Uniwersytet im. Adama Mickiewicza, Poznań, Poland
Wydział Biologii
Instytut Biologii Molekularnej i Biotechnologii
Zakład Biologii Obliczeniowej

Genetyczne determinanty odporności rzepaku (*Brassica napus* L.) na infekcję *Plasmodiophora brassicae* Wor.

Piotr Kopeć

Rozprawa doktorska

Pod kierunkiem Prof. dr. hab. Wojciecha Karłowskiego

Poznań 2023

Adam Mickiewicz University, Poznań, Poland
Faculty of Biology
Institute of Molecular Biology and Biotechnology
Department of Computational Biology

Genetic determinants of rapeseed (*Brassica napus* L.) resistance against *Plasmodiophora brassicae* Wor. infection.

Piotr Kopeć

Ph.D. Thesis

Supervised by Prof. dr hab. Wojciech Karłowski

Poznań 2023

Acknowledgements

I would like to express my gratitude to:

My supervisor, Prof. dr hab. Wojciech Karłowski, for consistently encouraging progress, maintaining openness to discussions, offering understanding, and providing a wealth of helpful advice.

My collaborators, particularly dr Katarzyna Mikołajczyk, for her willingness to help and her instrumental role in binding the project together, and to Ewa Jajor for her extensive support with the experiments. I would also like to acknowledge dr Przemysław Nuc for his invaluable assistance with the wet lab and for all the helpful, engaging, and entertaining conversations we have shared.

The entire Combio team - Aga Ch-P, Janek, Marta, Aga R, Karolina, Andrzej, Remigiusz, Deepti, Anand, Marek, Maciej, Marianna, Aga T, Medhat - for the time we have spent together in a friendly and supportive atmosphere. This journey would have been significantly more challenging without such a fantastic group of people around.

To dr Karolina Cerbin and Prof. Agnieszka Ludwików, for your constant search for, and management of opportunities that supported my development. Also, many thanks to Arleta Kucz and Iwona Kanonik-Jędrzejak - without your assistance with the administrative aspects, I would have found myself lost from day one, likely never to be found again.

The plant breeding team at the University of Gießen, not only for their methodological support and training, but also, for their warm welcome and the enjoyable times we have spent together.

To the people around IBMIB and the Faculty, for maintaining a friendly atmosphere and always being eager to help.

To my family for their unconditional belief in me and their continuous support throughout all these years.

My friends, for providing a sanctuary for venting and relaxation when needed the most.

And lastly, to Martynka, for encouraging my personal growth, and believing in me, especially during times when I had trouble doing so myself. Your understanding and support during the more stressful times have been invaluable. This work might not have happened without it.

Streszczenie

Brassica napus należy do najważniejszych roślin oleistych na świecie. Jej uprawom zagrażają liczne choroby roślin kapustowatych. Jedną z nich jest kiła kapusty, globalnie odpowiedzialna za około dziesięcioprocentową utratę plonów. Kiła kapusty wywoływana jest przez *Plasmodiophora brassicae* - obligatoryjnego endobiotrofa należącego do kladu Rhizaria. Patogen ten stymuluje rozrost zainfekowanych korzeni, w rezultacie wywołując charakterystyczne narośla. Porażone rośliny wykazują ogólnoustrojowe objawy, prowadzące do nieoptymalnego kwitnienia i przedwczesnego starzenia.

Ze względu na długowieczność i trwałość przetrwalników *P. brassicae*, ochrona upraw przed kiłą kapusty stanowi trudne wyzwanie. Jedną z najlepszych strategii ograniczania szkód jest wyprowadzanie odmian odpornych, opierające się na uprzedniej identyfikacji genetycznych czynników zaangażowanych w odpowiedź na infekcję.

W niniejszej pracy zmapowano *locus* odporności na kiłę kapusty w odmianie *B. napus* 'Tosca' do fragmentu genomu obejmującego około 100 kpz. Następnie zbadano miejscową zmienność sekwencji oraz różnice w ekspresji lokalnych genów, pomiędzy liniami odpornymi, a podatnymi na porażenie kiłą kapusty. W wyniku analiz zidentyfikowano wysoce polimorficzny, zduplikowany gen należący do rodziny TNL, który poddano analizie funkcjonalnej. Ektopiczna ekspresja sekwencji kodujących w *Arabidopsis thaliana* mocno wsparła hipotezę udziału zduplikowanego genu w odporności na chorobę.

