' was in the first branch (N1 and U1 of the NJ and UPGMA unrooted trees, respectively), while 'Laurentian' and 'Krasnoselskaya' both were grouped in the second branch (N2 and U2 of the NJ and UPGMA unrooted trees, respectively) (Figure 2.3C and 2.3D).

The NJ and UPGMA representation of the similarity matrices as a phylogram (Figures 2.5A and 2.5B) and a circular rooted (Figures 2.6A and 2.6B) diagram are included in the Supplementary Materials. These indicate even closer groupings of the accessions based on their geographic origins.

2.3.5 Population structure analysis

The *STRUCTURE* run of 20000 iterations and MCMC analysis of 50,000 permutations produced the most consistent peaks and hence was used for the final analysis. In this run, the ΔK value was highest at $K = 9$ (Figure 2.4A). Based on a threshold of $P \geq 0.70$, 45.2% of the accessions were placed into one of the nine clusters while 54.8% were classified as admixtures (Table 2.4). Excluding the admixture, 41.7% (5) of the accessions from Finland and 63.6 % (7) of the accessions from Iceland were present in only one cluster. In contrast, 44.0% (22) of the accessions from Sweden, 42.9% (12) of the accessions from Norway and 43.5% (10) of the accessions from Denmark were present in 6, 4 and 3 of the clusters, respectively (Table 2.4). The Canadian rutabaga 'Laurentian' was placed in cluster 5 along with some of the accessions from Denmark, Norway and Sweden. The origin of

‘Wilhemsburger’ from Germany was less ambiguous, since it was placed in cluster 9 together with five accessions from Denmark. The Russian rutabaga ‘Krasnoselskaya’ was an admixture. Overall, the number of clusters obtained in the *STRUCTURE* analysis (nine) was higher than the 4-6 subgroups obtained in the multivariate analysis.

2.4 Discussion

A comprehensive body of literature exists on rutabagas in the main Nordic languages (Personal communication, Prof. Ann-Charlotte Wallenhammar, Swedish Agricultural University). This probably reflects the transmission of seeds and information on agronomic practices for rutabaga cultivation in the Nordic region since medieval times (Harvey 1949). Turesson (1922 a and b, 1925) observed that when the same species of plants were grown in different habitats over many years, they differed from each other in stature, colour, morphology and texture of leaves, stem, flowers and seed. Consequently, rutabagas that are adapted to different climatic and geographic environments will develop different morphological traits.

In this study, SNP markers and combinations of allele- and distance-based population genetics statistics, multivariate clustering and Bayesian methods were used to examine genetic diversity and differentiation in rutabaga accessions from Norway, Sweden, Finland and Denmark and Iceland. Diers and Osborn (1994) used rutabaga accessions as an out-group in genetic diversity studies of *B. napus*, whereas Mailer et al. (1994) and Bus et al. (2011) compared rutabagas with spring oilseed rape, winter oilseed rape, fodder and vegetable types. Fewer than 100 SSR, RFLP and RAPD markers, however, were used in those studies compared with the 6861 SNP markers in the current study.

SNP markers offer enormous potential compared with SSR, RFLP, RAPD and other PCR-based markers, which often require agarose, polyacrylamide or capillary electrophoresis

for product size determination. Inconsistencies in allele size calling makes it difficult to compare data produced with PCR-based assays (Vignal et al. 2002). Furthermore, PCR product sizes vary with different fluorescent dyes, and products of identical sizes may include different alleles due to small deletions and insertions (Vignal et al. 2002). RAPD markers are dominant markers with low reproducibility and accuracy, while RFLP markers have a low discrimination power and can be costly (Vignal et al. 2002). In contrast, SNP markers offer the advantages of co-dominance, high-throughput processing of large numbers of samples, high reproducibility, high accuracy and cost effectiveness (Vignal et al. 2002).

The average of 2.012 ± 0.003 alleles per SNP locus obtained in this study was less than the 4.78 alleles per SSR locus found by Bus et al. (2011). This was expected, since SSR markers are multi-allelic codominant markers while SNP markers are often bi-allelic. The mean expected heterozygosity ($\bar{H}$e) or gene diversity (D) of 0.283 ± 0.001 obtained in this study was less than the 0.43 reported by Bus et al. (2011), likely because the 73 accessions examined in the latter represented a more diverse collection from 19 countries with a much wider geographical distribution (Europe, North America, Asia, New Zealand and the North Africa). The lower allelic diversity summary statistics with SNP markers compared with SSR markers also has been reported in rice (Singh et al. 2013; Gonzaga et al. 2015), barley (Varshney et al. 2008), mushrooms (Tsykun et al. 2017) and other species.

The pairwise fixation index (F_{ST}) obtained in current study ranged from 0.032 to 0.133, which was within the 0.054 reported by Bus et al. (2011). Therefore, the use of SNP and SSR markers confirmed that genetic differentiation in the rutabaga accessions is low and there is a high degree of genetic exchange within these accessions. The observation in earlier studies that rutabagas clustered separately from spring, winter, fodder and vegetable *Brassica* species (Diers and Osborn 1994; Mailer et al. 1994; Bus et al. 2011) could be due to many rutabagas being landraces with different morphological adaptions to various geographic and climatic

regions. The significantly higher pairwise F_{ST} values for ISL-DNK (0.133), ISL-FIN (0.124), ISL-SWE (0.106) and ISL-NOR (0.103) compared with the pairwise F_{ST} values for DNK, FIN, NOR and SWE (range 0.032 to 0.088) may reflect enrichment caused by natural selection in the Icelandic sub-population. As an island in the North Atlantic, Iceland is geographically isolated from the other Nordic countries. This isolation, combined with possible differences in microclimatic and soil conditions, may have resulted in less exchange of germplasm between Iceland and Norway, Denmark, Finland and Sweden.

The low levels of genetic variability and the high genetic exchange among rutabaga accessions from the Nordic countries was illustrated by the variable numbers of clusters in the PCoA, UPGMA and NJ. This was confirmed further by the high degree (~55%) of admixtures detected by the Bayesian population structure analyses. Thus, the population structure analyses were in agreement with the allele diversity summary statistics obtained in this study. In addition, the uneven sample sizes of the rutabaga accessions from the different countries did not appear to affect the measured population indices or clustering. For example, the measured population indices for Finland (n = 12) were not significantly different from those of Denmark (n = 23), Norway (n = 28) or Sweden (n = 50).

In conclusion, three multivariate analyses: principal coordinate analysis (PCoA), the unweighted pair group method with arithmetic mean (UPGMA) and neighbour-joining (NJ) clustering methods as well as the STRUCTURE results grouped the 124 accessions from the Nordic countries into four to nine subgroups. The majority of the genetic differences were present within the Icelandic subpopulations, while accessions from Norway, Sweden, Finland and Denmark were genetically very similar. Given these findings, which were based on molecular genetics analyses, there may be value in additional and more detailed study of the morphological traits of accessions originating from these different countries.

2.5 Tables

Table 2.1 Marker information and genetic diversity statistics (mean ± SE) for 124 rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden using 6861 SNP markers.

Pop	N	%P	N_a	N_e	I	H_e	UH_e
Denmark	23	93.8	2.117 ± 0.006	1.515 ± 0.004	0.474 ± 0.003	0.301 ± 0.002	0.308 ± 0.002
Finland	12	88.5	1.940 ± 0.005	1.484 ± 0.004	0.445 ± 0.003	0.288 ± 0.002	0.301 ± 0.002
Iceland	11	67.9	1.707 ± 0.006	1.299 ± 0.004	0.305 ± 0.003	0.191 ± 0.002	0.200 ± 0.002
Norway	28	95.3	2.063 ± 0.005	1.488 ± 0.004	0.456 ± 0.003	0.292 ± 0.002	0.297 ± 0.002
Sweden	50	99.6	2.236 ± 0.005	1.590 ± 0.004	0.535 ± 0.002	0.345 ± 0.002	0.348 ± 0.002
All	124	89.02 ± 5.57	2.012 ± 0.003	1.475 ± 0.002	0.443 ± 0.001	0.283 ± 0.001	0.291± 0.001

N = sample size; *%P* = percentage of polymorphic loci; N_a = mean number of alleles; N_e = number of effective alleles; *I* = Shannon's information index; H_e = expected heterozygosity and UH_e = unbiased expected heterozygosity.

Table 2.2 Pairwise correlation of the fixation index or F_{ST} values between subpopulations of rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden.

	DNK	FIN	ISL	NOR	SWE
DNK	0.000				
FIN	0.088	0.000			
ISL	0.133	0.124	0.000		
NOR	0.067	0.067	0.103	0.000	
SWE	0.050	0.032	0.106	0.042	0.000

F_{ST} values are below diagonal; DNK=Denmark; FIN=Finland; ISL=Iceland; NOR=Norway; and SWE=Sweden.

Table 2.3 Pairwise comparison of between and within population genetic variance of 124 rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden.

	DNK	FIN	ISL	NOR	SWE
DNK		84%	73%	86%	91%
FIN	16%		76%	88%	98%
ISL	27%	24%		79%	82%
NOR	14%	12%	21%		93%
SWE	9%	2%	18%	7%	

Values below the diagonal indicate genetic variance between populations, while those above the diagonal indicate genetic variance within populations. DNK=Denmark; FIN=Finland; ISL=Iceland; NOR=Norway; and SWE=Sweden.

Table 2.4 Inferred ancestry of 124 rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden.

Cluster	Origin of rutabaga					Total	Genetic Distance
(*K*)	DNK	FIN	ISL	NOR	SWE	Number	within subgroup
1	0	0	7	0	0	7	0.0810
2	1	0	0	0	2	3	0.4332
3	0	5	0	1	10	16	0.1533
4	0	0	0	0	3	3	0.0378
5†	4	0	0	3	4	11	0.2344
6	0	0	0	7	0	7	0.0299
7	0	0	0	1	1	2	0.0071
8	0	0	0	0	2	2	0.2642
9*	5	0	0	0	0	5	0.1917
In-Cluster	43.5%	41.7%	63.6%	42.9%	44.0%	45.2%	
Admixture♠	56.5%	58.3%	36.4%	57.1%	56.0%	54.8%	

Accessions were assigned to a specific clusters (*K*) if $P \geq 0.70$ and those that did not meet this threshold were considered as admixture. Placement of out-groups:†'Laurentian' (Canada) = *K*5; * 'Wilhemsburger' (Germany) = *K*9 and ♠ 'Krasnoselskaya' (Russia) = Admixture.

Table 2.5A Analysis of molecular variance (AMOVA) among and within 124 rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden based on genotypic distance matrix with the software GenAlEx.

Source of variance	DF	SS	MS	Estimated Variance	% Total	*P*
Among regions	2	19086.56	9543.28	59.75	1.4%	0.0214
Among populations	2	18447.25	9223.63	200.90	4.7%	0.0001
Within populations	119	477078.32	4009.06	4009.06	93.9%	0.0001

DF = Degree of freedom; SS = Sum of squares; MS = mean squared deviation. The probability is based on standard permutation across the full data set. The codom-genotypic option in GenAlEx, under the assumption of independence, was used to calculate sets of squared genetic distances summed across all loci.

Table 2.5B Analysis of molecular variance (AMOVA) among and within 124 rutabaga accessions from Denmark, Finland, Iceland, Norway and Sweden based on similarity matrix calculated with the software TASSEL.

Source of variance	DF	SS	MS	Estimated Variance	% Total	*P*
Among regions	2	0.735	0.367	0.002	1.4%	0.0201
Among populations	2	0.706	0.353	0.007	4.3%	0.0001
Within populations	119	19.176	0.161	0.161	94.3%	0.0001

DF = Degree of freedom; SS = Sum of squares; MS = mean squared deviation. The probability is based on standard permutation across the full data set. TASSEL calculates distance as 1 - IBS (identity by state) similarity, where IBS is defined as the probability that alleles drawn at random from two individuals at the same locus are the same.

Table 2.6 List of 124 rutabaga accessions from five Nordic countries (Denmark, Finland, Iceland, Norway and Sweden) included in this study of genetic diversity.

Accession	Common Name	Origin
NGB20826	AMERICAN PURPLE TOP	Denmark
NGB13795	BANGHOLM HINDERUPGAARD 9	Denmark
NGB13806	BANGHOLM HUNSBALLE 9	Denmark
NGB8372	BANGHOLM PAJBJERG REGENT	Denmark
NGB13812	BANGHOLM WIBOLTT	Denmark
NGB13813	BANGHOLM WILBY ØTOFTE	Denmark
NGB1610	DALO TRIFOLIUM	Denmark
NGB4128	DIMA TRIFOLIUM	Denmark
NGB13816	DÆHNFELDT	Denmark
NGB1611	FAMA DÆHNFELDT	Denmark
NGB1612	MAGRES PAJBJERG	Denmark
NGB13801	RECORD TASHUPGAARD	Denmark
NGB1613	RUTA ØTOFTE	Denmark
NGB8373	SAHNA PAJBJERG	Denmark
NGB1614	SATOR ØTOFTE	Denmark
NGB1615	WILBY ØTOFTE	Denmark
NGB13798	WILHELMSBUGER	Denmark
NGB13804	WILHELMSBURGER DAENO 9	Denmark
NGB8384	WILHELMSBURGER, DANILA	Denmark
NGB20827	WILHELMSBURGER DANILA TRIFOLIUM	Denmark
NGB8385	WILHELMSBURGER, REFORM	Denmark
NGB13815	WILHELMSBURGER TRIFOLIUM 9	Denmark
NGB13814	WILHELMSBURGER ØTOFTE 9	Denmark
NGB23908	BANGHOLM	Finland
NGB13791	BRYUKVA	Finland
NGB13808	BRYUKVA	Finland
NGB4408	LAITIALA AP0106	Finland
NGB4409	LYTTYLÄ AP0101	Finland
NGB24437	MESSUKYLÄ	Finland
NGB13792	MUSTIALA	Finland
NGB24436	SIMO (SIMO 1)	Finland
NGB14152	SIMO (SIMO 2)	Finland
NGB20823	SORTAVALA	Finland
NGB13805	TAMMISTO	Finland
NGB1176	VILLALA ME0101	Finland
NGB9915	HNAUSAROFA	Iceland
NGB9911	HVAMMSROFA	Iceland
NGB20824	KAFIFAFELLSROFUR	Iceland
NGB4142	KALFAFELLSROFA	Iceland
NGB9909	KOLSHOLTSHELLISROFA	Iceland
NGB9916	KORPUROFA	Iceland
NGB9914	LAUGABOLSROFA	Iceland
NGB9910	MARIUBAKKAROFA	Iceland
NGB9913	MÖGGUROFA	Iceland
NGB9907	SLETTUROFA	Iceland
NGB13811	VESTMANNAEYRA	Iceland

NGB13800	BANGHOLM ELITE	Norway
NGB7793	BANGHOLM GOKSTAD	Norway
NGB2657	BANGHOLM HAUKEBØ ST	Norway
NGB4568	BANGHOLM SANDNES	Norway
NGB4567	BANGHOLM VEREIDE	Norway
NGB10657	BANGHOLM WILBY ØTOFTE II	Norway
NGB5015	BRANDHAUG	Norway
NGB522	BRANDHAUG, MARKA STAMME	Norway
NGB5016	BREDIK	Norway
NGB7792	BRENDBERGLI	Norway
NGB13794	GRO	Norway
NGB11559	GRY	Norway
NGB9268	GØTA LEDAAL STAMME	Norway
NGB11558	KVIMAR	Norway
NGB7794	OLSGÅRD ELITE	Norway
NGB7795	RANAKÅLROT	Norway
NGB10000	REDY	Norway
NGB523	SALTA	Norway
NGB10656	SIMONETTE KVANDE	Norway
NGB5017	STENHAUG	Norway
NGB4133	STENHAUG	Norway
NGB9274	TRØNDERSK BRANDHAUG	Norway
NGB10659	TRØNDERSK HYLLA	Norway
NGB4569	TRØNDERSK KVITHAMAR	Norway
NGB11560	VALLDALSKÅLROT	Norway
NGB5018	VIGE	Norway
NGB4134	VIGE	Norway
NGB24906	VIGOD	Norway
NGB21723	BAGGENS	Sweden
NGB13793	BANGHOLM FENIX	Sweden
NGB7175	BANGHOLM FENIX	Sweden
NGB13797	BANGHOLM ORIGINAL	Sweden
NGB17916	Bjursås (Bortas Bertils kålrot)	Sweden
NGB13474	BJURSÅS	Sweden
NGB17910	BOLTJÄRN	Sweden
NGB13799	DROTTNING	Sweden
NGB17918	FARFAR	Sweden
NGB11748	FINNMARKENS VITA KÅLRÖTTER	Sweden
NGB13637	GLOBUS	Sweden
NGB13120	GLOBUS	Sweden
NGB13119	GULLÅKER III	Sweden
NGB11688	GULLÅKER III	Sweden
NGB13796	GUL SVENSK	Sweden
NGB13802	GUL SVENSK	Sweden
NGB13663	GUL SVENSK (from SESAM)	Sweden
NGB13818	GUL SVENSK SVALÖFS ORIGINAL	Sweden
NGB17914	JANNE	Sweden
NGB21676	KLINT KARIN	Sweden
NGB11742	KÅLROT FROM ASPÅS	Sweden
NGB11744	KÅLROT FROM TÄLLBERG	Sweden

NGB17905	LJUSTORP	Sweden
NGB13770	MARIEROKÅLROT	Sweden
NGB18050	NORJÖ	Sweden
NGB17909	NUSNÄS	Sweden
NGB13790	ODAL	Sweden
NGB13809	PATRIA	Sweden
NGB6791	ROTABAGGE, KÅLROT	Sweden
NGB17915	SOLLROT	Sweden
NGB20825	SVALÖF VICTORIA	Sweden
NGB13807	SVENSK	Sweden
NGB11689	TIFFANY	Sweden
NGB17908	TRUTSGÅRD	Sweden
NGB20822	WEIBULLS FODDER	Sweden
NGB23910	WEIBULLS ORGINAL BALDER	Sweden
NGB13810	WEIBULLS ORIGINAL BANGHOLM	Sweden
NGB17906	VIKSJÖ	Sweden
NGB7176	VIKTORIA	Sweden
NGB17911	VINTJÄRN	Sweden
NGB17913	ÖSTERGYLLEN	Sweden
NGB13803	ÖSTGÖTA	Sweden
NGB23909	ÖSTGÖTA II	Sweden
NGB13638	ÖSTGÖTA II	Sweden
NGB10658	ÖSTGÖTA II	Sweden
NGB13118	ÖSTGÖTA II	Sweden
NGB11164	ÖSTGÖTA II	Sweden
NGB17917	ÖSTNOR	Sweden
NGB10667	DELTA	Sweden
NGB10668	SIGMA	Sweden

2.7 Figures

Figure 2.1 The origin of 124 rutabaga accessions used in this study of genetic diversity. The Nordic region (Norway, Sweden, Finland, Denmark and Iceland) is often cited as the center of domestication and diversity of rutabaga. The map was retrieved from ArcGis Online (https://www.arcgis.com/) on August 30th, 2018.