Ponadto przeprowadzono analizę porównawczą struktury locus odporności wśród szerszego panelu genomów roślin kapustowatych. Wyniki podkreślają wysoką dynamikę ewolucyjną regionu, zwłaszcza w przypadku genu TNL.

Przeprowadzone badania dostarczyły cennych zasobów dla programów hodowli odpornościowej oraz otworzyły wiele nowych możliwości dalszych badań podstawowych i aplikacyjnych.

Abstract

Brassica napus is one of the world's most important oil crops. Its cultivation faces challenges from several significant *Brassica* diseases, including clubroot, which accounts for an estimated 10% yield loss worldwide. Clubroot is caused by the soil-borne obligate endobiothrophic rhizarian *Plasmodiophora brassicae.* This pathogen stimulates the outgrowth of the infected roots, leading to the development of prominent galls. The affected plants exhibit systemic symptoms, leading to non-optimal flowering and premature senescence.

Due to the high longevity and durability of *P. brassicae* resting spores, controlling the disease proves challenging. Consequently, breeding resistant varieties, which relies on the identification of the genetic factors involved in the response to infection, remains one of the most effective control strategies.

In this study, a clubroot resistance locus from *B. napus* cultivar 'Tosca' is mapped to a genomic region encompassing approximately 100 kbp. The local sequence polymorphism and differential gene expression patterns between clubroot-resistant and susceptible lines are evaluated. Based on the results, the duplicated, highly polymorphic TNL-family gene is designated for functional studies. The ectopic expression of the coding regions of the TNL genes in *Arabidopsis thaliana* strongly supports the hypothesis of their involvement in disease resistance.

Furthermore, a comparative analysis of the resistance locus structure among a broader panel of Brassicaceae genomes is conducted. The results highlight the high evolutionary dynamics of the region, particularly in the case of the TNL gene.

The results of this study provide valuable resources for direct incorporation into resistance breeding schemes and open up many new possibilities for further basic and applicational research.

List of scientific works included in the dissertation

This dissertation is a mixed-format thesis, comprising a published article and additional unpublished results. The published work included in this dissertation is as follows:

Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in *Brassica napus* cv. Tosca. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631. PMID: 33936130; PMCID: PMC8082685.

Ministry points (MNiSW): 100
Impact factor (2021): 6.627

Funding

This work was supported by:

1. The National Science Centre Grant HARMONIA 2016/22/M/NZ9/00604
   awarded to Prof. Wojciech Karłowski
2. The National Science Centre Grant PRELUDIUM 2021/41/N/NZ2/03595
   awarded to Piotr Kopeć
3. Passport to the future - Interdisciplinary doctoral studies at the Faculty of Biology, Adam Mickiewicz University. POWR.03.02.00-00-I022/16.
4. Minigrant from 'Uniwersytet Jutra' POWR.03.05.00-00-Z303/17
   awarded to Piotr Kopeć
5. Doctoral research grant 017/02/SNP/0024 from ID-UB awarded to Piotr Kopeć
6. International doctoral research grant 048/13/UAM/0034 from ID-UB
   awarded to Piotr Kopeć

Abbreviations

ABK	Ancestral *Brassica* karyotype
CNV	Copy number variation
DAI	Days after inoculation
DAMP	Damage-associated molecular pattern
DH	Doubled haploid
ETI	Effector-triggered immunity
FDR	False discovery rate
FP	Fluorescent protein
IGV	Integrative Genomics Viewer
LRR	Leucine-rich repeats
MAMP	Microbe-associated molecular pattern
MAS	Marker-assisted selection
NGS	Next-generation sequencing
NLR	Nucleotide-binding domain and leucine-rich repeat-containing receptor
ONT	Oxford Nanopore Technologies
PAMP	Pathogen-associated molecular pattern
PAV	Presence-absence variability
PRR	Pattern recognition receptor
PTI	Pattern-triggered immunity
QTL	Quantitative trait locus
R gene	Major effect disease resistance gene
ROS	Reactive oxygen species
RT-PCR	Reverse transcription PCR
SNP	Single nucleotide polymorphism
STP6	Sugar transport protein 6
TNL	Toll/interleukin-1 (TIR)-type NLR
WGT	Whole-genome triplication
WT	Wild type

1. Introduction

1.1. Brassica napus - a young species of global importance

Brassicas are among the most economically significant plants, second only to cereals (Dixon, 2014; Warwick, 2011). Since ancient times, they have been cultivated and selectively bred for their edible leaves, flowers, roots, and seeds. Different parts are directly consumed, used as fodder or forage, or processed further - for instance, seeds are used to produce edible and technical oils. Diverse selection criteria resulted in a spectacular array of morphotypes within the most prominent species. For example, *Brassica oleracea* L. has evolved into cabbage, Brussels sprouts, cauliflower, broccoli, kohlrabi, and kale. *Brassica rapa* L. has given rise to European turnips and Far Eastern leafy varieties, like Chinese cabbage, pak choi, and mizuna (Dixon, 2017). *Brassica nigra* L. was cultivated for its leaves and spicy seeds, which in addition to oil production, were used as a condiment mustard (Warwick, 2011).

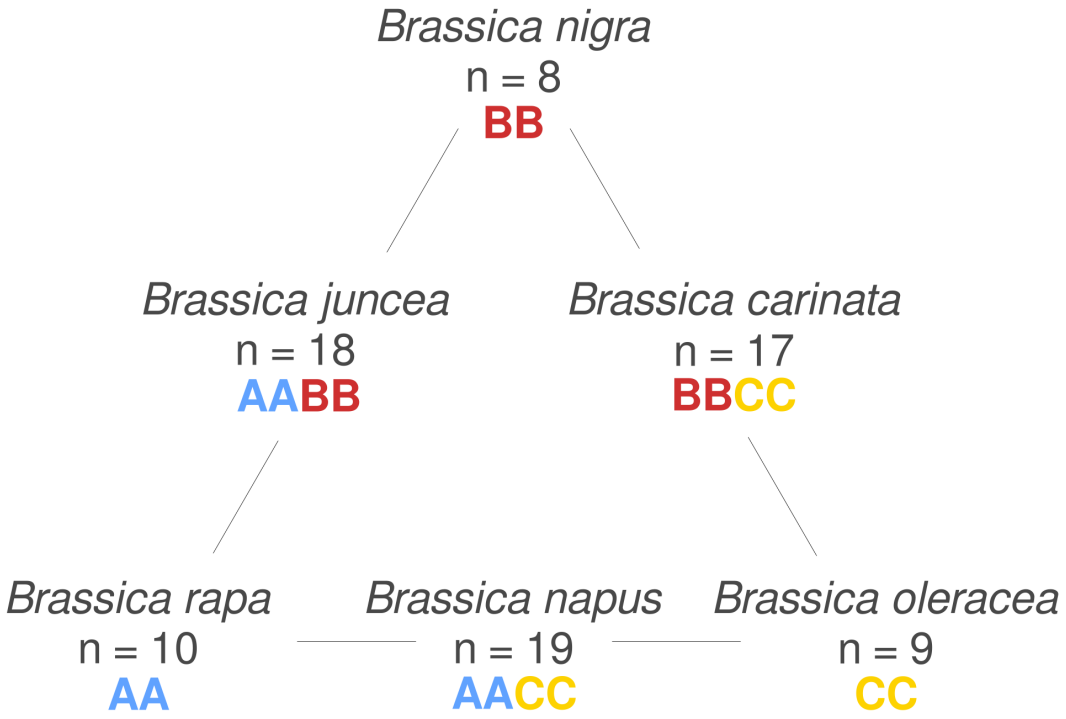


Figure 1. The triangle of U, after U Nagaharu, 1935

These three diploid *Brassica* species - *B. rapa* (n=10, A), *B. nigra* (n=8, B), and *B. oleracea* (n=9, C) - can naturally hybridize. This phenomenon, first described by Korean-Japanese botanist U Nagaharu, is commonly referred to as the 'Triangle of U' [Fig. 1] (U, Nagaharu, 1935). The hybridization resulted in the emergence of three new allotetraploid species: *Brassica juncea* L. (n=18, AB), *Brassica carinata* L. (n=17, BC), and *Brassica napus* L. (n=19, AC).

Brassica napus, the youngest of the three species, originated relatively recently from several independent hybridization events, most likely within the past 7,500 years (Chalhoub et al., 2014; K. Lu et al., 2019; X. Song et al., 2021), although some estimates reach up to 50,000 years (J. Yang et al., 2016). The multiplication of genomic content that results from hybridization is a recurrent feature in the evolutionary history of the *Brassica* lineage (Tang et al., 2018). Ancestral cycles of polyploidization and subsequent diploidization, together with recent allotetraploidization, have cumulatively multiplicated proto-angiosperm genomic blocks present in *B. napus* by up to 72-fold (Chalhoub et al., 2014). At the time of its sequencing, the *B. napus* genome was known to be the most duplicated plant genome, containing the highest number of annotated genes (Chalhoub et al., 2014).