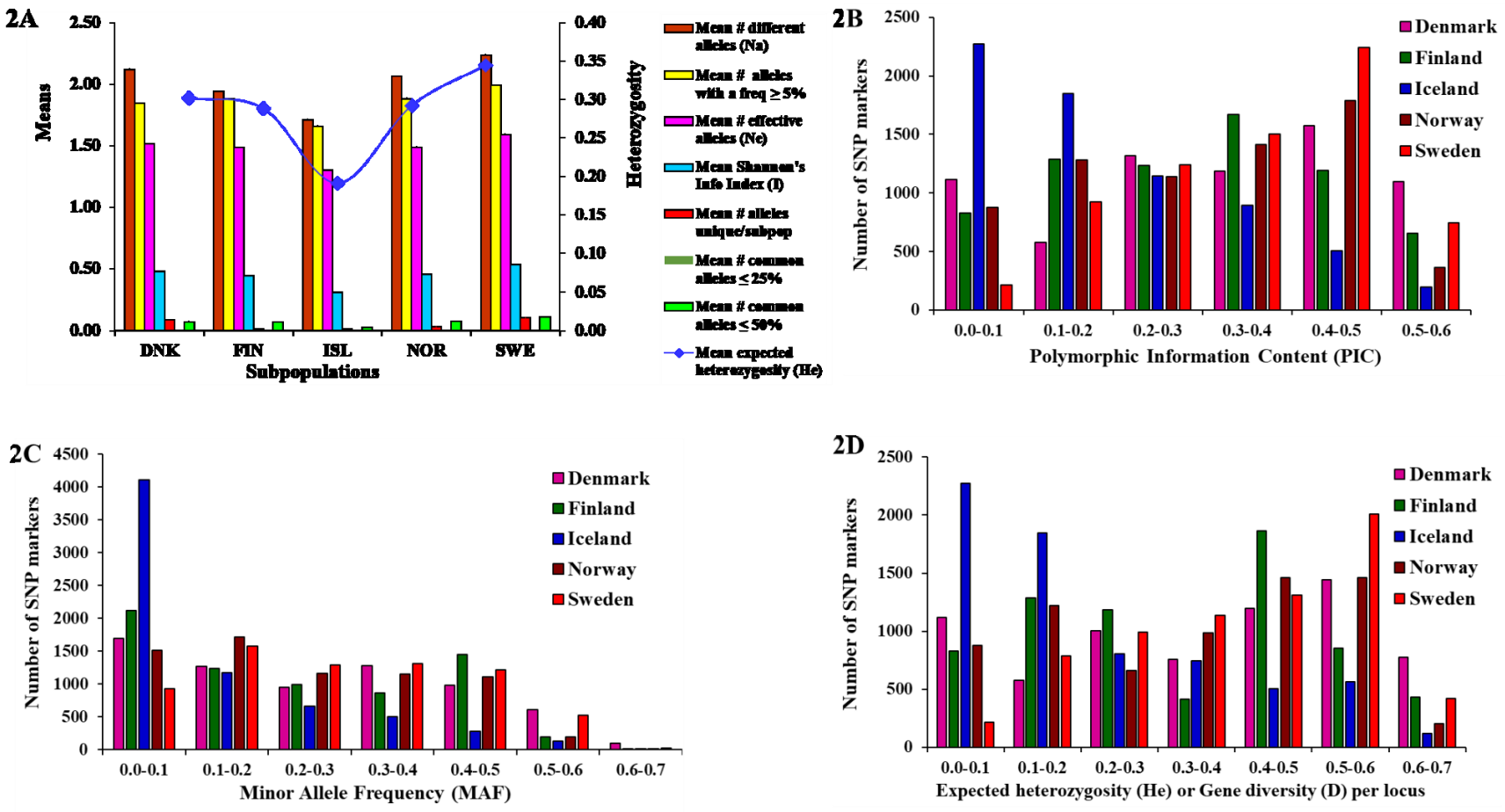

Figure 2.2 Distribution of allele frequency-based genetic diversity statistics **(A)**, Polymorphic Information Content (PIC) **(B)**, Minor Allele Frequency (MAF) **(C)**, and Expected heterozygosity (He) or gene diversity **(D)** of 6861 SNP markers across 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland.

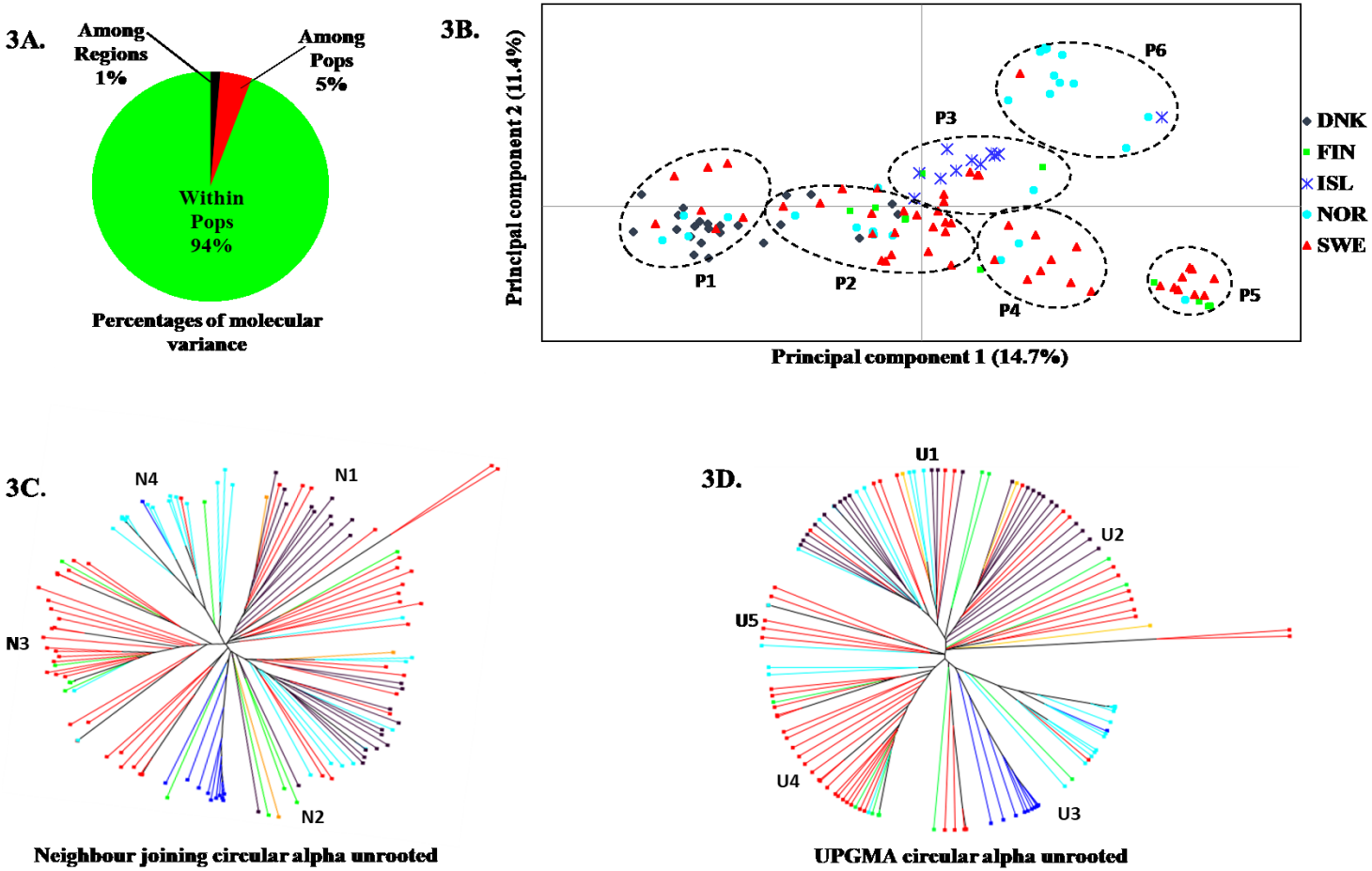

Figure 2.3 Analysis of molecular variance (AMOVA) partitioning of molecular variance among regions, populations and within accessions **(A)**. Principal coordinates analysis (PCoA) **(B)**, Neighbour joining (NJ) **(C)**, and Unweighted pair group method with arithmetic mean (UPGMA) **(D)** analyses with 6861 SNP markers grouped the 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland into 6, 4 and 5 subgroups, respectively.

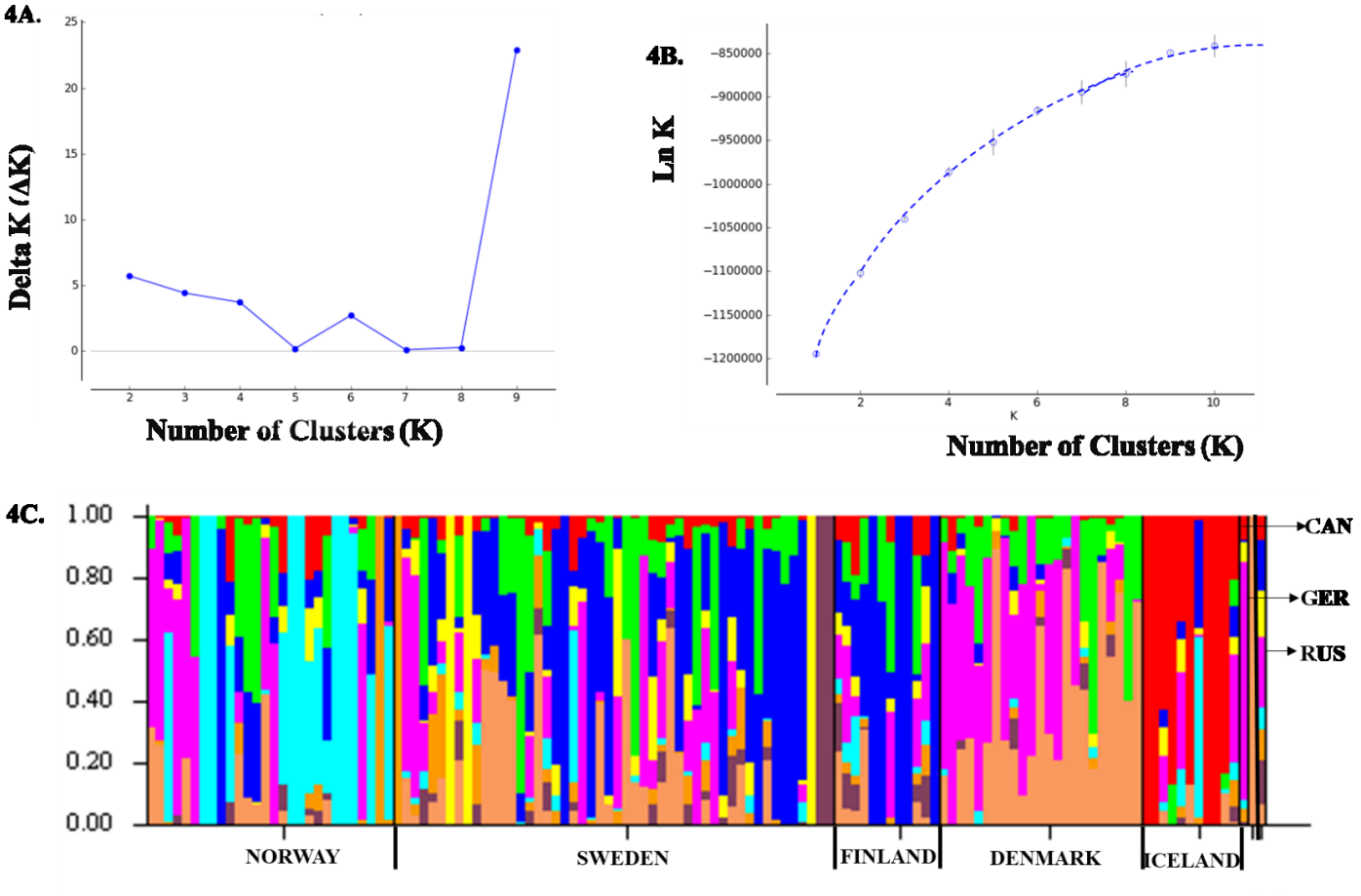

Figure 2.4 The optimal value of K determined by the Evanno et al. (2005) method suggested that the 124 rutabaga accessions could be placed into 9 clusters (*K*=9) **(A)**. Gradual variation of log likelihood with increase of *K* **(B)**, Population structure analysis of 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland. The column represents the individual rutabaga accessions used in this study while each colour represents one gene pool and the stacked bars with different colours represent admixtures with their shared ancestry components **(C).**

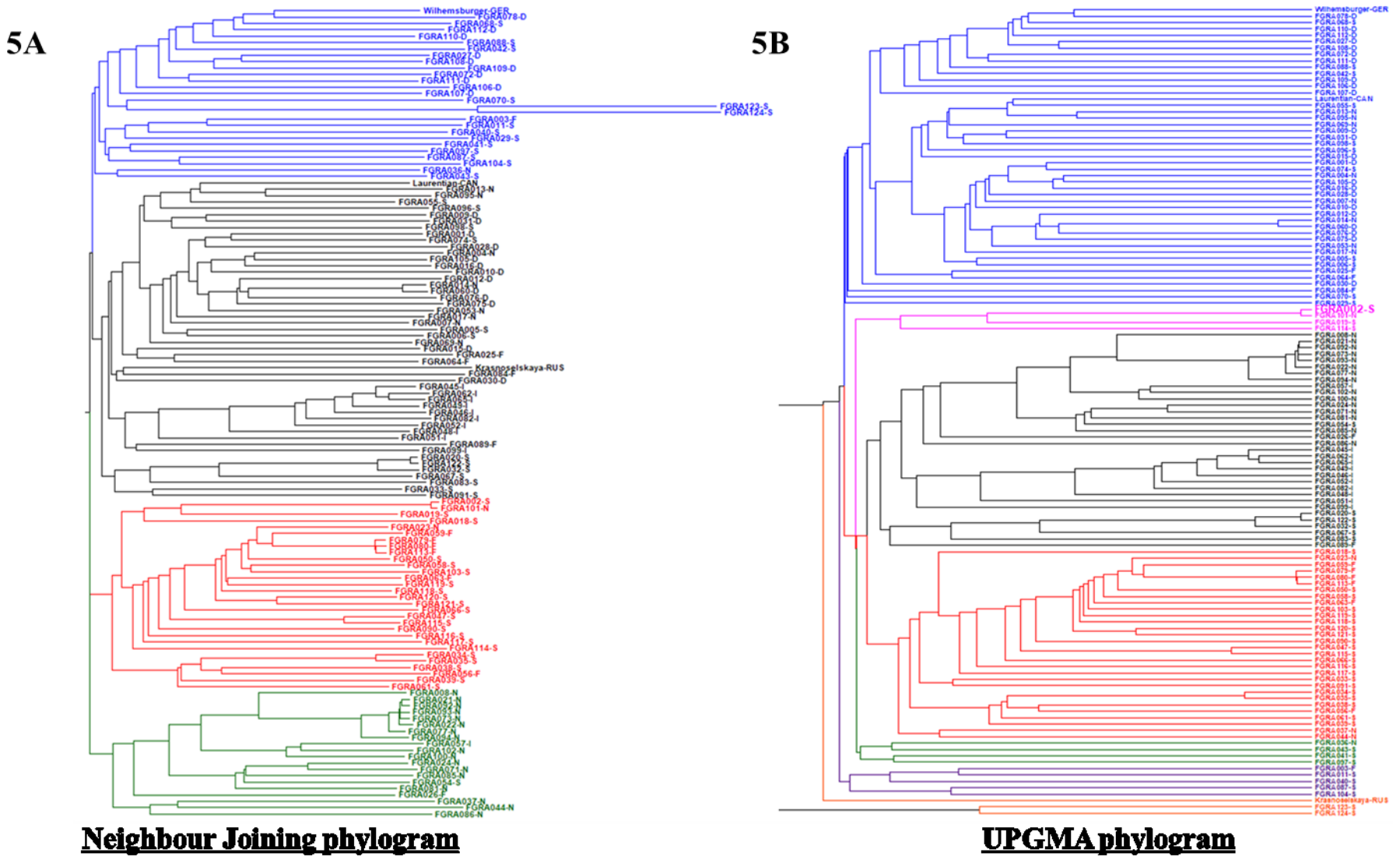

Figure 2.5 Neighbour-joining (NJ) **(A)** and Unweighted pair group method with arithmetic mean (UPGMA) **(B)** phylograms with 6861 SNP markers grouped the 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland into 4 and 5 subgroups, respectively.

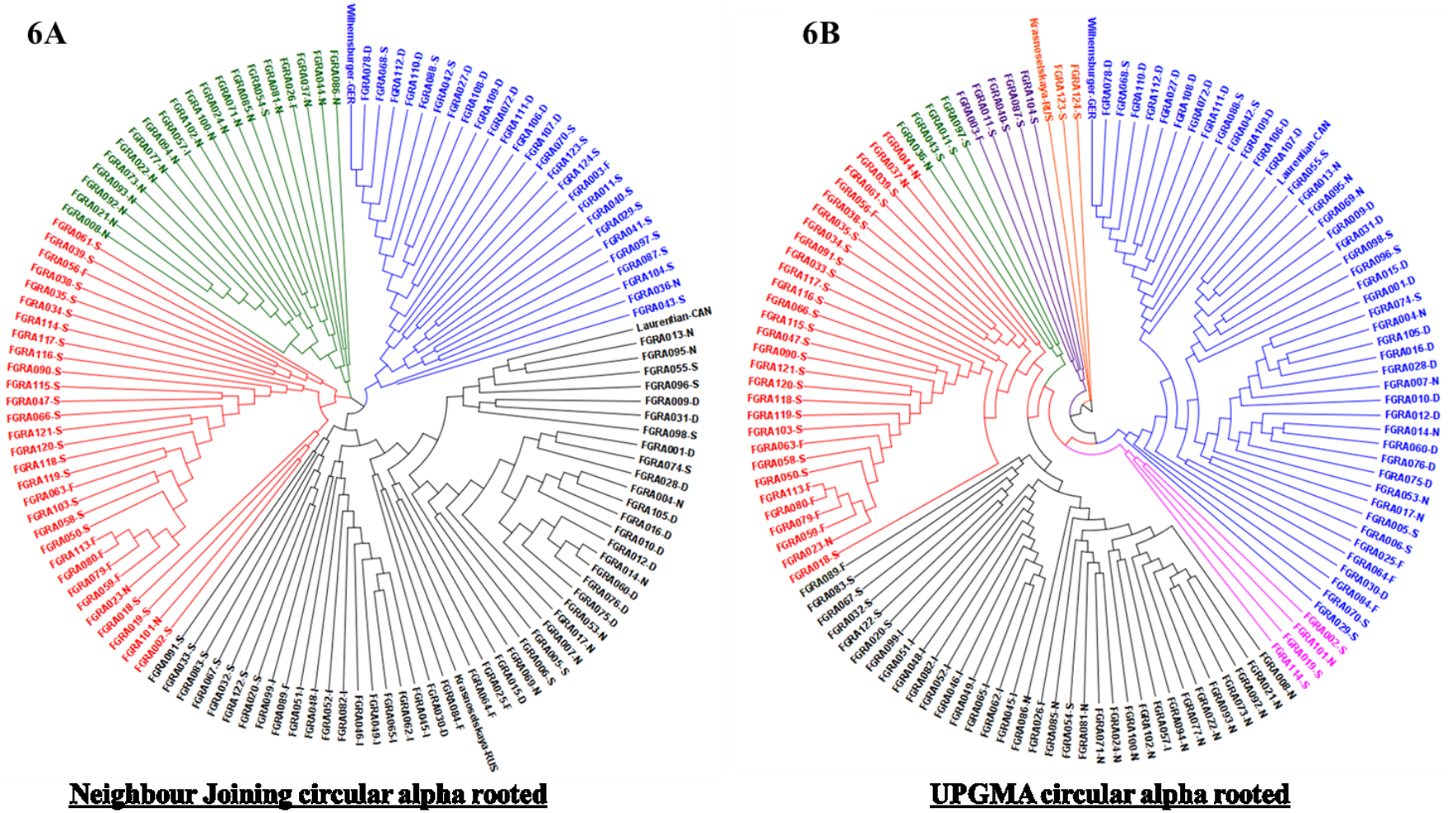

Figure 2.6 Neighbour-joining (NJ) **(A)** and Unweighted pair group method with arithmetic mean (UPGMA) **(B)** circular alpha rooted cladogram with 6861 SNP markers grouped the 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland into 4 and 5 subgroups; respectively.

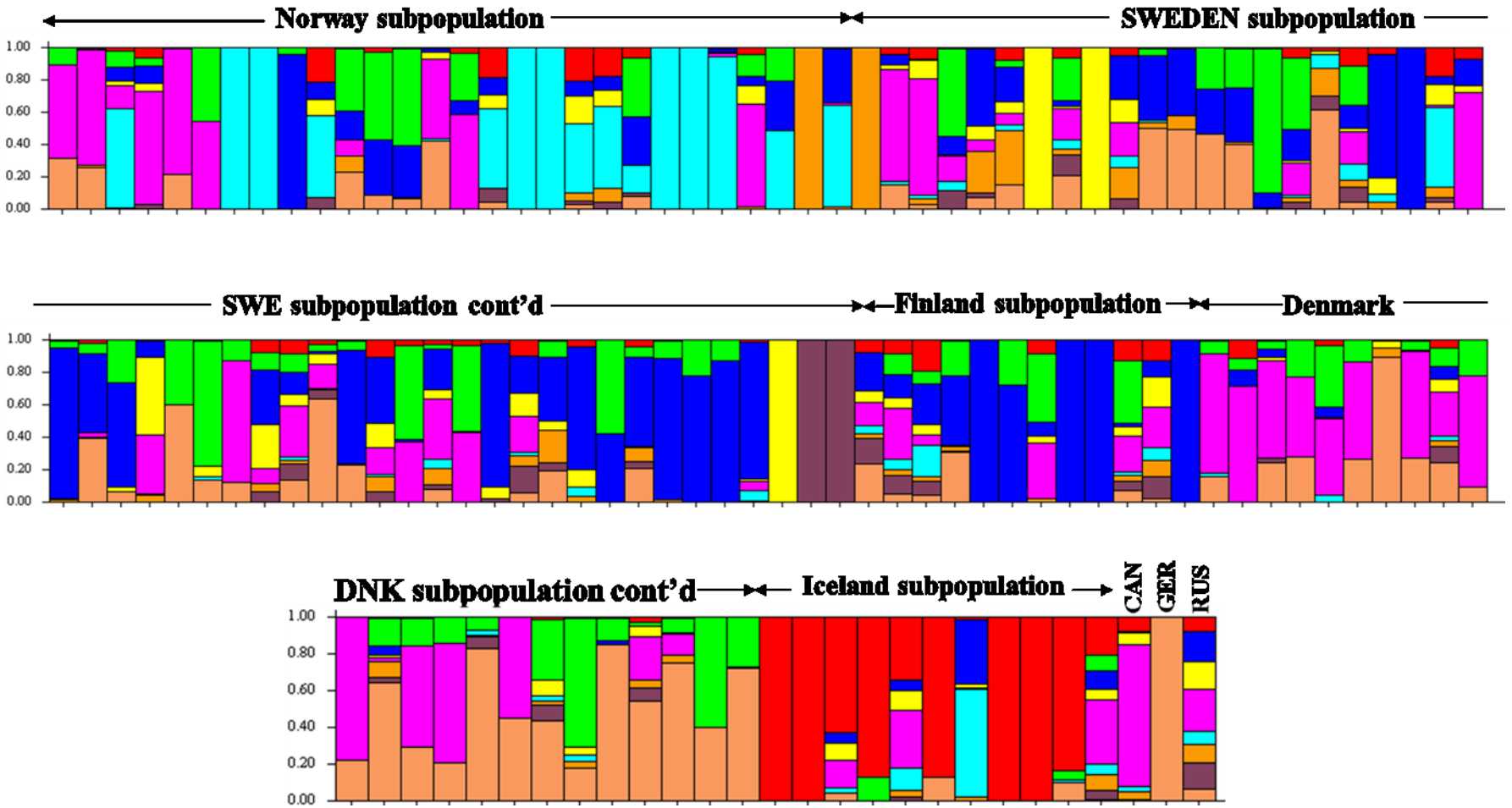

Figure 2.7 Detailed Bayesian clustering of 124 rutabaga accessions from Norway, Sweden, Finland, Denmark and Iceland**.** Each accession is represented by a rectangular bar.

Chapter 3: Genome-wide association mapping of clubroot resistance loci in rutabaga

3.1 Introduction

Rutabaga (*Brassica napus* spp. *napobrassica*) has been cultivated commercially as a vegetable crop in Canada since the 1950s and 1960s (Shattuck and Proudfoot 1990; Spaner 2002). This crop is affected by a number of diseases including clubroot of crucifers, caused by the obligate parasite *Plasmodiophora brassicae*. Clubroot is best managed by the use of host resistance (Rahman et al. 2014). The European winter oilseed rape (syn. canola; *B. napus*) 'Mendel' has been the main source of clubroot resistance gene(s) in many canola breeding programs in Canada (Fredua-Agyeman et al. 2018), but 'new' pathotypes of *P. brassicae* capable of overcoming this resistance in clubroot resistant (CR) canola cultivars have emerged recently in Alberta (Strelkov et al. 2016; Strelkov et al. 2018). Indeed, the *CRa*/*CRb*^{Kato} resistance gene(s) on the A03 chromosome of 'Mendel' were found to confer resistance to only about 50% of the new *P. brassicae* pathotypes identified in Alberta from 2013 to 2015 (Fredua-Agyeman et al. 2018). As such, efforts have intensified to identify new clubroot resistance genes from other Brassica sources to develop the next generation of CR canola cultivars.

Greenhouse resistance screening showed that rutabagas possess both qualitative and quantitative resistance to isolates representing the 'old' and 'new' pathotypes of *P. brassicae* (Hassan et al. 2012, Fredua-Agyeman et al. 2019). Therefore, rutabagas can be used in the breeding of CR canola cultivars. Examples of CR rutabaga cultivars developed in Canada include 'Chignecto', 'York', 'Fortune', 'Kingston' and 'Brookfield' (Shattuck and Proudfoot 1990). The Canola Breeding Program at the University of Alberta mapped the *Crr1* clubroot resistance gene in 'Brookfield' to the A08 chromosome of *B. rapa* and *B. napus* (Hassan and Rahman 2016). While rutabaga is a root vegetable and canola is an oilseed, they are both *B. napus* and share the same ploidy level and genome ($2n = 19$, AACC). Therefore, the introgression of clubroot resistance genes from rutabaga to

canola does not face many of the pre- and post-zygotic challenges associated with crosses with more distant relatives. On the other hand, many cycles of backcrossing and breeding will be needed to achieve canola quality (spring growth characteristics, plant architecture, reductions in the high content of erucic acid in the oil and glucosinolate in the seed meal, as well as days to flower). In addition, molecular markers linked to the clubroot resistance genes in rutabagas need to be identified for marker-assisted selection (MAS) to track the introgressed genes.

In plants, linkage-based mapping is often carried out on populations developed from bi-parental crosses, and so polymorphism is restricted to the contrasting genetic variability for the trait of interest in the parents (Brachi et al. 2011; Gupta et al. 2014). As such, markers detected for quantitative traits are restricted to the two parents and may not work in populations developed from other parents. An advantage of linkage-based QTL mapping is the often very high power of QTL detection (Gupta et al. 2014). Linkage disequilibrium (LD)-based association mapping (AM), such as genome-wide association mapping (GWAS), is an alternate approach for capturing recombination events to the gene level in natural populations, and has several advantages over traditional linkage-based mapping (Gupta et al. 2014). GWAS can be applied to a diverse set of genotypes of any crop species, does not require ancestry or pedigree information for QTL mapping, captures higher allelic diversity, provides higher resolution, can be used to study many traits of interest, and is cheaper and faster to complete since it does not require the development of a mapping population. However, LD-based AM is often limited by the detection of false positive associations (Type I error) that may arise as the result of linkage, population stratification and relatedness among individuals (Gupta et al. 2014). In addition, SNPs with minor alleles (< 5 to 10%) are filtered off during GWAS analyses, and hence GWAS lacks the power to detect these minor alleles (Brachi et al. 2011).

The objectives of this study were to phenotype and genotype rutabaga accessions collected from Norway, Sweden, Finland, Denmark and Iceland, and to use GWAS to identify SNP and SSR

markers significantly associated with clubroot resistance to a collection of *P. brassicae* isolates from Alberta, Canada, representing different pathotypes. In addition, the positions of the identified markers, as well as those from previous genetic mapping studies, were traced to the *B*. *rapa*, *B. oleracea*, *B. napus* reference genomes. Lastly, candidate genes associated with the identified genomic regions were identified based on the proteins they encode.

3.2 Materials and Methods

3.2.1 Plant Material

The materials used in this study comprised all 124 *B. napus* ssp. *napobrassica* (rutabaga) accessions used for the genetic diversity studies of Chapter 2. Seeds of the rutabaga were multiplied by planting 2-4 seeds of each accession in 13 × 13 × 15 cm pots filled with Sunshine Mix #4 Aggregate Plus growing mix (Sungro Horticulture, Seba Beach, Alberta, Canada), and placing the pots in a greenhouse at the Crop Diversification Centre North (CDCN), Alberta Agriculture and Forestry (AAF), Edmonton, Canada. After one week, the seedlings were thinned to two plants per pot and kept in the greenhouse for another 5 - 6 weeks at 20-25/15-18ºC day/night temperatures and a 16 h light/8 h dark photoperiod. The plants were then vernalized for 12 weeks in a cold room maintained at 4ºC with a 12 h photoperiod, and subsequently returned to the greenhouse for flowering, maturation and seed harvest.

3.2.2 *Plasmodiophora brassicae* pathotypes

The *P. brassicae* pathotypes used in the clubroot tests comprised 12 of the field isolates reported by Strelkov et al. (2016 and 2018) to cause elevated disease on clubroot resistant (CR) canola and five single-spore isolates (SSIs) identified by Xue et al. (2008) prior to the introduction of CR canola to Western Canada. The 12 virulent isolates (L-G2 + L-G3, D-G3, F183-14, F3-14, F175-14, CDCN#6,

F187-14, F10-15, F12-15, F381-16/C.C. and UofA/County#37) were classified according to the Canadian Clubroot Differential (CCD) Set as pathotypes 5X, 5L, 2B, 3A, 5C, 5G, 8E, 5K, 8J, 3O and 8P, respectively (Strelkov et al. 2018). Two of the field populations LG2 and LG3 (Strelkov et al. 2016) represented the same pathotype (5X). Therefore, the 12 virulent isolates represented 11 pathotypes.

The five SSIs were classified as pathotypes 2, 3, 5, 6 and 8 on the differentials of Williams (1966) and as pathotypes 2F, 3H, 5I, 6M and 8N, respectively on the CCD Set (Strelkov et al. 2018). The numbers in the CCD designations correspond to the classification according to the differentials of Williams, while the letters designate the unique virulence patterns on the hosts of the CCD. The 17 isolates (representing isolates of 16 pathotypes) used in this study were maintained as galls on the universally susceptible European Clubroot Differential (ECD) 05 (Buczacki et al. 1975) at -20ºC.

3.2.3 Clubroot Phenotyping

Greenhouse clubroot tests were conducted at the CDCN following Fredua-Agyeman et al. (2019). Briefly, 4-8 seeds of each rutabaga accession were placed on Whatman No. 1 filter paper wetted with distilled water in Petri dishes, and kept at room temperature (18 - 22ºC) and a 12 h light/12 h dark photoperiod provided by white fluorescent light. The filter paper was moistened daily with water for 7 days, at which point the seeds had germinated. The universally susceptible *B. napus* (canola) cv. 'Westar' was included in each experiment as a positive control.

Inoculum of isolates representing each *P. brassicae* pathotype (2F, 3H, 5I, 6M, 8N, 5X (LG2 and LG3) , 5L, 2B, 3A, 5C, 5G, 8E, 5K, 8J, 3O and 8P) was prepared by macerating the frozen galls with a Waring LB10G variable-speed blender (Cole-Palmer) and filtering the resulting homogenate through two layers of cheesecloth. The resting spore concentration in the filtrate was measured with a hemocytometer and adjusted to 1.0×10^7 resting spores mL^{-1} with water. Seven-day old seedlings of

each of the 124 rutabaga accessions were inoculated by dipping the roots in the spore suspension as described by Nieuwhof and Wiering (1961).

The inoculated seedlings were transplanted into 13 cm ×13 cm ×15 cm pots filled with Sunshine Mix #4 Aggregate Plus soilless mix (Sungro Horticulture Canada Ltd., Seba Beach, Canada), followed by the addition of 1 mL of inoculum into the potting mix at the base of each plant (Lamers and Toxopeus, 1977 as cited by Voorrips and Visser, 1933). The inoculated rutabaga seedlings were planted at the periphery of each pot while one seedling of the susceptible check 'Westar' was placed at the centre. Pots inoculated with the same isolate were placed randomly on one bench and re-randomized twice on the 3rd and 6th weeks after transplanting.

The experiments were repeated 3-4 times based on seed availability. Maintenance of the plants (watering, fertilizing and insect pest control) and growing conditions (photoperiod and temperature) in the greenhouse were as described by Fredua-Agyeman et al. (2019).

3.2.4 Clubroot disease rating

The severity of clubroot was assessed 8 weeks after inoculation on a 0-3 scale (Kuginuki et al., 1999) where: 0 = no galling, 1 = a few small galls on the lateral roots, 2 = moderate galling on the lateral roots but not on the main root, and 3 = severe galling on both the lateral and main root. Ratings for the rutabaga accessions in each pot were considered valid only if the disease rating of the susceptible check 'Westar' in the same pot was 2 or 3. An index of disease (ID, 0-100%) was calculated for each isolate/host genotype combination following Strelkov et al. (2006), with the mean IDs and standard deviations (SDs) of the repeated experiments determined for all 124 rutabaga accessions for each of the 17 *P. brassicae* isolates.

The rutabaga accessions were designated as resistant (Mean ID + SD ≤ 30%), moderately resistant (30% < Mean ID + SD ≤ 50%) or susceptible (Mean ID + SD > 50%) to isolates representing each pathotype based on Fredua-Agyeman et al. (2019) and the clubroot resistance screening guidelines of the Western Canada Canola/Rapeseed Recommending Committee (WCC/RCC). The grand mean (GM) IDs (means of the means of IDs with all 17 isolates) of each accession were also calculated and used as an indicator of broad-spectrum resistance.

3.2.5 DNA extraction

DNA was extracted from ca. 0.25 g leaf tissue of each accession using the cetyltrimethyl ammonium bromide (CTAB) method (Sambrook and Russell 2001). The DNA concentration was measured with a ND-2000C spectrophotometer (NanoDrop, Technologies, Inc., Wilmington, DE, USA) and the concentration diluted to 10-20 ng μL^{-1} for the working solution each DNA.

3.2.6 SNP genotyping

SNP genotyping was performed using the *Brassica* array 15K SNP at TraitGenetics GmbH, Gatersleben, Germany, according to the manufacturer's protocols (Clarke et al. 2016). After removing monomorphic, low coverage site markers, markers with MAF ≤ 0.05 and those missing data for > 5% of the accessions, 6861 SNP markers, comprising 4390 A-genome and 2471 C- genome markers, were used for GWAS analyses.

3.2.7 PCR and genotyping with PCR-based markers linked to known CR genes

PCR reactions, genotyping, amplified product detection and calling of allele sizes were performed as described by Fredua-Agyeman et al. (2014). The genotyping was carried out with 63 PCR-based primers linked to nine previously identified clubroot resistance genes. Primers screened from the A03 chromosome of *B. rapa* consisted of the following: six SCAR markers linked to the *CRa* gene (Ueno

et al. 2012), three SSR, four SCAR and one CAP marker(s) linked to the *CRb* gene (Piao et al. 2004; Zhang et al. 2014), 32 SSR and three InDel markers linked to the CRb^{Kato} gene (Kato et al. 2012; 2013), two SSR and one STS markers linked to the *Crr3* gene (Saito et al. 2006; Fredua-Agyeman and Rahman 2016), one SSR, STS and SCAR marker each reported to be linked to the *CRk* gene (Matsumoto et al. 2012).

Primers screened from other chromosomes comprised: three SSR markers on the A08 chromosome of *B. rapa* linked to the *Crr1* gene (Suwabe et al. 2006; Hassan and Rahman 2016; Hobson and Rahman 2016), one SSR marker on the A01 chromosome linked to the *Crr2* gene, two SCAR and one STS markers on the A02 chromosome linked to the *CRc* gene (Sakamoto et al. 2008; Matsumoto et al. 2012) and one SSR marker on the A06 chromosome linked to the *Crr4* gene (Suwabe et al. 2006). The primer sequences and chromosomal locations are listed in Table 3.2. Each allele was coded as a separate site with each allele in turn recoded as one of two nucleotides (A or T; C or G) as specified in the TASSEL Manual (Bradbury et al. 2007).

Genotyping was repeated for 10% of the individual samples to confirm the reproducibility of the molecular data (Fredua-Agyeman et al. 2014). In addition, alleles with a frequency of ≤ 5% were removed to reduce false positive associations (Nie et al. 2016). One hundred and eight alleles linked to known CR genes were used for the GWAS analyses.

3.2.8 Marker-clubroot association

To identify loci in the rutabaga accessions associated with the response to each of the 17 *P. brassicae* isolates, genome-wide association analyses were conducted with TASSEL 5.0 (Bradbury et al. 2007) using the 6861 SNP marker data and the mean ID values of each accession and pathotype. A separate analysis was performed using the 108 alleles derived from the 63 PCR-based markers and the mean ID data from the clubroot phenotyping experiments.

Marker-trait association analyses were carried out following Li et al. (2014) with slight modifications. Four models each were tested with the general linear models (GLM) and mixed linear models (MLM) procedures implemented in TASSEL 5.0 (Bradbury et al. 2007). The GLM tested comprised the N (Naïve), Q-only (population structure), K-only (Kinship) and PCA-only (Principal Component Analysis) models. The MLM tested comprised the Q + K and PCA + K models (Li et al. 2014) and two additional MLM association tests of Q and PCA using Distance matrices (D) (i.e., Q + D and PCA +D).

The best models were identified from Quantile-Quantile (Q-Q) plots, which plot the $-\log_{10} P$ value of the test of association (observed) with that expected given the null hypothesis of no marker-trait associations. Markers with strong associations to clubroot resistance were identified from the best fitting Q-Q plots and also from peaks on Manhattan plots. The Bonferroni correction was used to set the significance cut-off (α/n, where α = level of significance, n = number of markers) for both the SNP and the PCR-based markers. The SNP markers were considered to be significantly associated with the traits if $P \leq 1 \times 10^{-4}$ (i.e. $-\log_{10} P = 4.0$). In the case of the PCR-based markers, a slightly lower threshold of $P \leq 5 \times 10^{-4}$ (i.e. $-\log_{10} P = 3.3$) was used to indicate the significance of associations between the markers and the traits.

The physical positions of the SNP and the SSR markers with strong associations to clubroot resistance were then mapped to the *B. rapa* and *B. oleracea* genome sequences to identify their association with previously characterized clubroot resistance genes.

3.2.9 Candidate gene prediction

The sequences of the PCR-based and SNP markers found to be associated with clubroot resistance loci were used in BlastN searches of the *B. rapa* (AA), *B. oleracea* (CC), *B. napus* (AACC) and *Arabidopsis thaliana* genome sequences available in the NCBI (www.ncbi.nlm.nih.gov) GenBank

database to determine their possible functions. An E-value ≤ E-20 and a percentage identity of ≥ 95% were used as the threshold cut-offs to confirm the putative functions of the candidate genes.

3.2.10 Statistical analyses

Statistical analyses of the phenotype data were conducted with SAS v. 9.4 (SAS Institute, USA). The PROC SORT function was used to select the R and MR (i.e. ID ± SD ≤ 50%) accessions. Duncan's test (Steel and Torrie, 1960) was used to test ($P \leq 0.05$) for differences among the ID mean values of all 17 isolates and to quantify these differences among the rutabaga accessions. The distribution of the IDs for each R and MR rutabaga accession were visualized with box-and-whisker plots.