Due to its recent polyploidization, *B. napus* serves as an excellent model for studying post-polyploidy genome evolution. In particular, it provides insight into the mechanisms that stabilize a genome following the 'genomic shock' of polyploidization. These mechanisms involve extensive genomic reorganization, including genetic and epigenetic changes, resulting in widespread gene content and regulation alterations. The changes are often biased towards one of the subgenomes, leading to gene fractionation and genome dominance (Samans et al., 2018; Tang et al., 2018; Ziegler et al., 2023).

The extensive multiplication of genomic content facilitates the incidence of homology-based exchange events between genomic regions, catalyzing significant evolutionary dynamics. Notably, the extensive homology between progenitor subgenomes provides a foundation for homoeologous exchanges. This results in widespread structural variability conferring regulatory changes, variations in gene copy number, including presence-absence variations, as well as gene conversion events (Hurgobin et al., 2018; Samans et al., 2018). Collectively, these factors serve as key drivers of plants' adaptability

and hold potential for crop improvement (Gabur et al., 2019, 2020; Schiessl et al., 2017; Stein et al., 2017; Vollrath et al., 2021).

Aside from the duplications resulting from polyploidization, the *B. napus* genome is characterized by a substantial proportion (48%) of repetitive elements. These elements further contribute to the extensive structural variability observed in the genome (Waminal et al., 2018).

Like its progenitors, *B.napus* has undergone diversifying selection, leading to the emergence of two subspecies comprising several morphotypes. These include *B. napus* ssp. *napobrassica*, embodying taproot-forming swedes or rutabagas, and *B. napus* ssp. *napus,* which includes leafy fodder and kale-like varieties, along with morphologically distinct *B. napus* ssp. *napus var. pabularia*, as well as oilseed rape. Within oilseed rape, two ecological forms have emerged: winter and spring, which differ in their cold hardiness and the need for vernalization (Friedt & Snowdon, 2010).

Oilseed rape has, in recent times, risen to global importance. It is the third most important source of vegetable oil after palm and soybean, and the second most important source of protein meal, following soybean, with its production expected to grow (USDA Foreign Agricultural Service, 2023). Historically, rapeseed oil was primarily used as a lamp oil and later as a technical lubricant. Due to its then-unpalatable taste, it was used for food mainly in times of scarcity. Furthermore, it contained a high content of potentially cardiotoxic erucic acid as well as glucosinolates, which make the protein-rich post-extraction rapeseed meal unsuitable for monogastric animal feed (Daun, 2011; Piętka et al., 2014; Snowdon et al., 2007). The advent of the modern food-grade oilseed rape occurred over the past few decades, following two major advancements: the development of the low-erucic-acid German variety 'Liho' and the low-glucosinolate Polish variety 'Bronowski' (Krzymański, 1968; Stefansson & Hougen, 1961). Those two traits were combined in Canada in 1974 in the first double-low ('00') variety 'Tower', marking the rise of the trademarked 'Canola'-quality oilseed rape (Friedt & Snowdon, 2010; Stefansson & Kondra, 1975).

These introgressions constituted a major genetic bottleneck, reducing the already limited genetic diversity of oilseed rape. In a natural setting, low genetic diversity can hinder adaptability to changing environment. Similarly, in artificial selection, scarce genetic

diversity can restrict the potential for improving traits such as yield quantity and quality, ease of cultivation, or stress response. Numerous breeding efforts have been undertaken to expand the gene pool by utilizing hybridization within the Triangle of U, and beyond (e.g., with *Sinapis* or *Raphanus*), through both natural crossing and techniques like embryo rescue and somatic hybridization (Katche et al., 2019). Such 'resynthesized' *Brassica napus* were the foundation for the development of 'Mendel' and 'Tosca' varieties, which have been created to address the lack of adequate clubroot resistance in '00' oilseed rape (Diederichsen et al., 2006; Diederichsen & Sacristan, 1996; Friedt & Snowdon, 2010).