3.3 Results

3.3.1 Phenotypic variation of clubroot resistance in rutabaga accessions

The frequency distribution of 124 rutabaga accessions evaluated for resistance to 17 *P. brassicae* isolates (representing 16 pathotypes) is shown in Figure 3.1. Between 0.8-10.4% of the rutabaga accessions were resistant (R) and 4.0-12.0% were moderately resistant (MR) to 14 of the 17 *P. brassicae* isolates. The 14 isolates represented 13 pathotypes and this included all five SSIs (representing pathotypes 2F, 3H, 5I, 6M, and 8N) and 9 of the 12 field isolates (pathotypes 2B, 3A, 3O, 5C, 5G, 8E, 8P, and 5X (L-G2 and L-G3)). Among these, isolates representing pathotype 5C appeared to be the most virulent on the rutabaga accessions, since only 0.8% (one of the 124 accessions) was resistant. In contrast, greater percentages (33.6%, 62.4% and 66.4%) of the rutabaga accessions were resistant or moderately resistant to isolates representing the remaining three pathotypes (8J, 5L and 5K, respectively) (Figure 3.1). There was no significance difference in the percentages of the rutabaga accessions R (5.6-8.8%), MR (5.6-9.6%) and S (84.0-84.8%) to the L-G2 and L-G3 isolates representing pathotype 5X.

Resistance and moderate resistance were consistent among nine accessions (Figure 3.2). Accession FGRA107 (GM ID = 22.7%) was resistant or moderately resistant to isolates representing all pathotypes (i.e., R to 11 and MR to 6 isolates). Three accessions FGRA037 (GM ID = 26.1%) FGRA109 (GM ID = 27.6%) and FGRA073 (GM ID = 33.3%) were resistant or moderately resistant to isolates representing 15 of the 16 pathotypes (i.e., R + MR to 16 of the 17 isolates) but were each susceptible to one pathotype; i.e., 3O, 8N and 8N, respectively. Two accessions, FGRA069 (GM ID = 31.0%) and FGRA044 (GM ID = 32.1%) were resistant to isolates representing 14 of the 16 pathotypes (i.e., R + MR to 15 of the 17 isolates) but susceptible to two pathotypes; 6M and 5C and 2F and 5X (LG2), respectively. Another three accessions FGRA036 (GM ID = 44.5%), FGRA113 (GM ID = 45.0%) and FGRA110 (GM ID = 47.7%) exhibited considerable resistance or moderately resistance (i.e., R + MR to 9 -11 of 17 isolates).

Overall, 12.1% of the rutabaga accessions were resistant or moderately resistant to ≥ 8 of the isolates, while the vast majority (87.9%) exhibited very little resistance (Figure 3.1). The latter comprised 33 (26.6%) accessions susceptible to all 17 *P. brassicae* isolates, and 76 (61.3%) accessions that were resistant or moderately resistant to isolates representing 1-7 pathotypes (data not shown). Therefore, the rutabaga accessions showed a wide variation in their reaction to isolates representing all the *P. brassicae* pathotypes used in this study. Resistance in the accessions with broad spectrum resistance was of the order FGRA107 > FGRA037 > FGRA109, FGRA069, FGRA044, FGRA073 > FGRA036, FGRA113 > FGRA110.

3.3.2 Association mapping of clubroot resistance loci

Analyses of the SNP marker genotype data with TASSEL 5.0 (Bradbury et al. 2007) detected population structure (Q) and kinship (K) among the 124 rutabaga accessions. Based on the Q-Q plots of the four GLM models, the deviation of the observed $-\log_{10} P$ distribution from the expected

distribution was of the order Naïve > K-only > PCA-only and Q-only for both the PCR-based (data not shown) and SNP markers (Figure 3.3a to 3.3d). All four MLM (Q + K, PCA + K, Q + K and PCA + D) models (Figure 3.3e to 3.3h) gave a minimal deviance of the observed $-\log_{10} P$ from the expected distribution compared with the aforementioned GLM models. In addition, the MLM models that utilized Kinship matrix (i.e., PCA + K and Q + K) departed the least from the expected distribution compared with the MLM models that utilized the Distance matrix (i.e., PCA + D and Q + D). Therefore, based on the Q-Q plots of the eight GWAS models, the PCA + K and Q + K MLM models were used for SNP-clubroot association analyses.

Forty-three (43) SNP markers were detected by the PCA + K model to be significantly associated with resistance to 11 *P. brassicae* pathotypes. This comprised 4, 13, 3, 4, 6, 2, 1, 3, 1, 5 and 1 SNPs which were associated with resistance to pathotypes 2F, 3H, 5I, 6M, 8N, 2B, 3A, 3O, 5C, 5K and 8P, respectively (Table 3.1). The SNP marker Bn_A01_p161237 was significantly associated with resistance to pathotypes 3H and 6M, while the remaining 42 were associated with resistance to only one pathotype each. Thirteen (13) SNP markers were found by the Q + K model to be significantly associated with resistance to six *P. brassicae* pathotypes (Table 3.1). This comprised one SNP marker each which was associated with resistance to pathotypes 5C and 5K, and 10 SNPs which were associated with resistance to pathotype 3H. The last SNP marker, Bn_scaff_18338_1_p871455, was significantly associated with resistance to pathotypes 2F, 3A and 5I. Eleven of the 13 SNPs were detected both by the Q + K and PCA + K models, while two SNPs (Bn_A03_p21487106 and Bn_A05_p3191390) were detected only by the Q + K model (Table 3.1).

Only the PCA + K model was used in the case of the GWAS with the PCR-based markers. Two SSR markers KB29N19 and B0903 (Kato et al. 2012) on the *B. rapa* chromosome A03 were significantly associated with resistance to eight (3H, 6M, 8N, 2B, 3A, 5C, 5G and 5I) and 3 pathotypes (5C, 5L and 5K), respectively. The InDel marker B1005 (Kato et al. 2012), also on the

A03 chromosome, was associated with resistance to pathotype 5K. In addition, the SSR marker A08-5021 (Hobson and Rahman 2016) on the A08 chromosome was significantly associated with resistance to pathotype 5L. Therefore, a total of 45 (32 by PCA + K only, 2 by Q + K only and 11 by both PCA +K and Q + K) SNP markers and 4 PCR-based markers were detected to be significantly associated with resistance loci to the five 'old' and eight of the 12 'new' *P. brassicae* pathotypes used in this study. None of the molecular markers were found to be associated with resistance to pathotypes 5X (L-G2 and L-G3), 8E or 8J.

About 85% of the markers identified mapped to the A-genome of *B. rapa* and *B. napus*. The highest number of markers (22%) mapped to the A03 chromosome, where the *Crr3* (Hirai et al. 2004; Saito et al. 2006), *CRk* (Matsumoto et al. 2012), *CRd* (Pang et al. 2018), *CRa* (Matsumoto et al. 1998, 2012), *CRb* (Piao et al. 2004, Zhang et al. 2014), CRb^{Kato} (Kato et al. 2012, 2013), *Rcr1* (Chu et al. 2014; Yu et al. 2016) and *BraA.CR.a* (Hirani et al. 2018) genes have been characterized previously (Figure 4). This next highest number of markers mapped to the A08 chromosome (13%), where the *Crr1* gene has been reported (Hatakeyama et al. 2013; Hassan and Rahman 2016; Hirani et al. 2018) (Figure 5). About 5-8% of the markers were identified on the A01, A02, and A06 chromosomes, where the *Crr2*, *CRc* and the *Crr4* genes have been identified, respectively. In addition, 11% of the markers mapped to each of the A04 and A05 chromosomes, while about 2-5% of the markers mapped to the A07, A09 and the A10 chromosomes, where no major clubroot resistance genes have been mapped thus far. About 8% of the markers mapped to the C-genome (C02, C03 and C05) of *B. oleracea*, while 7% mapped to scaffolds. The physical positions and the chromosomal locations of the associated SNP and the PCR-based markers on the *B. rapa* and *B. oleracea* genome sequences are provided in Table 3.1.

3.3.3 Putative functions of proteins encoded by identified sequences

The sequences identified in this study matched entries in GenBank corresponding to ATP binding proteins, kinases, hydrolases, transferases, transcription factors, DNA topoisomerases, chaperone proteins, and translocation-related proteins, which are involved in cellular and biological processes as well as plant growth and development. Genes that encoded nucleotide-binding site leucine-rich repeat (NBS-LRR) proteins, WD-40 repeat family proteins, and syntaxin and histone deacetylases, which are associated with plant defense against pathogens, also matched some of the sequences identified in this study. Other matching genes encoded proteins including the RING/U-box superfamily proteins and nodulin-like transporter family proteins, which have been reported to play a role in plant growth and organ size development. Overall, about 60% of the genomic regions associated with the identified clubroot loci were previously characterized. The remaining 40% encoded proteins of unknown molecular function (Table 3.1).

3.4 Discussion

The complex genetic basis of clubroot disease and the emergence of new virulent isolates of *P. brassicae* make it imperative to identify molecular markers significantly associated with resistance to different *P. brassicae* pathotypes for use in marker-assisted selection (MAS) and for the breeding of clubroot-resistant Brassica crops. Genome wide association studies have without doubt proven to be one of the most useful methods for finding significant marker-trait associations (Rafalski 2010). In this study, 124 rutabaga accessions collected from Norway, Sweden, Finland, Denmark and Iceland were evaluated for resistance to pathotypes of 17 *P. brassicae* identified in Canada from 2003 to 2018 (Strelkov et al. 2003, 2016 and 2018). This is the most comprehensive clubroot GWAS or genetic mapping study carried out to date, since the previous studies utilized from one to five pathotypes (Hirai et al. 2004; Piao et al. 2004; Saito et al. 2006; Kato et al. 2012, 2013; Matsumoto et al. 1998,

2012; Chu et al. 2014; Fredua-Agyeman and Rahman 2016; Hassan and Rahman 2016; Li et al. 2016; Yu et al 2016; Pang et al. 2018; Zhang et al. 2014).

Different GWAS models were tested to find the best fit model for identifying associations between the SNP and PCR-based markers and clubroot resistance. Clearly, the PCA + K model, which found 43 SNP markers to be strongly associated with clubroot resistance, overestimated the number of markers. In contrast the Q + K model, which detected 13 markers, underestimated the number of markers associated with clubroot resistance. Therefore, we combined the results from the two models to harness the strengths of the two methods. The 11 markers captured by the two MLM models were the most consistent, but we found that markers detected by either the PCA + K or the Q + K method only still had value. For example, the SNP marker Bn_A03_p21487106, which was detected only by the Q + K model, was associated with Leucine-rich repeat receptor kinases (LRR-RKs).The InDel marker B1005, which was detected only by the PCA + K model, was associated with the Toll and interleukin-1 receptor nucleotide binding site-Leucine rich repeat (TIR-NBS-LRR) protein. These disease resistance proteins play roles in host-specific and non-host-specific defense and wounding responses, as well as in the control of development, the stress response and senescence in plants (Torri 2004; McHale et al. 2006).

'Electronic PCR' (e-PCR) is a useful procedure to search for and position DNA sequences on reference genomes (Schuler 1998; Rotmistrovsky et al. 2004; Fredua-Agyeman and Rahman 2016). By positioning the markers identified in this study and markers from previous studies that co-segregated with clubroot resistance on the *B. rapa* reference genome sequences, we identified the top and bottom segments of the A03 chromosome and the middle segment of the A08 chromosome of rutabaga as genomic hotspots associated with resistance to the different *P. brassicae* pathotypes.

The first genomic hotspot was located at the top half of the A03 chromosome and was found by e-PCR to be flanked by the SCAR marker HC688-6 (14,396,950 nt) (Matsumoto et al. 2012) and the STS marker BrSTS-061 (15,161,430 nt) (Pang et al. 2018). In this study, the SNP markers Bn_A03_p13610858 (13,610,459 nt) and Bn_Scaffold000164 _p55747 (16,537,330 nt) flanked the *Crr3* (Hirai et al. 2004; Saito et al. 2006), *CRk* (Matsumoto et al. 2012) and *CRd* (Pang et al. 2018) genes. The SNP marker Bn_A03_p13610858 was located ~786,000 nt upstream of the SCAR marker HC688-6, while the SNP marker Bn_Scaffold000164_p55747 was located ~1,375,000 nt downstream of BrSTS61. The genomic region based on the SNP markers spanned ~3,000,000 nt compared with that based on the PCR-based markers which spanned ~765,000 nt and conferred resistance to the field isolates representing pathotypes 2B and 8P. The large physical distances between the SNP and the PCR-based markers located at the top half of the A03 chromosome, and the fact that the genes controlling clubroot resistance in this region conferred resistance to multiple *P. brassicae* pathotypes, suggest that the genes controlling clubroot resistance in this region were independent.

The second genomic hotspot was located at the bottom half of the A03 chromosome conferred resistance to isolates representing 10 pathotypes. Fredua-Agyeman and Rahman (2016) identified the major resistance gene(s) in this genomic region to be the *CRa* and or the CRb^{Kato} genes based on markers from Matsumoto et al. (1998, 2012) and Kato et al. (2012, 2013). In the literature, three other clubroot resistance genes *CRb* (Piao et al. 2004, Zhang et al. 2014), *Rcr1* (Chu et al. 2014; Yu et al. 2016) and *BraA.CR.a* (Hirani et al. 2018) were also mapped to the same genomic region. Marker TCR37 (23,826,564 nt) (Zhang et al. 2014) and SC2930-2 (24,814,696 nt) (Matsumoto et al. 1998, 2012) flanked all five major clubroot resistance genes (*CRa*, *CRb*, CRb^{Kato}, *Rcr1* and *BraA.CR.a*) on the bottom half of the A03 chromosome and this genomic region spanned about 1 million (~ 988,000 nt) nucleotides. In this study, a much smaller genomic region (~ 260,000 nt) with flanking markers B1005 (24,376,816 nt) and KB29N19 (24637310 nt) conferred resistance to four SSIs (pathotypes

3H, 5I, 6M and 8N) and six field isolates (pathotypes 2B, 3A, 5C, 5G, 5L and 5K). In contrast to the top half, this study and the previous studies mapped the *CRa*, *CRb*, *CRb*^{Kato}, *Rcr1* and *BraA.CR.a* genes located at the bottom half of the A03 chromosome to < 1 million nucleotides from each other. This strongly suggests that the aforementioned reported genes could be alleles. It also is possible, however, that they could be part of a gene cluster with each conferring resistance to the different pathotypes as reported by Zhang et al. (2014) and Yu et al. (2016).

The third genomic hotspot was located at the middle segment of the A08 chromosome with SSR marker A08-5021 (11,614,477 nt) being closest to the overlapping gene (Bra034629) and the SNP markers Bn_A08_p10123561 (10,122,198 nt) and Bn_scaff_16110_1_p2556157 (13,408,834) as the flanking markers. This genomic region conferred resistance to pathotypes and 3H, 5L and 5K in this study. Hirani et al. (2018) used linkage analysis to position the *Crr1* gene between SSR markers S11R11 and S06R06, which by e-PCR was between 10,779598 to 10,970,334 nt. Hassan and Rahman (2016) reported that the *Crr1* gene in the rutabaga cultivar 'Brookfield' was located in the genomic region (10,692,602 to 11,617,968 nt) and it conferred resistance to pathotypes 2F, 3H, 5I, 6M and 8N. Unfortunately, e-PCR could not position any of the *Crr1* markers used by Hatakeyama et al. (2013) on the *B. rapa* genome. However, markers for the amplification of the 2nd Intron and 3rd intron of the *Crr1* gene were positioned between S11R11 (8,529,231 nt) and S27R27 (6,871,891 nt) on the *B. napus* genome. This study positioned the *Crr1* gene in the same genomic region as was reported by the other three studies. Hatakeyama et al. (2013) reported two gene loci for the *Crr1* gene: *Crr1a* which encodes a TIR-NB-LRR class of R proteins and has major effects, and *Crr1b* with minor effects but necessary to confer resistance to some *P. brassicae* isolates. The overlapping genes identified in this genomic region encoded a cyclase-associated protein, a DDB1-CUL4 associated factor and proteins of unknown molecular function. Therefore, functional analysis studies need to be carried out to get a full understanding of the role of the overlapping genes in this genomic region.

In conclusion, genomic regions on the A- and C-genomes of rutabaga associated with resistance to different *P. brassicae* pathotypes from Alberta were identified. As expected, a greater percentage of the markers were located on the A-genome, especially the A03 chromosome, where previous studies have reported several clubroot resistance genes. In addition, the study identified novel clubroot resistance loci on the C-genome of rutabaga. Markers identified in this study need to be validated to find out whether they directly co-segregate with clubroot resistance or they are in linkage disequilibrium with a QTL that contributes to the resistance. The markers identified in this study will be valuable in MAS and the breeding of clubroot resistant cruciferous crops.

3.5 Tables

Table 3.1. SNP and PCR-based markers in rutabaga accessions, their chromosomal location and linkage association with clubroot caused by 17 single-spore and field isolates representing 16 different pathotypes of *Plasmodiophora brassicae*.

SNP Marker[α]	Linkage group	Pathotype [Ω]	Description of gene functions in *Brassica rapa, B. oleracea, B. napus* and *Arabidopsis thaliana*[¥]
Bn_A01_p161237	A01	3H and 6M	Molecular function unknwon
Bn_A01_p3070657 [a]	A01	5K	Topoisomerase family protein
Bn_scaff_16394_1_p847636	A01	2F	Molecular function unknwon
Bn_scaff_16876_1_p908602 [b]	A02	3H	Nodulin-related; WAT1-related protein At5g40230-like
Bn_A02_p4210188	A02	3H	Molecular function unknwon
Bn_A02_p6615821 [c]	A02	3H	Molecular function unknwon
Bn_A03_p21205471 [d]	A02	3H	GDSL-motif lipase/hydrolase family protein
Bn_scaff_17721_1_p272248	A02	8N	Molecular function unknwon
Bn_A03_p7088375	A03	6M	EPIDERMAL patterning factor-like protein 6
Bn_A03_p7094698	A03	6M	Molecular function unknwon
Bn_A03_p8764481	A03	5I	Molecular function unknwon
Bn_A03_p13610858	A03	8P	WD-40 repeat family / beige-related; BEACH domain-containing protein
Bn_Scaffold000164_p55747 [e]	A03	2B	Transcription activator; two-component response regulator-like APRR1
Bn_A07_p6850383 [f]	A03	3O	Catalytic/ cation binding / hydrolase
Bn_A03_p21205471 [d]	A03	3H	Histone deacetylase
Bn_A03_p21377430	A03	3H	Molecular function unknwon
Bn_A03_p21487106	A03	3H	Leucine-rich repeat receptor-like serine/threonine-protein kinase BAM2
Bn_scaff_16110_1_p2556157 [g]	A03	3H	DDB1-CUL4 associated factor 1; nucleotide binding
Bn_A01_p3070657 [a]	A03	5K	DNA topoisomerase family protein
Bn_A04_p248884	A04	3H	Molecular function unknwon
Bn_A04_p251383 [h]	A04	3H	OTU-like cysteine protease family protein
Bn_A04_p15492182	A04	5K	Chaperone protein dnaJ 13; DNAJ heat shock N-terminal
Bn_scaff_16394_1_p842382	A04	2F	Condensin complex subunit 3

Bn_scaff_16394_1_p920749	A04	2F	Molecular function unknwon
Bn_scaff_16876_1_p908602 [b]	A04	3H	Transcriptional elongation regulator; WAT1-related protein At5g40230
Bn_scaff_15585_1_p978781 [i]	A04	3O	PLL1 (POLTERGEIST LIKE 1); catalytic/ protein serine/threonine phosphatase
Bn_A05_p894768	A05	3H	DEAD/DEAH box helicase, putative; P-loop containing hydrolases protein
Bn_A05_p3191390	A05	3H	Jacalin-related lectin 22-like
Bn_scaff_15585_1_p978781 [i]	A05	3O	PLL1 (POLTERGEIST LIKE 1); catalytic/ protein serine/threonine phosphatase
Bn_A05_p16738871	A05	6M	Translocation protein-related; nuclear transcription factor Y subunit A-9
Bn_A05_p17894045 [j]	A05	6M	4-galactosyl-N-acetylglucosaminide 3-alpha-L-fucosyltransferase
Bn_A05_p19650278	A05	8N	Molecular function unknwon
Bn_A05_p19650965	A05	8N	Molecular function unknwon
Bn_A05_p17894045 [j]	A06	6M	Fucosyltransferase, transferring glycosyl groups
Bn_A06_p17037739	A06	5C	ATP binding / endoribonuclease, serine/threonine kinase
Bn_A06_p18000461	A06	3H	Endonuclease, putative, Flap endonuclease
Bn_A07_p6850383 [f]	A07	3O	BGLU9 (BETA GLUCOSIDASE 9); catalytic/ cation binding / hydrolase
Bn_A07_p6542254 [k]	A07	3O	Molecular function unknwon
Bn_A07_p22005058	A07	5I	Alpha-1,3/1,6-mannosyltransferase ALG2
Bn_A05_p17894045 [j]	A08	6M	Fucosyltransferase, transferring glycosyl groups
Bn_A08_p8869180	A08	3H	SH3 domain-containing protein 1 (SH3P1)
Bn_A08_p10123561	A08	5K	Molecular function unknwon
Bn_scaff_16110_1_p2556157 [g]	A08	3H	DDB1-CUL4 Associated factor 1; nucleotide binding
Bn_A07_p6542254 [k]	A08	3O	SYP61; syntaxin-61-like; SNAP receptor
Bn_A08_p17061248	A08	2F	Molecular function unknwon
Bn_A08_p17393018	A08	8N	Molecular function unknwon
Bn_A08_p18310412	A08	5K	Molecular function unknwon
Bn_A03_p21205471 [d]	A09	3H	Histone deacetylase
Bn_A07_p6542254 [k]	A09	3O	SYP61 syntaxin-61-like; SNAP receptor
Bn_A04_p251383 [h]	A09	3H	OTU-like cysteine protease family protein

Bn_A02_p6615821 [c]	A10	3H	Molecular function unknwon
Bn_scaff_16665_1_p188604	C03	3H	Molecular function unknwon
Bn_scaff_16665_1_p199303	C03	3H	Molecular function unknown
Bn_scaff_17721_1_p273764	C02	8N	Vegetative storage protein 2
Bn_scaff_17721_1_p309137	C02	8N	Molecular function unknown
Bn_scaff_18338_1_p871455	C05	2F, 5I and 3A	Molecular function unknown
Bn_scaff_16876_1_p908602 [b]	Scaffold000522	3H	Molecular function unknwon
Bn_Scaffold000247_p28610	Scaffold000247	5K	Molecular function unknwon
Bn_Scaffold000164_p46170	Scaffold000164	2B	ATMRP8; ATPase, coupled to transmembrane movement, ABC transporter C family member 6-like
Bn_Scaffold000164_p55747 [e]	Scaffold000164	2B	Transcription activator; two-component response regulator-like APRR1
KB29N19 (SSR)	A03	3H, 6M, 8N, 2B, 3A, 5C, 5G , 5I	Molecular function unknown (Overlapping gene Bra019372)
B0903 (SSR)	A03	5C, 5L and 5K	Molecular function unknown (overlapping gene Bra019408)
B1005 (InDel)	A03	5K	Disease resistance protein (TIR-NBS-LRR class) (overlapping gene Bra019410)
A08-5021 (SSR)	A08	5L	Cyclase-associated protein 1-like (overlapping gene Bra034629)

[α] SNP markers denoted with the same superscript letter mapped to multiple chromosomes on the reference genomes.

[β] The type of PCR-based markers showing trait association has been specified.

[Ω] Pathotype designations are as per the Canadian Clubroot Differential Set (letters) and the system of Williams (numbers).

[¥] Putative functions are based on matching entries in the GenBank database of NCBI

Table 3.2. List of 55 PCR-based primers linked to previously identified clubroot resistance genes used to genotype 124 rutabaga accessions.

Marker	Chrom	CR gene	Forward	Reverse	Reference
GC2360-2	A03	*CRa*	AGTTTTGTAATTTTC ACCCAAAGTATCA	CAGCTGGAGGAGCAC TGCAACGGAGAGA	Matsumoto et al. (2012)
GC2920-1	A03	*CRa*	ATTTTCGAATCATCC AAGCTGAAAGT	TGTTCAACAAGTTCCC ATCTCCAT	Matsumoto et al. (2012)
GC2920-2	A03	*CRa*	CAAAGAACTGCCTG TTGTAAGTAAA	TGTTCAACAAGTTCCC ATCTCCAT	Matsumoto et al. (2012)
SC2930-1	A03	*CRa*	TAGACCTTTTTTTTG TCTTTTTTTTTACCT	AAGGCCATAGAAATC AGGTC	Matsumoto et al. (2012)
SC2930-2	A03	*CRa*	CAGACTAGACTTTTT GTCATTTAGACT	AAGGCCATAGAAATC AGGTC	Matsumoto et al. (2012)
GC1430-1	A03	*CRa*	AGAACAATCTTACCT TGTGCCTCTTCATAC	TGATAGGTATCTACCT GAGCTAAGGCGTGA	Matsumoto et al. (2012)
TCR05	A03	*CRb*	AGAATCATGACCGG GGAAAT	GCAGCTAAGTCATCG ACCAA	Piao et al. (2004)
TCR09	A03	*CRb*	GCAGCAACCGATAA TATAAGGA	AACCAGAAGAAGAAA AACAAAAA	Piao et al. (2004)
TCR10	A03	*CRb*	AACTCTTGAAGAAA GCAAAGAAGC	GCAGGAATAAGAAGG AACACCA	Piao et al. (2004)
TCR08	A03	*CRb*	GCAGAATTATAACCT GAGCGTGT	ATTACCGGAGTATGCG ATCC	Piao et al. (2004)
TCR02	A03	*CRb*	GCTCCATTCAGTTAC GGTGA	GCAGAGAATTTTGGA AGAGGA	Piao et al. (2004)
TCR17	A03	*CRb*	GCACATCACTTTGAG GACGA	TTTCCGTTGTCCTTTGT GAA	Zhang et al. (2014)
TCR37	A03	*CRb*	GGGAATTTGAGGAG GGACTC	CAAAGAAAGAGCTCC AATCACA	Zhang et al. (2014)
TCR74	A03	*CRb*	ATGGATGATGGATG GATAGAGTG	TTGAACCATAGGAGG GATAGTTG	Zhang et al. (2014)
KBrH129J0 8Rc	A03	*CRb*^{Kato}	ATGAGATTGAAGAG GGAAACACAA	GTTTCCAATGGTGAAA CCAATCCTA	Kato et al. (2012)
KB69N08	A03	*CRb*^{Kato}	ATTACACGGTCCATG AAAAGAT	GTTTGGTTACCGAAAC AGAAGGAA	Kato et al. (2013)
KB69N05	A03	*CRb*^{Kato}	ATCACAACCAAAAT GGAATGAC	GTTTCTCAAGCACCGA GACTCATAA	Kato et al. (2013)
KBrH059N2 1F	A03	*CRb*^{Kato}	ATCGACGCCGTTTAT TAGAACTCG	GTTTACGCCACGTCAG CTCACTAACTC	Kato et al. (2012)
KB59N08	A03	*CRb*^{Kato}	AGCATCTTGCAAATT ATTTACGTT	GTTTGACCGTGTTATT GTTGTAGGG	Kato et al. (2013)
KB59N07	A03	*CRb*^{Kato}	ATGTACTCGGGTGTC CCCTAGA	GTTTGACACGATGAAC CAGAC	Kato et al. (2013)
KB59N06	A03	*CRb*^{Kato}	ATGAAATTGCAACTC TCAAAATG	GTTTAGGCTTTCTCCA TCAACCACTA	Kato et al. (2013)
KB59N05	A03	*CRb*^{Kato}	AGTCAACGAAACAA AGATATGC	GTTTCTTTTTCTCCAC AAAAGGAGAGC	Kato et al. (2013)
KB59N03	A03	*CRb*^{Kato}	AGGTAAATCCTCAA AAAGCCAT	GTTTGGCGAAATTCAG TTGACA	Kato et al. (2013)
B4701	A03	*CRb*^{Kato}	AGATTCTTGTTCTTC TCGCTGG	GTTTACGGAGACCATG AAGGATAATG	Kato et al. (2013)
B4732	A03	*CRb*^{Kato}	ATCTGATGTACCTTT GTGCTGG	GTTTGTCAATCATTCA AGCTAAGTGG	Kato et al. (2013)
B1324	A03	*CRb*^{Kato}	ATAATGGCTTCAAAT AGTCAAAA	GTTTGCATATACACGT TGAGGAAAC	Kato et al. (2013)

B1210	A03	*CRb*^{Kato}	ATTGAAAAGTTGACTCCGTTGA	GTTTCTTCCTTGAAGTTGCTTCAGCTCTTC	Kato et al. (2013)
B1005	A03	*CRb*^{Kato}	AGGAAGTTGTGGTGTTTTGAA	GTTTATATCCTCGATCATGGCAGC	Kato et al. (2013)
B0902	A03	*CRb*^{Kato}	AGCCTTGCGTAAAAGCAACTAC	GTTTGGAATCCGACAAATACATCCAT	Kato et al. (2013)
B0903	A03	*CRb*^{Kato}	ACTTCCTCTGCTTTTCTCAGGT	GTTTGAAACTCTTCTCCCCCTTC	Kato et al. (2013)
BGA01	A03	*CRb*^{Kato}	ATCTGACTGTTTGTGAAAGCGA	GTTTAGAGTTTTTGGGTGCAAATGTT	Kato et al. (2013)
BGA02	A03	*CRb*^{Kato}	ACAAATTCACAAGTCTTCCTCC	GTTTACATGCAATTGATGGGAAAA	Kato et al. (2013)
BGA06	A03	*CRb*^{Kato}	AGAAATAGCAAAGCTCAAACGG	GTTTCCAGAAAAGAGATGCAGACAA	Kato et al. (2013)
BGA10	A03	*CRb*^{Kato}	AGATTACAAAATTTTCAAAGTGAGA	GTTTCTCACACTTTCCTTAAATAAAAGCTA	Kato et al. (2013)
BGA12	A03	*CRb*^{Kato}	ACCCCCTCTCTTTCCTACTTTC	GTTTCATTCATTGGTCATAAGGCAA	Kato et al. (2013)
BGA15	A03	*CRb*^{Kato}	ACCAAAAACATCAGCTTTCGTA	GTTTCAGAATTCTTTATGAATAGGTTGC	Kato et al. (2013)
BGB29	A03	*CRb*^{Kato}	ATTTCGCTCTACACTTTTCCCC	GTTTGTTTCTGAGGAGGCTCATT	Kato et al. (2013)
BGB36	A03	*CRb*^{Kato}	AGCTAACATTGCAGACTTTGCT	GTTTGATAACCATGCTGTAGCGAG	Kato et al. (2013)
BGB41	A03	*CRb*^{Kato}	ATCGCATAAACTAATAAAAATCAAAA	GTTTGACCCACATGATTAACAA	Kato et al. (2013)
KB29N19	A03	*CRb*^{Kato}	ATGAGATCGTCAGCCATTTCTC	GTTTCCAGTCCGGTTTTTATTACCTT	Kato et al. (2013)
KB29N17	A03	*CRb*^{Kato}	ACAGCTCCTTTTTAGGTAACGA	GTTTGGATTGCAAGTGTTATTTCCA	Kato et al. (2013)
KB29N16	A03	*CRb*^{Kato}	AGACTCGACAAGGTATCGATCT	GTTTGACGCCATTATGACACAACT	Kato et al. (2013)
KB29N11	A03	*CRb*^{Kato}	ACTCTCCACCAACACTTCCTAA	GTTTGAAGCTATCTTAGACCACC	Kato et al. (2013)
KB29N05	A03	*CRb*^{Kato}	ATACAAGCTCTCAGAGGAGGAA	GTTTCAGCTTGACACTCTTGACTTGC	Kato et al. (2013)
KBrH129J18R	A03	*CRb*^{Kato}	AGAGCAGAGTGAAACCAGAACT	GTTTCAGTTCAGTCAGGTTTTTGCAG	Kato et al. (2012)
KB91N13	A03	*CRb*^{Kato}	AGACGGAGACTTTGAGATCTGG	GTTTCGAGTACTTCCAGAAACACG	Kato et al. (2013)
KB91N06	A03	*CRb*^{Kato}	AGAATTTCCTTGTTAGCCAAAT	GTTTGTGTTTGTTCATTTTCTATTTCAGA	Kato et al. (2013)
KBrB091M11R	A03	*CRb*^{Kato}	ACTTAAAGCACGAGAATGCAAA	GTTTGGTGTCGAAGCTATGTGTG	Kato et al. (2012)
TCR02-F	A03	*CRb*^{Kato}	AGCTCCATTCAGTTACGGTGA	GTTAAGAAACTTGCAGAAACTCG	Kato et al. (2012)
BRMS-088	A08	*Crr1*	TATCGGTACTGATTCGCTCTTCAAC	ATCGGTTGTTATTTGAGAGCAGATT	Suwabe et al. 2003
A08_5021	A08	*Crr1*	TCGATTGATCCGACAAAACA	ATCGTCGTAAGCCAGAATCG	Hobson and Rahman (2016)
A08_5024	A08	*Crr1*	GAACACGAAGCGTGTCTGAA	AAGAAACCATCGGTGTCGAG	Hobson and Rahman (2016)
BRMS-096	A01	*Crr2*	AGTCGAGATCTCGTTCGTGTCTCCC	TGAAGAAGGATTGAAGCTGTTGTTG	Suwabe et al. 2003
BrSTS-078	A03	*Crr3*	CTCTCCTCTAACCTGTTCCAAGAA	GGTGTATCCACACACTCATCAAGT	Saito et al. (2006)
BRM-125F	A06	*Crr4*	GTTCTCAAAGGGAAACCGAAAAACA	GAGTTGGCCAGAGATTTACATGCGT	Suwabe et al. (2006)

3.6 Figures

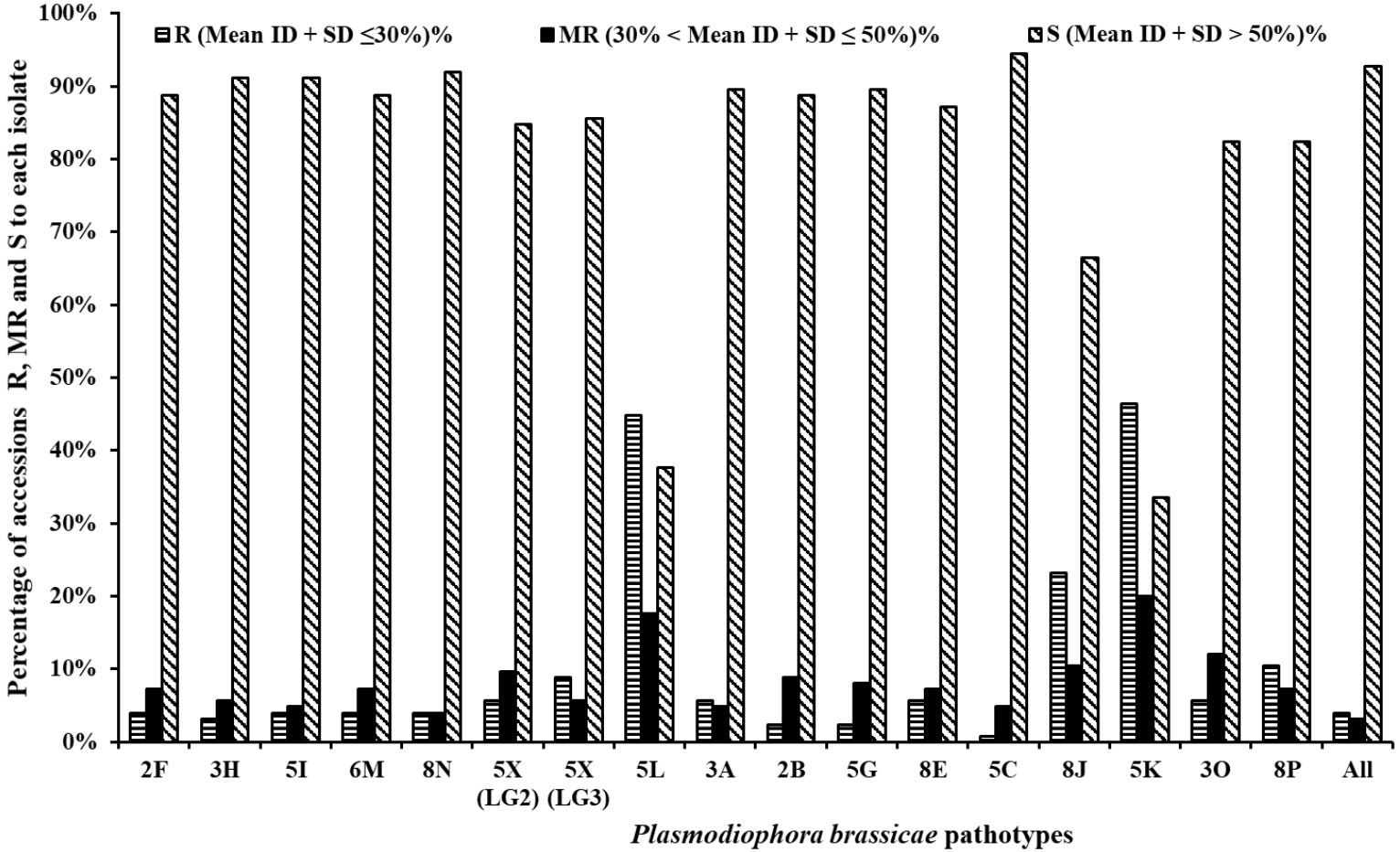

Figure 3.1 Frequency distribution of 124 rutabaga (*Brassica napus* ssp. *napobrassica*) accessions evaluated in greenhouse experiments for resistance to 17 *Plasmodiophora brassicae isolates* representing 16 different pathotypes. Pathotypes 2F, 3H, 5I, 6M and 8N are single-spore isolates identified prior to the introduction of clubroot resistant (CR) varieties in Canada, while pathotypes 5X (LG2 and LG3), 5L, 3A, 2B, 5G, 8E, 5C, 8J, 5K, 3O and 8P are represented by field isolates identified after the introduction of CR varieties in Canada.