1.2. Clubroot disease

1.2.1. The pathogen

Clubroot is a disease caused by the infection of *Plasmodiophora brassicae* (Woronin, 1878). Plasmodiophorids are unicellular eukaryotes within Phytomyxea class of the Rhizaria clade. As obligate biotrophs, they infect a wide range of organisms, including brown algae, diatoms, and oomycetes (Neuhauser et al., 2014). However, their primary targets are plants, encompassing several economically important species. *P. brassicae* specifically affects members of the Brassicaceae family, including *B. napus*.

P. brassicae exhibits a complex life cycle consisting of a dormant phase and a host-dependent infectious phase, which can be subdivided into an epidermal primary infection and a cortical secondary infection (Kageyama & Asano, 2009; L. Liu, Qin, Zhou, et al., 2020).

During dormancy, *P. brassicae* persists in the soil as chitin-encapsulated resting spores. These spores are very resilient and long-lived, with a viability half-life of 3.6 years (Wallenhammar, 1996). Under favorable conditions and in response to certain root exudates (Rashid et al., 2013), these spores release biflagellate primary zoospores that invade root hairs and epidermal cells of the elongation zone through a puncture in the cell wall (Aist & Williams, 1971).

Once inside the root, the pathogen enters a primary infection phase, which, intriguingly, is compatible with a broader range of non-host species beyond Brassicaceae

(Dixon, 2014). During this phase, a series of acytokinetic mitoses occur, leading to the development of coenocytic 'plasmodium'. This plasmodium undergoes cytokinesis, forming a zoosporangium that subsequently releases uninucleate secondary zoospores into the lumen of epithelial cells. Despite evidence of meiosis (Garber, 1979), the existence of a sexual stage in *P. brassicae* life cycle has been debated (Dixon, 2014; Javed et al., 2023; Kageyama & Asano, 2009). Current evidence suggests that secondary zoospores conjugate and undergo karyogamy, leading to the formation of a zygote (L. Liu, Qin, Zhou, et al., 2020). This zygote invades the plant's cortical cells, marking the onset of the secondary infection.

The secondary infection is generally specific to Brassicaceae, although some exceptions have been reported (Ludwig-Müller et al., 1999). It largely parallels the primary stage, with the sequential development of a coenocytic plasmodium and, following meiosis, a sporangium. This phase concludes the cycle with the production of haploid resting spores. Unlike the relatively brief primary infection, the vegetative phase of cortical infection extends over a longer period, and during this phase, the characteristic symptoms of clubroot disease develop.

1.2.2. The disease

The reproductive success of *P. brassicae* depends on its ability to acquire nutrients from its host. During the cortical infection, the pathogen manipulates the host's signaling pathways, effectively 'hijacking' certain metabolic modules. These alterations transform the infection site into a carbohydrate sink and stimulate the growth of the subterranean parts of the plant, leading to the development of large galls or 'clubs' [Fig. 2] (Ludwig-Müller et al., 2009; Malinowski et al., 2019).

Figure 2. Clubroot disease on 7-week-old *B. napus* plants. Healthy plant to the left. Affected plants exhibit severely altered root systems.

Gall formation involves a coordinated reprogramming of multiple meristematic tissues (Malinowski et al., 2012, 2019). It initiates with the stimulation of lateral root formation, followed by increased cambial proliferation paired with suppressed differentiation. Subsequent growth is predominantly driven by hormone-induced hypertrophy and endocycling-induced polyploidization (Olszak et al., 2019). The cambial differentiation patterns are also altered, with xylogenesis suppression and stimulation of phloem production (Malinowski et al., 2019; Olszak et al., 2019). The sugar sink is established by modulating specific aspects of carbohydrate logistics. In particular, the proteins that facilitate transport and modulate solubility are impacted, with their expression changed in a way that channels sugar toward the newly created environment, which provides a safe, nutrient-rich space for *P. brassicae* to thrive (Walerowski et al., 2018).

Such drastic alteration of a plant's homeostasis can significantly impact its overall fitness. The redirection of nutrient flow towards the infection site limits resources available for the host. At the same time, structural changes to the roots, including gall formation and reduced xylogenesis, hinder water uptake, causing significant water stress (Javed et al., 2023; Malinowski et al., 2019). In the later stages of severe disease, galled roots disintegrate, releasing spores into the soil. Affected plants exhibit wilting, chlorosis, and reddening of the leaves. The growth is stunted; the plants flower earlier and produce fewer seeds of lower

quality (Dixon, 2009; Korbas et al., 2009). In severe cases, the plants prematurely die (Cao et al., 2009).