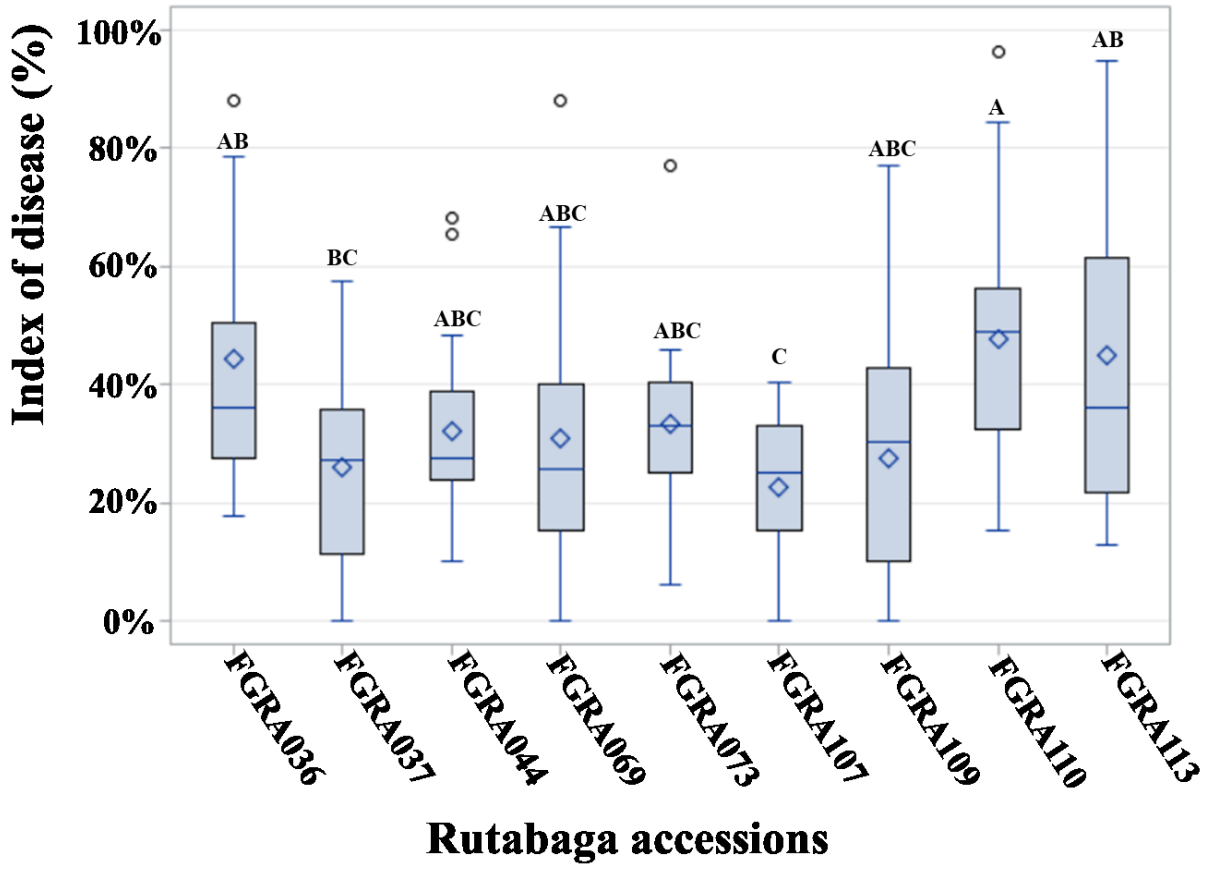

Figure 3.2 Distribution of indices of disease (IDs) among rutabaga (*Brassica napus* ssp. *napobrassica*) accessions resistant (R) or moderately resistant (MR) to 17 *Plasmodiophora brassicae* isolates representing 16 different pathotypes. The grand mean (GM) ID (◊), median (line inside box), 75th percentile (upper end of box), 25th percentile (lower end of box) as well as the maximum and minimum observations for all 17 isolates are presented by Box-and-Whiskers plots. The GM is the mean ID for an accession across all 17 isolates. The genotypes were considered R if the GM ID + Standard Deviation (SD) ≤ 30% and MR 30% < GM ID + SD ≤ 50%. Accessions with the same letters are not significantly different.

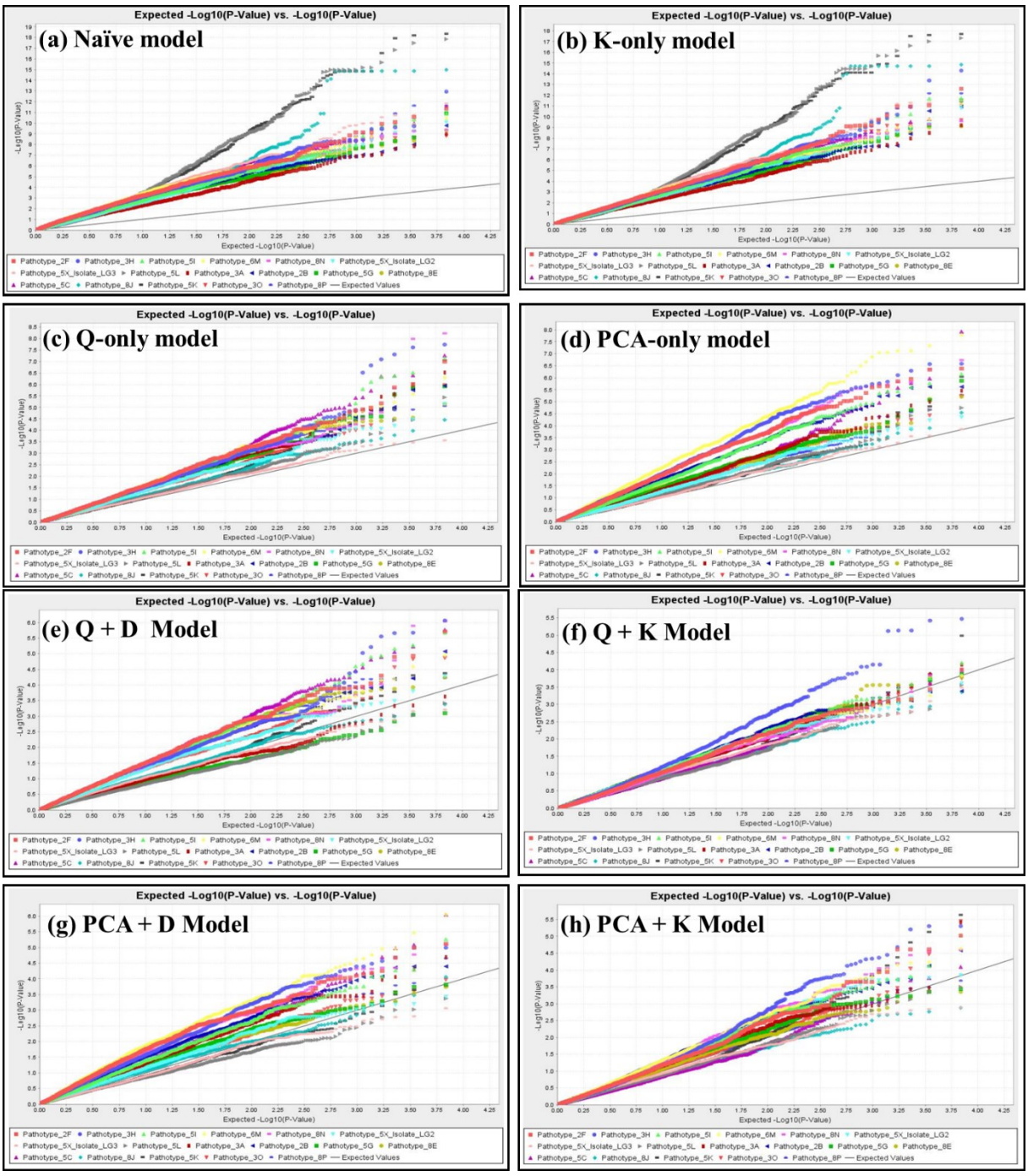

Figure 3.3 Quantile-Quantile comparison of eight GWAS models for identifying clubroot resistance loci in 124 rutabaga (*Brassica napus* ssp. *napobrassica*) accessions tested with 17 field and single-spore isolates of *Plasmodiophora brassicae* representing 16 different pathotypes. The four GLM tested comprised the Naïve(N)(a), Kinship (K)-only (b), Population structure (Q)-only (c) , and the Principal Coordinate Analysis (PCA)-only (d). The four MLM tested comprised Q + D (e), Q + K (f), PCA + D (g), PCA + K (h) models; where D is the Distance Matrix. The black line is the expected $-\log_{10} P$ distribution while colored lines are the observed $-\log_{10} P$ distribution for each of the 17 *P. brassicae* pathotypes.

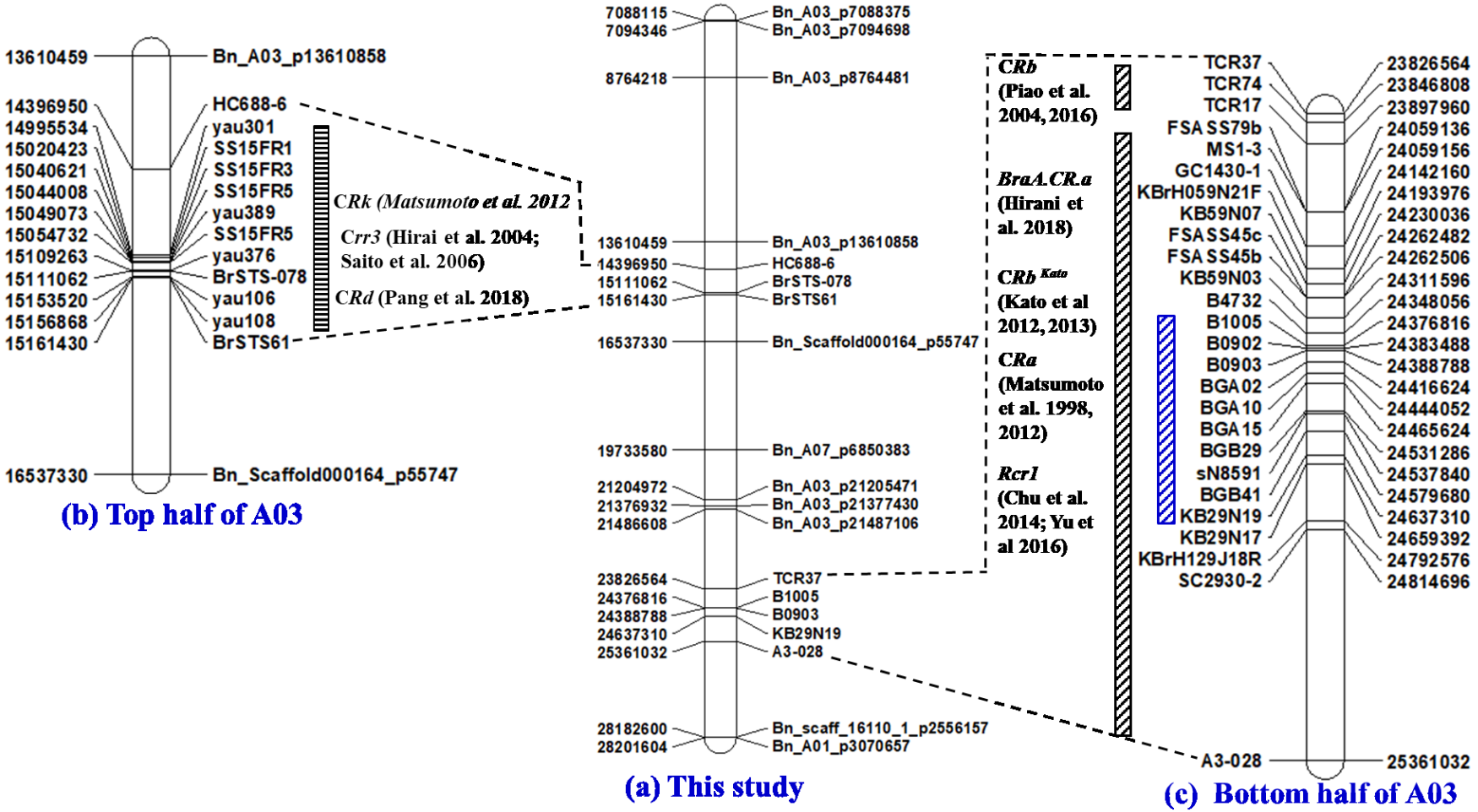

Figure 3.4 Physical maps of the A03 chromosome of *B. rapa* constructed by the use of SSR and SNP markers identified to be associated with clubroot resistance in this study (a) and the PCR-based markers previously identified to be linked to the *Crr3* (Hirai et al. 2004; Saito et al. 2006), *CRk* (Matsumoto et al. 2012) and *CRd* (Pang et al. 2018) gene(s) located on the top half of chromosome A03 (b), as well as the *CRa* (Matsumoto et al. 1998, 2012), *CRb* (Piao et al. 2004, Zhang et al. 2014), *CRb*^{Kato} (Kato et al. 2012, 2013), *Rcr1* (Chu et al. 2014; Yu et al. 2016) and *BraA.CR.a* (Hirani et al. 2018) gene(s) located on the bottom half of chromosome A03 (c).

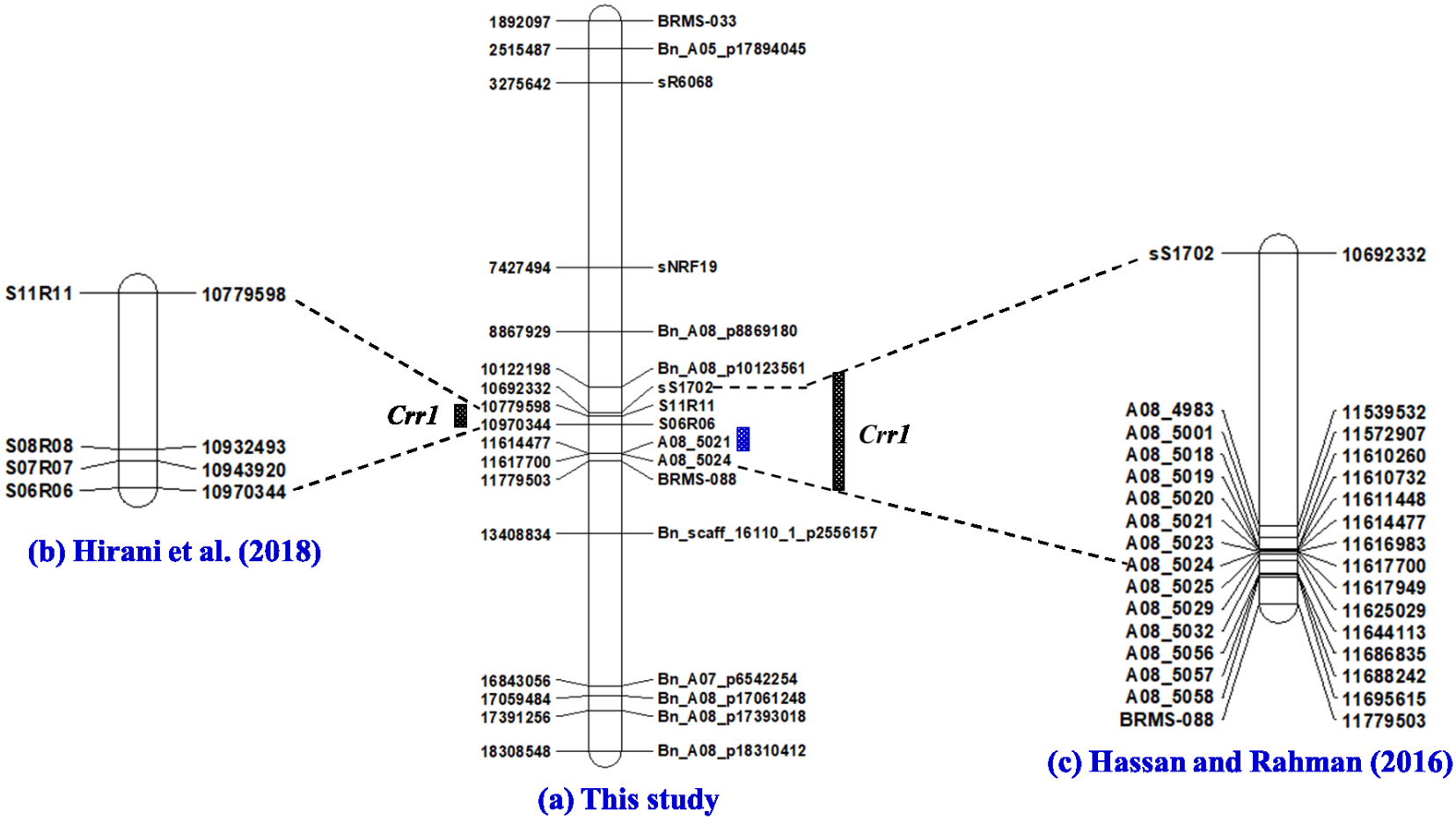

Figure 3.5 Physical maps of chromosome A08 of *Brassica rapa* constructed by the use of SSR and SNP markers identified to be associated with clubroot resistance in this study (a), and the PCR-based markers previously identified to be linked to the *Crr1* gene (Hirani et al. 2018) (b) and Hassan and Rahman (2016) (c). Fine mapping of the *Crr1* genomic region was conducted with SSR markers developed by Hobson and Rahman (2016).

Chapter 4: General Conclusions

4.1 Introduction

Clubroot (*Plasmodiophora brassicae*) is a major problem globally, although it can be managed by combining cultural and chemical strategies with the deployment of resistant cultivars (Donald and Porter 2009). In Canada, the planting of clubroot resistant (CR) canola (*Brassica napus*) has been one of the most successful approaches for clubroot management (Hwang et al., 2014); unfortunately, the repeated cultivation of cultivars with similar resistance places selection pressure on pathogen populations to overcome that resistance (LeBoldus et al. 2012; Strelkov et al. 2016; Fredua-Agyeman et al. 2018). In Alberta, novel virulent strains of *P. brassicae* have been detected in increasing numbers, as the clubroot outbreak itself continues to spread (Strelkov et al. 2018; Strelkov et al. 2019).

There have been a number of large-scale resistance screening studies to identify Brassica genotypes resistant to old and new pathotypes of *P. brassicae* in western Canada (Hasan et al. 2012; Peng et al. 2014; Fredua-Agyeman et al. 2019). While accessions of *B. nigra* showed good resistance, the introgression of B-genome genes or traits into AC-genome *B. napus* canola is technically difficult (Rahman et al. 2014). In contrast, the AC-genome rutabaga (*B. napus* ssp. *napobrassica*) is highly fertile with canola. Hasan et al. (2012) found that some rutabaga accessions carry broad-spectrum resistance to different *P. brassicae* pathotypes, which could prove valuable to the breeding of CR canola.

While rutabaga is grown worldwide (Gowers 2010), the extent of genetic diversity in accessions from different regions was not known, or at least reported, prior to the work presented here. To the best of my knowledge, this is the first study to examine the genetic relationships

between rutabaga accessions, and to extensively evaluate and characterize a diverse collection of accessions for resistance to a suite of *P. brassicae* pathotypes. This work provides a foundation for future research, including studies on the molecular basis of clubroot resistance in rutabaga and in fully exploiting this subspecies of *B. napus* as a resistance source.

4.2 Genetic diversity and population structure of Nordic rutabaga accessions

One hundred twenty-seven rutabaga accessions, including 124 Nordic accessions and three German, Canadian and Russian accessions included as external references, were genotyped with 13,714 SNP markers. The Nordic rutabaga accessions were found to have a relatively narrow genetic background, with the population structural analysis placing more than half of the accessions (51.4%) in an admixture group. The reference accession from Russia, 'Krasnoselskaya', was grouped in the admixture subpopulation, while 'Wilhemsburger' (Germany) and 'Laurentian' (Canada) clustered in a group with the Danish accessions and another group that included accessions from Denmark, Norway and Sweden, respectively. This indicated different pedigrees for these accessions. The diversity analysis also distinguished the geologically isolated Icelandic subpopulation from the other Nordic accessions, while high genetic similarity was found between the accessions from Denmark, Finland, Norway and Sweden. This information may help to understand the history of rutabaga domestication, as well as being useful for breeding purposes when associated with phenotypic traits.

4.3 Clubroot resistance of rutabaga accessions

The 124 Nordic rutabaga accessions were screened for resistance to a suite of 17 isolates representing 16 'old' and 'new' (resistance-breaking) pathotypes of *P. brassicae*. Nearly three-quarters (72.6%) of the accessions showed resistance to at least one of these isolates. Higher

frequencies of resistant accessions were observed with pathotypes 5L, 5K and 8J (23.2-46.4%) compared with isolates representing other pathotypes (0.8-10.4%). More than half of the accessions (58.1%) showed isolate- or pathotype-specific resistance, while 14.5% appeared to have broad-spectrum resistance, indicating that the clubroot resistance traits within this population are under different genetic control.

Among the rutabaga subpopulations from the five Nordic countries, the Danish accessions generally showed the greatest level of resistance, with over one-third of the accessions (34.8%) found to be resistant or moderately resistant to anywhere from 7-17 isolates, and only 17.4% found to be completely susceptible. Clubroot resistance responses were most diverse in the Swedish subpopulation, with 88.0% of the accessions showing resistance to different pathotypes and only 14.0% having broad-spectrum resistance. In contrast, 60.7% of Norwegian accessions were completely susceptible, although three of the accessions (10.7%) showed resistance to 11-16 of the isolates.

Compared with PCR-based approaches such as RAPD, RFLP, AFLP, and SSR markers, DNA-chip based SNP arrays provide more consistent and specific detection of genome-wide polymorphisms (Vignal et al. 2002). In this study, 45 SNPs across the A and C genomes were linked to resistance to one to three isolates. Three genomic regions on chromosomes A03 and A08 were identified as hotspots of clubroot resistance genes, which linked to 19 SNPs from this study and four previously published PCR-based markers. Candidate genes identified from these markers included not only those encoding plant defense-related proteins such as NBS-LRR proteins, WD-40 repeat family proteins, and syntaxin and histone deacetylases, but also proteins associated with intra- or intercellular functions in plant development such as ATP binding

proteins, transferases, transcription factors, and DNA topoisomerases. The functions of these genes should be validated in future studies.

4.4 Final remarks

This research was limited to rutabaga accessions from the Nordic region and to a *P. brassicae* collection mainly from western Canada. A wider selection of rutabagas from across the world would be expected to be more diverse, while the Canadian *P. brassicae* isolates examined did not include all of the pathotypes that have been found elsewhere (Bus et al. 2011; Piao et al. 2009; Shen et al. 2009; Strelkov et al. 2018). In addition, the SNP array used in this study was based on the *B. rapa* and *B. napus* genomes, and may not provide fully appropriate genomic information specific to rutabaga (*B.napus* ssp. *napobrassica*). Further study of rutabaga and its interactions with *P. brassicae* may, therefore, yield additional genes for resistance and serve as a model for exploring the clubroot pathosystem.

References

Adhikari KKC (2010) Effect of temperature, biofungicides and fungicides on clubroot of selected *Brassica* crops. M.Sc. thesis. University of Guelph, Guelph, ON. 207 pp

Ahmad A (1994) Clubroot disease of Chinese cabbage in the Cameron Highlands. Teknol Sayur-sayuran Jil 10:17–20

Ahmed HU, Hwang SF, Strelkov SE, Gossen BD, Peng G, Howard RJ, Turnbull GD (2011). Assessment of bait crops to reduce inoculum of clubroot (*Plasmodiophora brassicae*) of canola. Can J Plant Sci 91(3): 545-551.

Ahokas H (2004) On the evolution, spread and names of rutabaga. In: MTT-agrifood research Finland 2004. Kave, Helsinki, Finland, pp: 32.

Arie T, Kobayashi Y, Kono Y, Gen O,Yamaguchi I (1999) Control of clubroot of crucifers by *Phoma glomerata* and its product epoxydon. Pesticide Science 55: 602–4.

Arie T, Kobayashi Y, Okada G, Kono Y,Yamaguchi I (1998) Control of soilborne clubroot disease of cruciferous plants by epoxydon from *Phoma glomerata*. Plant Pathol 47, 743–8.

Asano T, Kageyama K (2006) Growth and movement of secondary plasmodia of *Plasmodiophora brassicae* in turnip suspensionculture cells. Plant Pathol 55:145–151

Ayers GW (1944) Studies on the life history of the club root organism, *Plasmodiophora brassicae*. Can J Res Sect C 32:143– 149

Ayers GW, Lelacheur KE (1972) Genetics of resistance in rutabaga to races of *Plasmodiophora brassicae*. Can J Plant Sci 52: 897- 900.

Ayers GW., Lelacheur KE (1972) Genetics of resistance in rutabaga to two races of *Plasmodiophora brassicae*. Can J Plant Pathol 52(6): 897-900.

Botero A, García C, Gossen BD, Strelkov SE, Todd CD, Bonham-Smith PC, Pérez-López E (2019). Clubroot disease in Latin America: distribution and management strategies. Plant Pathol 68(5): 827-833.

Brachi B, Morris GP, Borevitz JO (2011). Genome-wide association studies in plants: the missing heritability is in the field. Genome biol 12(10), 232.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633-2635.

Bradshaw JE, Gemmell DJ, Wilson RN (1997) Transfer of resistance to clubroot (*Plasmodiophora brassicae*) to swedes (*Brassica napus* L. var. *napobrassica* Peterm) from B. rapa. Ann Appl Biol 130: 337–48.

Bradshaw JE, Griffiths DW (1990) Sugar content of swedes for stock feeding. J Sci Food and Agric 50(2): 167-172.

Buczacki ST, Cadd SE (1976) Glasshouse evaluation of systemic compounds, derivatives of dithiocarbamic acid and other fungicides for the control of clubroot. Ann Appl Biol 84:43–50

Buczacki ST, Ockendon JG (1979) Preliminary observations on variation in susceptibility to clubroot among collections of some wild crucifers. Ann Appl Biol 92(1): 113-118.

Buczacki ST, White JG (1979) The value of soil sterilants for the control of clubroot on a field scale. Plant Pathol 28:36–39

Buczacki, ST, Toxopeus H, Mattusch P, Johnston TD, Dixon GR, Hobolth LA (1975) Study of physiological specialization in *Plasmodiophora brassicae*: proposals for attempted rationalization through an international approach. Transactions of the British Mycological Society, 65: 295-303.

Burki F, Keeling PJ (2014) Rhizaria. Curr Biol 24: R103–R107.

Bus A, Körber N, Snowdon RJ, Stich B (2011) Patterns of molecular variation in a species-wide germplasm set of *Brassica napus*. Theor Appl Genet 123(8): 1413-1423.

Cao T, Manolii VP, Hwang SF, Howard RJ and Strelkov SE (2009) Virulence and spread of *Plasmodiophora brassicae* [clubroot] in Alberta, Canada. Can J Plant Pathol 31: 321-329.

Cao T, Tewari J, Strelkov SE (2007) Molecular detection of *Plasmodiophora brassicae*, causal agent of clubroot of crucifers, in plant and soil. Plant Dis 91:80– 87.

Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Sci 345(6199): 950-953.

Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X, et al. (2014) Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. Sci 345: 950–953.

Chittem K, Mansouripour SM, del Río Mendoza LE (2014) First report of clubroot on canola caused by Plasmodiophora brassicae in North Dakota. Plant dis 98(10): 1438-1438.

Chu M, Song T, Falk KC, Zhang X, Liu X, Chang A, Lahlali R, McGregor L, Gossen BD, Yu F, Peng G (2014) Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. BMC genomics 15(1):1166.

Clarke WE, Higgins EE, Plieske J, Wieseke R, Sidebottom C, Khedikar Y, et al. (2016) A high-density SNP genotyping array for *Brassica napus* and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. Theor Appl Genet 129: 1–13.

Colhoun J (1958) Clubroot disease of crucifers caused by *Plasmodiophora brassicae* Woron. A monograph. Phytopathological Paper No 3. Commonwealth Mycological Institute, Surrey, UK

Conforth IS (1971) Calcium cyanamide in agriculture. Soils Fertilisers 34:463–470

Creelman DW (1958) Thirty-eighth annual report of the Canadian plant disease survey. Can. Plant Dis. Surv 38: 48–84.

Crête R (1981) Worldwide importance of clubroot. Clubroot Newslett No. 11: pp 6–7

Crisp P, Crute IR, Sutherland RA et al. (1989) The exploitation of genetic resources of Brassica oleracea in breeding for resistance to clubroot (*Plasmodiophora brassicae*). Euphytica 42: 215–26.

Crute IR, Pink DAC (1989) The characteristics and inheritance of resistance to clubroot in *Brassica oleracea*. Aspects Appl Biol 23:57–60

Diederichsen E, Beckmann J, Schondelmeier J, Dreyer F (2006). Genetics of clubroot resistance in *Brassica napus*' Mendel'. Acta horticulturae (706): 307-311.

Diederichsen E, Deppe U, Sacristan MD (2003). Characterization of clubroot resistance in recent winter oilseed rape material. Proceedings of the 11th International Rapeseed Congress. 1: 68-70.

Diers BW, Osborn TC (1994) Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphisms. Theor Appl Genet 88: 662–668.

Dixon GR (2009) The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. J Plant Growth Regul 28: 194-202.

Dixon GR (2017) Managing clubroot disease (caused by *Plasmodiophora brassicae* Wor.) by exploiting the interactions between calcium cyanamide fertilizer and soil microorganisms. J Agric Sci 155(4): 527-543.

Dixon GR, Brokenshire T (1981) Chemical control of clubroot (*Plasmodiophora brassicae*). Crop protection in Northern Britain. Proceedings of a conference held at Dundee University 17-19 March 1981. 325-329. Farnham: The British Crop Protection 795 Council.

Dixon GR, Wilson F (1983) Evaluation of calcium cyanamide for control of *Plasmodiophora brassicae* (clubroot). Ann Appl Biol 102(Suppl.): 50-51.

Dixon RA (2001) Natural products and plant disease resistance. Nature, 411: 843-847.

Dobson RL, Gabrielson RL (1983) Role of primary and secondary zoospores of *Plasmodiophora brassicae* in the development of clubroot in Chinese cabbage. Phytopathol 73:559–561

Donald C, Porter I (2009) Integrated control of clubroot. J Plant Growth Regul 28: 289-303.

Donald EC, Lawrence JM, Porter IJ (2004) Influence of particle size and application method on the efficacy of calcium cyanamide for control of clubroot of vegetable brassicas. Crop Prot 23:297–303

Donald EC, Porter IJ (2014). Clubroot in Australia: the history and impact of *Plasmodiophora brassicae* in Brassica crops and research efforts directed towards its control. Can J Plant Pathol 36(sup1): 66-84.

Donald EC, Porter IJ, Lancaster RA (2001) Band incorporation of fluazinam (Shirlan) into soil to control clubroot of vegetable *Brassica* crops. Aust. J. Exp. Agric. 41: 12231226.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4(2): 359-361.

Ernst TW, Kher S, Stanton D, Rennie DC, Hwang SF, Strelkov SE (2019) Plasmodiophora brassicae resting spore dynamics in clubroot resistant canola (Brassica napus) cropping systems. Plant Pathol 68(2):399–408.

Estey RH (1994). Essays on the early history of plant pathology and mycology in Canada. Montreal, Québec and Kingston, Ontario: McGill– Queen's University Press.

Evanno G, Regnaut S, Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol, 14(8): 2611-2620.

Excoffier L., Smouse PE., Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Faggian R, Strelkov SE (2009) Detection and measurement of *Plasmodiophora brassicae*. J Plant Growth Regul 28(3): 282-288.

Fredua-Agyeman R, Coriton O, Huteau V, Parkin IA, Chèvre A, Rahman H (2014) Molecular cytogenetic identification of B genome chromosomes linked to blackleg disease resistance in *Brassica napus*× *B. carinata* interspecific hybrids. Theor Appl Genet 127(6): 1305-1318.

Fredua-Agyeman R, Hwang SF, Strelkov SE, Zhou Q, Feindel D (2018) Potential loss of clubroot resistance genes from donor parent *Brassica rapa* subsp. *rapifera* (ECD 04) during doubled haploid production. Plant Pathol 67(4): 892-901.

Fredua-Agyeman R, Hwang SF, Strelkov SE, Zhou Q, Manolii VP, Feindel D (2019) Identification of Brassica accessions resistant to 'old'and 'new'pathotypes of *Plasmodiophora brassicae* from Canada. Plant Pathology, 68(4), 708-718.

Fredua-Agyeman R, Rahman H (2016) Mapping of the clubroot disease resistance in spring *Brassica napus* canola introgressed from European winter canola cv.'Mendel'. Euphytica 211(2): 201-213.

Friberg H, Lagerlof J, Ramert B (2005) Germination of *Plasmodiophora brassicae* resting spores stimulated by a non-host plant. Eur J Plant Pathol 113:275–281

Friberg H, Lagerlof J, Ramert B (2006) Usefulness of nonhost plants in managing Plasmodiophora brassicae. Plant Pathol 55:690–695

Froese RD, Derksen H, Guo X, McLaren DL (2019) Monitoring and occurrence of clubroot in Manitoba in 2018. Can Plant Dis Surv. 99: 224-225.

Gemmell DJ, Griffiths DW, Bradshaw JE (1990) Effect of cultivar and harvest date on dry-matter content, hardness and sugar content of swedes for stockfeeding. J Sci Food Agric 53(3): 333-342.

Gibbs JG (1931) Dissemination of clubroot in the dung from stock. NZ J Agric 42:193–198

Gibbs JG (1932) Weed host plants of club-root in New Zealand. NZ J Agric 44: 273-276.

Gonzaga ZJ, Aslam K, Septiningsih EM, Collard, BCY (2015) Evaluation of SSR markers for Molecular breeding in rice. Plant Breed Biotech 3(2):139-152

Govindaraj M, Vetriventhan M, Srinivasan M (2015) Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genet Res Int pp 1-14.

Gowers S (2010) Swedes and turnips. In: Bradshaw JE (ed.) Root and tuber crops. Handbook of plant breeding, vol 7. Springer, New York, pp. 245-289

Grant I, Harney PM, Christie BR (1982) Inheritance of yield and other quantitative characters in *Brassica napus* var. *napobrassica*. Can J Genet Cytol 24(4): 459-465.

Griffiths DW, Bradshaw JE, Taylor J, Gemmell DJ (1991) Effect of cultivar and harvest date on the glucosinolate and S-methylcysteine sulphoxide content of swedes (*Brassica napus* ssp. *rapifera*). J Sci Food Agric 56: 539–549.

Gupta S, Kumari K, Muthamilarasan M, Parida SK, Prasad M (2014). Population structure and association mapping of yield contributing agronomic traits in foxtail millet. Plant cell reports 33(6): 881-893.

Harling R, Kennedy SH (1991) Biological control of *Plasmodiophora brassicae* using a bait crop. Mededelingen Faculteit Landbouwwetenschappen Rijksuniversiteit Gent 56:159–170

Harvey N (1949) The coming of the Swede to Great Britain: An obscure chapter in farming history. Agric His 23(4): 286-288.

Hasan MJ, Rahman H (2016) Genetics and molecular mapping of resistance to *Plasmodiophora brassicae* pathotypes 2, 3, 5, 6, and 8 in rutabaga (*Brassica napus* var. *napobrassica*). Genome 59(10): 805-815.

Hasan MJ, Strelkov SE, Howard RJ, Rahman H (2012) Screening of Brassica germplasm for resistance to *Plasmodiophora brassicae* pathotypes prevalent in Canada for broadening diversity in clubroot resistance. Can J Plant Sci 92(3): 501-515.

Hatakeyama K, Suwabe K, Tomita RN, Kato T, Nunome T, Fukuoka H, Matsumoto S (2013) Identification and characterization of Crr1a, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. PloS one. 8(1): e54745.

Hirai M (2006) Genetic analysis of clubroot resistance in Brassica crops - Review. Breed Sci 56: 223-229.

Hirai M, Harada T, Kubo N, Tsukada, Suwabe K, Matsumoto S (2004) A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. Theor Appl Genet 108: 639-643.

Hirani AH, Gao F, Liu J, Wu C, McVetty PB, Duncan RW, Li G (2018) Combinations of independent dominant loci conferring clubroot resistance in all four turnip accessions (*Brassica rapa*) from the European clubroot differential set. Front Plant Sci 9: 1628.

Hobson N, Rahman H (2016) Genome-wide identification of SSR markers in the Brassica A genome and their utility in breeding. Can J Plant Sci 96(5): 808-818.

Howard RJ, Strelkov SE, Harding MW (2010) Clubroot of cruciferous crops – new perspectives on an old disease. Can J Plant Pathol 32: 43–57.

Hwang SF (2018) Management of clubroot of canola in Alberta, Canada. The International Clubroot Workshop 2018.

Hwang SF, Cao T, Xiao Q, Ahmed HU, Manolii VP, Turnbull GD, Gossen BD, Peng G, Strelkov SE (2012) Effects of fungicide, seeding date and seedling age on clubroot severity, seedling emergence and yield of canola. Can. J. Plant Sci 92: 1175.

Hwang SF, Howard RJ, Strelkov SE, Gossen BD, Peng G (2014) Management of clubroot (*Plasmodiophora brassicae*) on canola (*Brassica napus*) in western Canada. Can J Plant Pathol 36(sup1): 49-65.

Hwang SF, Strelkov SE, Gossen BD, Turnbull GD, Ahmed HU, Manolii VP (2011) Soil treatments and amendments for amelioration of clubroot of canola. Can. J. Plant Sci 91(6): .999-1010.

Iñiguez Luy FL, Federico ML (2011) The genetics of *Brassica napus* L. In: Bancroft I and Schmidt R (ed.) Genetics and genomics of the Brassicaceae. Springer, New York Dordrecht, Heidelberg, London, pp. 291–322

Ikegami H (1985) Decrease of clubroot fungus by cultivation of different crops in heavily infested soil. Res Bull Fac Agric Gifu Univ 50:19–32

Ikegayami H, Naiki T, Ito T, Imuro Y (1982) Ultrastructural growth process of Plasmodiophora brassicae in infected cells of Chinese cabbage and turnip (Studies on the Clubroot of Cruciferous Plants IV). Res Bull Fac Agric Gifu Univ 46:9–19

Ingram DS, Tommerup IC (1972) The life history of Plasmodiophora brassicae Woron. Proc R Soc Lond Ser B 180:103–112

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics: 23(14): 1801-1806.

Ji HW, Ren L, Chen KR, Xu L, Liu F, Sun CC, Li J, Liu SY, Fang XP (2013) Identification of physiological races of club root and resistance of rape cultivars to Plasmodiophora brassicae. Chin J Oil Crop Sci 35(3): 301-206.

Jung GA, Byers RA, Panciera MT, Shaffer JA (1986) Forage dry-matter accumulation and quality of turnip, swede, rape, Chinese-cabbage hybrids and kale in the eastern USA. Agron J 78: 245–253.

Kageyama K, Asano T (2009) Life Cycle of Plasmodiophora brassicae. J Plant Growth Regul 28: 203-211.

Karling JS (1968) The *Plasmodiophorales*. *The Plasmodiophorales*. 2nd edition, Published by the author, New York, USA. 256 pp.

Kato T, Hatakeyama K, Fukino N, Matsumoto S (2012) Identificaiton of a clubroot resistance locus conferring resistance to a *Plasmodiophora brassicae* classified into pathotype group 3 in Chinese cabbage (*Brassica rapa* L.). Breed sci 62(3):282-7.

Kato T, Hatakeyama K, Fukino N, Matsumoto S (2013) Fine mapping of the clubroot resistance gene CRb and development of a useful selectable marker in *Brassica rapa*. Breed Sci 63: 116–24.

Katsura K, Tamura M, Yamaguchi N (1970) Some observations on the infection of *Plasmodiophora brassicae* Woronin in root hair of Cruciferae plants. Proc Kansai Plant Prot 12:23–29

Khoury WE, Makkouk K (2010) Integrated plant disease management in developing countries. J. Plant Pathol 4: S35.

Kim CH, Cho WD, Kim HM (2000) Distribution of *Plasmodiophora brassicae* causing clubroot disease of Chinese cabbage in soil. Plant Dis Res 6(1):27–32

Kim S, Shin C, Moon S et al. (2004) Isolation and characterization of *Streptomyces* sp. KACC 91027 against *Plasmodiophora brassicae*. J Microbiol Biotech 14: 220– 3.

Krupinsky J, Gossen BD, Turkington T, Bailey K, McMullen M (2002) Managing plant disease risk in diversified cropping systems. Agron J 94: 198-209.

Kuginuki Y, Ajisaka H, Yui M, Yoshikawa H, Hida K, Hirai M (1997) RAPD markers linked to a clubroot-resistance locus in Brassica rapa. Euphytica 98: 149-154.