1.2.3. Economic cost

Given the importance and widespread cultivation of Brassicaceae, and clubroot's substantial impact on the plant's health, the economic implications of the disease are significant. In oilseed rape production, it ranks among the most critical diseases, along with *Sclerotinia* stem rot and *Phoma* stem canker (Zheng et al., 2020). Globally, the disease is responsible for a 10-15% reduction in yield (Dixon, 2009), with local crop failures occurring in heavily infested fields (Hwang et al., 2012).

Poland is one of the key European producers of Brassicaceae crops, ranking 4th in the EU and 7th worldwide in oilseed rape production. Approximately 20% of all cultivated vegetables in Poland are Brassicas, and 13% of arable land is dedicated to rapeseed. Recent PCR screenings revealed the presence of clubroot in all 16 provinces of Poland, with the highest prevalence in the southwestern and northern regions. The pathogen was detected in 62% of tested fields (Czubatka-Bieńkowska et al., 2020). This widespread prevalence of the disease highlights its significant impact on the agricultural sector and food production, emphasizing the need for exploring effective solutions.

1.2.4. Managing clubroot disease

Managing the disease is a very challenging task. Its prevention is arguably the most effective strategy. Resting spores of *P. brassicae* are primarily dispersed by farm equipment, transfer of infected soil or crops, and to a lesser extent, by water and wind (Cao et al., 2009; Hwang et al., 2014; Rennie et al., 2015). Therefore, the implementation of robust agronomic practices such as field monitoring and proper sanitation is commonly recommended (Hwang et al., 2014; Jajor et al., 2017). However, once established, the disease is difficult to eradicate from the field as viable resting spores can be detected in the soil even 15 years following the last *Brassica* growth (Wallenhammar, 1996).

Nevertheless, some strategies can help manage the disease (Donald & Porter, 2009; Hwang et al., 2014). For instance, regular crop rotations can keep the pathogens population in the soil low. However, the effectiveness of this approach might be limited by the presence of common cruciferous weeds and volunteer plants (Zamani-Noor & Rodemann, 2018). Furthermore, adjusting the seeding dates can mitigate the risk of infection by minimizing the exposure of plants to pathogens during their vulnerable growth stages, taking advantage of *P. brassicae* spores' optimal temperature and humidity ranges for germination (Hwang et al., 2014).

Certain fungicides, calcium cyanamide, and lime application can be helpful, though their effectiveness can be limited in heavily infested soils (Hwang et al., 2014). Additionally, biocontrol measures, such as applying specific *Bacillus* and *Trichoderma* strains, have shown some efficacy in limiting the disease severity (Ahmed et al., 2020). Moreover, given the nonspecificity of the primary infection, using non-host bait plants can also be effective in reducing the spore load. However, the disease-limiting efficiency of such an approach is rather low (Hwang et al., 2014).

Unfortunately, many of the proposed solutions are impractical or not economically viable, especially in the case of large-scale rapeseed cultivation (Hwang et al., 2014). Hence, one of the best strategies to combat the disease is to utilize and augment plants' innate defense mechanisms. Therefore, breeding and employment of resistant varieties, either alone or integrated with some of the methods mentioned above, presents as a primary choice to manage the disease effectively (Diederichsen et al., 2009; Hwang et al., 2014; Jajor et al., 2017).

1.3. Plants' endogenous responses - PTI, ETI

Facing a continuous threat from a multitude of pathogens, plants have evolved a diverse array of defense mechanisms. These include physical barriers such as cell walls, waxes, trichomes, and regulated stomatal openings. Biochemical responses, including the synthesis of phytoalexins and the release of reactive oxygen species, are also employed. In some scenarios, plants resort to localized programmed cell death. Given that some of these defenses are costly, they are deployed only as needed, in response to the activation of an

elaborate danger recognition system composed of two layers - pattern-triggered immunity (PTI) and effector-triggered immunity (ETI) (Ngou et al., 2022; Yu et al., 2017).

PTI serves as the 'first line' of plant immune response, recognizing and responding to specific molecular patterns associated with pathogens or damage, known as PAMPs/MAMPs (pathogen/microbe-associated molecular patterns) or DAMPs (damage-associated molecular patterns). MAMPs encompass common, conserved molecules found in many pathogens, such as chitin or flagellin, while DAMPs are endogenous molecules indicative of cellular damage, such as extracellular ATP, oligosaccharides, or peptides (Boutrot & Zipfel, 2017). These molecular patterns are recognized by the specific ectodomains of transmembrane pattern recognition receptors (PRRs), often acting in association with co-receptors (Boutrot & Zipfel, 2017; Yu et al., 2017).