Lamers W, Toxopeus H (1977) New 'Multipot'clubroot resistance screening method in use at the SVP. In Woronin+ 100 International Conference on Clubroot, Madison, Wisconsin (pp. 122-123).

Lammerink J (1967) The inheritance of clubroot resistance *in Brassica napus* L. N Z J Agric Res 10:109–115

Landry BS, Hubert N, Crete R, Change MS, Lincoln SE, Etoh T (1992) A genetic map for Brassica oleracea based on RFLP markers detected with expressed DNA sequences and mapping of resistance genes to race 2 of *Plasmodiophora brassicae* (Woronin). Genome 35: 409-420.

Laurens F, Thomas G (1993) Inheritance of resistance to clubroot (*Plasmodiophora brassicae* Wor.) in kale (*Brassica oleracea* ssp. *acephala*). Hereditas 119:253–262

Li L, Luo Y, Chen B, Xu K, Zhang F, Li H, Huang Q, Xiao X, Zhang T, Hu J, Li F, Wu X (2016) A Genome-Wide Association Study Reveals New Loci for Resistance to Clubroot Disease in *Brassica napus*. Front Plant Sci 7: 1483.

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21(9): 2128-2129.

Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin I, Parkin IAP, et al. (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. Nat Commun 5:3930.

Liu B, Wang Y, Zhai W, Deng J, Wang H, Cui Y, Cheng F, Wang X, Wu J (2013) Development of InDel markers for *Brassica rapa* based on whole-genome re-sequencing. Theo App Genet 126(1):231-9.

Livingstone M, Jones AS, Mennie I (1977) Swedes (*Brassica napus*) for growing pigs: chemical composition and use as a replacement for barley in the diet. Anim Feed Sci Tech 2:31-40.

Ludwig-Muller J, Prinsen E, Rolfe SA, Scholes JD (2009) Metabolism and hormone action during clubroot disease. J Plant Growth Regul 28: 229-244.

Mailer RJ, Scarth R, Fristensky B (1994) Discrimination among cultivars of rapeseed (*Brassica napus* L.) using DNA polymorphisms amplified from arbitrary primers. Theor Appl Genet 87(6): 697-704.

Matsumoto E, Ueno H, Aruga D (2012) Accumulation of three clubroot resistance genes through marker-assisted selection in Chinese cabbage (Brassica rapa ssp. pekinensis). J Jpn Soc Hortic Sci 81, 184–90.

Matsumoto E, Yasui C, Ohi M, Tsukada M (1998) Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Euphytica 104(2): 79.

McHale L, Tan X, Koehl P, Michelmore RW (2006). Plant NBS-LRR proteins: adaptable guards. Genome biol 7(4): 212.

Mitani S, Sugimoto K, Hayashi H, Takii Y, Ohshima T, Matsuo N (2003) Effects of cyazofamid against Plasmodiophora brassicae Woronin on Chinese cabbage. Pest Manage Sci 59:287–293

Murakami H, Tsushima S, Akimoto T, Kuroyanagi Y, Shishido Y (2004) Quantitative studies on the relationship between plowing into soil of clubbed roots of preceding crops caused by *Plasmodiophora brassicae* and disease severity in succeeding crops. Soil Sci Plant Nutr 50:1307–1311

Murakami H, Tsushima S, Akimoto T, Murakami K, Goto I, Shishido Y (2000) Effects of growing leafy daikon (*Raphanus sativus*) on populations of *Plasmodiophora brassicae* (clubroot). Plant Pathol 49:584–589

Myers DF, Campbell RN (1985) Lime and the control of clubroot of crucifers: effects of pH, calcium, magnesium, and their interactions. Phytopathology 75:670–673

Naiki T, Dixon GR (1987). The effects of chemicals on developmental stages of *Plasmodiophora brassicae* (clubroot). Plant Pathol 36: 316-327

Naiki T, Kawaguchi C, Ikegami H (1984) Root hair reinfection in Chinese cabbage seedlings by the secondary zoospores of *Plasmodiophora brassicae* Woronin. Ann Phytopathol Soc Jpn 50:216–220

Narisawa K, Shimura M, Usuki F,Fukuhara S, Hashiba T (2005) Effects of pathogen density, soil moisture, and soil pH on biological control of clubroot in Chinese cabbage by *Heteroconium chaetospira*. Plant Dis 89: 285–90.

Narisawa K, Tokumasu S, Hashiba T (1998) Suppression of clubroot formation in Chinese cabbage by the root endophytic fungus, *Heteroconium chaetospira*. Plant Pathol 47: 206–10.

Nurhasanah N, Ecke W (2016) Construction of a high-fidelity genetic linkage map using AFLP and SSR markers in rapeseed (Brassica napus L.). SABRAO J Breed Genet 48: 189-199.

Neik TX, Barbetti MJ and Batley J (2017) Current status and challenges in identifying disease resistance genes in *Brassica napus*. Front Plant Sci 8: 1788.

Neuhauser S, Bulman S, Kirchmair M (2010) Plasmodiophorids: The challenge to understand soil-borne, obligate biotrophs with a multiphasic life cycle. Mol Identif Fungi: 51-78.

Neuhauser S, Kirchmair M, Gleason FH (2011) Ecological roles of the parasitic Phytomyxea (plasmodiophorids) in marine ecosystems-a review. Mar Freshw Res 62: 365-371.

Nie G, Huang L, Zhan X, Taylor M, Jiang Y, et al. (2016) Marker-trait association for biomass yield of potential bio-fuel feedstock *Miscanthus sinensis* from southwest China. Frontiers in Plant Science, 7, 802.

Nieuwhof M, Wiering D (1961). Testing cabbage plants for clubroot resistance. Euphytica. 10(2):191-200.

Noling JW (2008) Soil Fumigation. In: Capinera J.L. (eds) Encyclopedia of Entomology. Springer, Dordrecht, pp: 224

Overton J (1996) Ecologically Based Pest Management-New Solutions for a New Century. National Academy Press, Washington DC, USA.

Pang W, Fu P, Li X, Zhan Z, Yu S, Piao Z (2018) Identification and mapping of the Clubroot resistance gene CRd in Chinese cabbage (Brassica rapa ssp. pekinensis). Front Plant Sci 9:653.

Pasko P, Bukowska-Strakova K, Gdula-Argasinska J, Tyszka-Czochara M (2013) Rutabaga (*Brassica napus* L. var. *napobrassica*) seeds, roots, and sprouts: a novel kind of food with antioxidant properties and proapoptotic potential in Hep G2 hepatoma cell line. J Med Food 16 (8): 749-759.

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2: 2074–2093.

Peakall R and Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288-295.

Peakall R and Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28: 2537-2539.

Peng G, Falk KC, Gugel RK et al. (2014) Sources of resistance to *Plasmodiophora brassicae* (clubroot) pathotypes virulent on canola. Can J Plant Pathol 36: 89–99.

Peng G, Mcgregor L, Lahlali R, Gossen BD, Hwang SF, Adhikari KK, Strelkov SE, McDonald MR (2011) Potential biological control of clubroot on canola and crucifer vegetable crops. Plant Pathol 60: 566- 574.

Peng G, Pageau D, Strelkov SE, Gossen BD, Hwang SF, Lahlali R (2015) A> 2-year crop rotation reduces resting spores of *Plasmodiophora brassicae* in soil and the impact of clubroot on canola. Eur J Agron 70, 78-84.

Piao Z, Ramchiary N, Lim YP (2009) Genetics of clubroot resistance in *Brassica* species. J Plant Growth Regul 28(3): 252-264.

Piao ZY, Deng YQ, Choi SR, Park YJ, Lim YP (2004) SCAR and CAPS mapping of CRb, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* var. *pekinensis*). Theor Appl Genet 108: 1458-1465.

Pivovarova NS (1979) Breeding material of culinary turnip and swede and methods of evaluating it. Byulleten' Vsesoyuznogo Instituta Rastenievodstva 90: 32–38.

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am J Hum Genet 67(1): 170-181.

Rahman H, Peng G, Yu F, Falk KC, Kulkarni M, Selvaraj G (2014) Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus* L.). Can J Plant Pathol 36(sup1): pp.122-134.

Rahman H, Peng G, Yu F, Falk KC, Kulkarni M, Selvaraj G (2014). Genetics and breeding for clubroot resistance in Canadian spring canola (*Brassica napus* L.). Can J Plant Pathol 36(sup1): pp.122-134.

Rankin WH, Fraser WP (1920) Survey of the prevalence of common plant diseases in the Dominion of Canada, 1920. First Annual Report. Ottawa, Ontario: Dominion of Canada Department of Agriculture.

Ren L, Xu L, Liu F, Chen K, Sun C, Li J, et al. (2016) Host range of *Plasmodiophora brassicae* on cruciferous crops and weeds in China. Plant Dis 100: 933–939.

Rennie DC, Holtz MD, Turkington TK, Leboldus, JM, Hwang SF, Howard RJ, Strelkov SE (2015). Movement of Plasmodiophora brassicae resting spores in windblown dust. Can J Plant Pathol 37(2): 188-196.

Rennie DC, Manolii VP, Cao T, Hwang SF, Howard RJ, Strelkov SE (2011) Direct evidence of surface infestation of seeds and tubers by *Plasmodiophora brassicae* and quantification of spore loads. Plant Pathol 60: 811–819.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4(1): 137-138.

Rotmistrovsky K, Jang W, Schuler GD (2004). A web server for performing electronic PCR. Nucleic Acids Res 32(sup2): W108-12.

Saito M, Kubo N, Matsumoto S, Suwabe K, Tsukada M, Hirai M (2006) Fine mapping of the clubroot resistance gene, *Crr3*, in *Brassica rapa*. Theor Appl Genet 114:81–91

Saitou N and Nei M (1987) The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. Mol Biol and Evol 4:406–425

Sakamoto K, Saito A, Hayashida N, Taguchi G, Matsumoto E (2008) Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Theor Appl Genet 117: 759–767.

Sambrook J, Russell DW (2001). Molecular cloning: A laboratory manual. Cold Spring.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7: 465–474.

Schuler GD (1998). Electronic PCR: bridging the gap between genome mapping and genome sequencing. Trends in biotechnol 16(11):456-9.

Schwelm A, Fogelqvist J, Knaust A, Jülke S, Lilja T, Bonillarosso G, et al. (2015) The *Plasmodiophora brassicae* genome reveals insights in its life cycle and ancestry of chitin synthases. Sci Rep 5:11153.

Shattuck VI, Proudfoot KG (1990) Rutabaga breeding. Plant Breed Rev 8: 217–48.

Shen XQ, Nie K, Wu Q, Zhang YQ, Meng XH (2009) Initial research report on differentiation identification of the Chinese cabbage clubroot main physiological races. China Veg 1(8): 59-62.

Siemens J, Bulman S, Rehn F, Sundelin T (2009) Molecular biology of *Plasmodiophora brassicae*. J Plant Growth Regul 28: 245-251.

Singh N, Choudhury DR, Singh AK, Kumar S, Srinivasan K (2013) Comparison of SSR and SNP Markers in Estimation of Genetic Diversity and Population Structure of Indian Rice Varieties. PLoS ONE 8(12): e84136.

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. University of Kansas Scientific Bulletin 28:1409–1438

Somé A, Manzanares MJ, Laurens F, Baron F, Thomas G, Rouxel F (1996) Variation for virulence on *Brassica napus* L. amongst *Plasmodiophora brassicae* collections from France and derived single-spore isolates. Plant Pathol 45: 432–439.

Spaner, D. (2002). Agronomic and horticultural characters of rutabaga in eastern Canada. Can J Plant Pathol *82*(1): 221-224.

Steel RGD, Torrie JH (1960) Principles and procedures of statistics. Principles and procedures of statistics.

Strelkov SE, Hwang SF (2014) Clubroot in the Canadian canola crop: 10 years into the outbreak. Can J Plant Pathol 36(sup1): 27-36.

Strelkov SE, Hwang SF, Manolii VP, Cao T, Fredua-Agyeman R, Harding MW, Peng G, Gossen BD, Mcdonald MR, Feindel D (2018) Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. Can J Plant Pathol 40(2): 284-298.

Strelkov SE, Sheau-Fang Hwang, S-F, Manolii VP, Cao T, Feindel D (2016) Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. Eur J Plant Pathol 145(3): 517-529.

Strelkov SE, Manolii VP, Lageyre J, Hwang SF, Harding MW, Daniels GC (2019) Occurrence and spread of clubroot on canola in Alberta in 2018. Can Plant Dis Surv. 99:160–163.

Strelkov SE, Manolii VP, Cao T, Hwang SF, Orchard D (2007) Incidence of clubroot on canola in Alberta in 2006. Can Plant Dis Surv. 87:109–111.

Strelkov SE, Tewari JP, Smith-Degenhardt E (2006) Characterization of *Plasmodiophora brassicae* populations from Alberta, Canada. Can J Plant Pathol 28(3): 467-474.

Sturtevant EL (1919) Sturtevant's notes on edible plants. New York Agr. Exper. Sta., Geneva, New York, pp. 304-305.

Sundelin T, Christensen CB, Larsen J, Møller K, Lübeck M, Bødker L, Jensen B (2010) In planta quantification of *Plasmodiophora brassicae* using signature fatty acids and real-time PCR. Plant Dis 94(4), 432-438.

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Fujimura M, Nunome T, Fukuoka T, Matsumoto S, Hirai M (2003) Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa*. Theor Appl Genet 107: 997-1002.

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Kondo M, Fujimura M, Nunome T, Fukuoka H, Hirai M, Matsumoto S (2006) Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics 173:309–319

Takahashi K, Yamaguchi T (1987) An improved method for estimating the number of resting spores of *Plasmodiophora brassicae* in soil. Ann Phytopathol Soc Jpn 53: 507–515.

Tewari JP, Strelkov SE, Orchard D, Hartman M, Lange RM, Turkington TK (2005) Identification of clubroot on canola (*Brassica napus*) in Alberta. Can J Plant Pathol 27(1):143–144

Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S (2002) Agricultural sustainability and intensive production practices. Nature 418: 671-677.

Torii KU (2004) Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. International Review of Cytology 234:1-46.

Tremblay N, Be´lec C, Coulombe J, Godin C (2005) Evaluation of calcium cyanamide and liming for control of clubroot disease in cauliflower. Crop Prot 24:798–803

Tsykun T, Rellstab C, Dutech C, Sipos G, Prospero S (2017) Comparative assessment of SSR and SNP markers for inferring the population genetic structure of the common fungus *Armillaria cepistipes*. Heredity 119: 371–380.

Turesson G (1922a) The species and the variety as ecological units. Hereditas 3: 100-113.

Turesson G (1922b) The genotypic response of the plant species to habitat. Hereditas 3: 211-350.

Turesson G (1925) The plant species in relation to habitat and climate: contributions to the knowledge of genecological units. Hereditas 6(2): 147-236.

U N (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Jpn J Bot 7: 389−452.

Ueno H, Matsumoto E, Aruga D, Kitagawa S, Matsumura H, Hayashida N (2012) Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. Plant Mol. Biol. 80, 621–629.

Usuki F, Narisawa K (2007) A mutualistic symbiosis between a dark septate endophytic fungus, *Heteroconium chaetospira*, and a non-mycorrhizal plant, Chinese cabbage. Mycologia 99: 175–84.

Varshney RK, Thiel T, Sretenovic-Rajicic T, Baum M, Valkoun J, Guo P, Grando S, Ceccarelli S, Graner A (2008) Identification and validation of a core set of informative genic SSR and SNP markers for assaying functional diversity in barley. Mol Breeding 22:1–13 DOI 10.1007/s11032-007-9151-5

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genet Sel Evol 34(3): 275.

Voorrips RE, Visser DL (1993) Examination of resistance to clubroot in accessions of *Brassica oleracea* using a glasshouse seedling test. Neth J P1ant Pathol 99:269–276

Wallenhamar AC (1999) Monitoring and control of *Plasmodiophora brassicae* in spring oilseed brassica crops. Acta Universitatis Agriculturae Sueciae, Agraria 183. Swedish University of Agricultural Sciences, Uppsala, p 53

Wallenhammar AC (1996) Prevalence of *Plasmodiophora brassicae* in spring oil-seed rape growing area in Central Sweden and factors influencing soil infestation levels. Plant Pathol 45: 710–719

Wallenhammar AC, Almquist C, Söderström M, Jonsson A (2012) In-field distribution of *Plasmodiophora brassicae* measured using quantitative real-time PCR. Plant Pathol 61(1): 16-28.

Wang J, Huang Y, Hu XL, Niu YZ, Li XL, Liang Y (2008) Study of symptoms, yield loss of clubroot and modality of *Plasmodiophora brassicae* in rape. Chin J Oil Seed Sci 30(1):112–115

Wang J, Huang Y, Li XL, Li HZ (2011) Research progress in clubroot of crucifers. Plant Protection 37(6): 153-158

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, et al. (2011) *Brassica rapa* genome sequencing project consortium. The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43: 1035–1039. doi: 10.1038/ng.919

Watson AG, Baker KF (1969) Possible gene centers for resistance in the genus Brassica to *Plasmodiophora brassicae*. Econ Bot 23: 245–252.

Webster MA, Dixon GR (1991) Calcium, pH and inoculum concentration influencing colonization by *Plasmodiophora brassicae*. Mycol Res 95:65–73

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

White JG, Buczacki ST (1977) The control of clubroot by soil partial sterilization: a review. Ann Appl Biol 85:287–300

Wiggell P, Hawken RH, Wiggell D, Cock LJ, Bant JH (1961) Field trials on the control of clubroot in Brassicae. Ann Appl Biol 49:110–119

Williams PH (1966) A system for the determination of races of Plasmodiophora brassicae that infect cabbage and rutabaga. Phytopathol 56: 624–626.

Woronin M (1878) *Plasmodiophora brassicae*, Urheber der Kohlpflanzen - Hernie. Jahrb Wiss Bot 11:548-574 [translated by Chupp C (1934) Phytopathological classics no 4. American Phytopathological Society, St. Paul]

Wright S (1965). The interpretation of population structure by F-statistics with special regard to system of mating. Evolution 19: 395-420.

Xue S, Cao T, Howard RJ, Hwang SF, Strelkov SE (2008) Isolation and variation in virulence of single-spore isolates of *Plasmodiophora brassicae* from Canada. Plant Dis 92(3):456-62.

Yamagishi H, Yoshikawa H, Ashizawa KH, Yui S (1986) Effects of resistant plants as a catch crop on the reduction of resting spores of clubroot (Plasmodiophora brassicae Worn.) in soil. J Jpn Soc Hortic Sci 54:460–466

Yang J, Liu D, Wang X, Ji C, Cheng F, Liu B, et al. (2016) The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. *Nat. Genet.* 48, 1225–1232.

Yano S, Tanaka S, Kameya-Iwaki M, Katumoto K (1991) Relation of Ca2+ efflux to germination of resting spores of clubroot fungus. Bull Fac Agric Yamaguchi Univ 39:105–112

Yoshikawa H (1983) Breeding for clubroot resistance of crucifer crops in Japan. Jpn Agric Res Quarterly 17: 6-11.

Yu F, Zhang X, Huang Z, Chu M, Song T, Falk KC, Deora A, Chen Q, Zhang Y, McGregor L, Gossen BD (2016). Identification of genome-wide variants and discovery of variants

associated with Brassica rapa clubroot resistance gene Rcr1 through bulked segregant RNA sequencing. PLoS One 11(4): e0153218.

Yuan YK (1995) Research in clubroot of rapeseed. Plant Doctor 8(5): 25-27.

Zhang T, Zhao Z, Zhang C, Pang W, Choi SR, Lim YP, Piao Z (2014). Fine genetic and physical mapping of the CRb gene conferring resistance to clubroot disease in *Brassica rapa*. Mol Breed 34:1173–1183.

Ziesman B, Avila R, Ursu K, Kwasnicki J, Roszell L, Baraniecki C, et al. (2019) 2018 clubroot survey and known occurrence of clubroot in Sasketchewan. Can Plant Dis Surv. 99: 200-203.