Upon ligand binding, PRRs initiate a cascade of downstream responses (Ngou et al., 2022; Yu et al., 2017). These include the activation of mitogen-activated protein kinase (MAPK) signaling pathways, calcium influx, reactive oxygen species (ROS) burst, and the synthesis of stress hormones like salicylic acid, ethylene, and jasmonic acid. This signaling results in the implementation of defensive measures, ranging from immediate reactions like stomatal closure or ROS-mediated oxidative crosslinking of cell walls that prevent the pathogen entry, to long-term transcriptional reprogramming, which activates defense-related genes. These can lead to metabolic and structural changes, for example, by inducing the synthesis of phytoalexins or callose deposition in cell walls. In addition to the localized responses, PTI triggers a systemic response known as systemic acquired resistance (SAR), which primes the entire plant for enhanced defense against future pathogen attacks (Vlot et al., 2021).

While PTI provides plants with broad-spectrum defense against a wide range of pathogens, many pathogens have evolved means to surpass this initial response. They secrete proteins, and other molecules called 'effectors' that target and disable certain components of the PTI (Ngou et al., 2022). A fascinating example of such a strategy is the secretion of protein that contains the LysM domain - a chitin-binding module of PRRs. By binding competitively, this protein effectively shields the chitin from recognition by plants' PRRs, thereby suppressing the PTI response (de Jonge et al., 2010).

The recognition and reaction to effectors is the premise of plants' second layer of immunity, namely ETI (effector-triggered immunity) (Ngou et al., 2022). In general, ETI is facilitated by cytoplasmic nucleotide-binding domain and leucine-rich repeat-containing receptors (NLRs) that can be classified into three main types (depending on their N-terminal domain): coiled-coil type (CNLs), Toll/interleukin-1 receptor (TIR)-type (TNLs) and Resistance to powdery mildew 8-like domain type (RNLs) (Ngou et al., 2022). NLRs recognize effectors directly or indirectly by monitoring the state of host proteins targeted by effectors, so-called 'guardees' (Y. Lu & Tsuda, 2021; X. Zhang et al., 2017). Furthermore, in addition to acting alone, some NLRs can act in pairs or more complex assemblies (Adachi et al., 2019).

Although ETI and PTI have historically been regarded as somewhat parallel pathways with overlapping downstream effects, it has since been discovered that ETI does not function independently of PTI. Instead, ETI appears to potentiate the basal PTI response, often to the point of inducing a localized programmed cell death called hypersensitive response (HR) (Bjornson & Zipfel, 2021; Ngou et al., 2021; Yuan et al., 2021).

1.4. Arms race

Any mutation that allows a pathogen to evade recognition by the host or dampen its defense response confers a significant fitness gain and would be likely to spread in the pathogen population. While the MAMPs are relatively evolutionarily 'fixed', effectors are generally disposable and can evolve more dynamically. Conversely, any new mutation that allows for the recognition of an undetectable pathogen would spread among the host population. This constant interplay between pathogen's effectors and plants' receptors drives a case of the evolutionary arms race. As a result, NLRs are the largest, most diverse, and fastest-evolving class of proteins in plants, and the landscape of NLRs has been shaped by extensive copy number variation, structural variation, and gene conversion events (Barragan & Weigel, 2021; Han, 2019; Michelmore et al., 2013). Consequently, although a single plant must rely on its innate defenses, plants' immune system generally exhibits significant adaptability in space and time.

1.5. Pathotypes

From an agricultural perspective, the emergence or spread of strains pathogenic to varieties so far regarded as resistant is of the greatest concern. Such resistance breakdowns are a regular occurrence in various host-pathogen systems. For example, the breakdown of resistance to *P. brassicae* infection in the *B. napus* 'Mendel' variety happened relatively quickly after its introduction to the market (Diederichsen et al., 2014).

Genetically distinct strains of pathogens that can infect specific sets of ecotypes or cultivars are referred to as races or pathotypes. In *P. brassicae,* several race classification systems are used, such as Williams, European Clubroot Differential (ECD), and Somé systems (Buczacki et al., 1975; Somé et al., 1996; Williams & Others, 1966). They involve assessing an infection response of a panel of variably responsive hosts, with four, fifteen, and three accessions, respectively. The geographic distribution of pathotypes is not uniform. For example, in northern Poland, Somé pathotype P1 prevails while P3 is more common in the south (Řičařová et al., 2016). This substantial variability has prompted the development of more locally appropriate differential systems, like Canadian and Sinitic Clubroot Differentials (Pang et al., 2020; Strelkov et al., 2018). Moreover, it necessitates local cultivation of specific, genetically compatible varieties.

Given the high evolutionary dynamics of the host-pathogen interaction, the identification of novel genetic sources of resistance is a key effort in breeding resistant varieties.

1.6. Clubroot resistance

Several genetic sources of clubroot resistance have been found in Brassicaceae, primarily in the genus *Brassica*, including *B. oleracea*, *B. rapa*, *B. napus*, *B. nigra* (Chang et al., 2019; Lv et al., 2020; Neik et al., 2017), as well as in *Raphanus sativus* and *Arabidopsis thaliana* (Fuchs, 1996; Gan et al., 2022; Ochoa et al., 2022). Most of the identified loci come from *Brassica oleracea* and *Brassica rapa*, particularly European fodder turnip varieties (*B. rapa* ssp. *rapifera*).

1.6.1. *Brassica oleracea* - the source of quantitative resistance

Resistance in *B. oleracea* is predominantly quantitative, with broad-spectrum and isolate-specific effects that often act in a recessive manner (Piao et al., 2009). The exact mechanisms of this resistance remain unknown, as no functional studies of *B. oleracea* resistance genes have been published to date. In general, quantitative resistance could arise from adapting various aspects of morphology, metabolism, or basal defense (Pilet-Nayel et al., 2017). For instance, in *B. napus*, C - genome-derived quantitative trait loci (QTLs) have been linked to specific metabolic modules, and partial resistance was found to be associated with higher glucosinolate gluconasturtiin and a decrease of citric acid content in roots (Wagner et al., 2019). In *A. thaliana*, allelic substitutions within clubroot resistance QTLs lead to the upregulation of camalexin synthesis, an important phytoalexin (Lemarié et al., 2015), and increased ability to tolerate exogenous trehalose (Gravot et al., 2011), a key component of successful infection that can reach phytotoxic concentrations in affected tissues (Malinowski et al., 2019).

To date, the only major qualitative resistance gene found in *Brassica oleracea* is Rcr7, which is believed to be introgressed from *Brassica rapa* (Dakouri et al., 2018).

1.6.2. The qualitative resistance of *Brassica rapa*

B. rapa exhibits a qualitative type of resistance determined by strong-effect, race-specific genes. This source of resistance has been extensively utilized in breeding, both intra- and interspecifically (Piao et al., 2009). The cloned and characterized clubroot R genes from *B. rapa* belong to the TNL class, the receptive component of the ETI pathway (Hatakeyama et al., 2013, 2017; Ueno et al., 2012; Z. Yang et al., 2022). While the complete molecular mechanism of their action remains unclear, evidence suggests that the plant's response prevents a secondary stage of infection, during which gall development occurs. For resistant varieties 'ECD04' and 'ECD10', the microscopic appearance of the infection during the primary phase is identical to susceptible infections. However, the cortical invasion and subsequent development of secondary plasmodia are absent (L. Liu, Qin, Cheng, et al., 2020)

1.6.3. Limited resistance in *Brassica napus*

Sources of resistance within the primary genetic pool of *Brassica napus* are known from swede (*B. napus* ssp. *napobrassica*) varieties such as 'Wilhelmsburger', 'York', and 'Ditmars S2' (Ayers & Lelacheur, 1972). Further, the swede's have been enhanced by incorporating resistance derived from 'ECD-04' through resynthesis and recurrent backcrossing (Bradshaw et al., 1997). The swede resistance loci have been mapped in 'Brookfield'-derived 'Rutabaga-BF' (Hasan & Rahman, 2016) as well as 'Wilhelmsburger' and resynthesis-derived 'Invitation' cultivars (Lüders, 2017).

The first clubroot-resistant oilseed rape varieties to reach the European market, 'Mendel' and 'Tosca', were also derived from resynthesis, with resistance introduced from the ancestral species (Diederichsen et al., 2006).

The genetic basis of 'Tosca' resistance was not disclosed and is a main subject of this work. In the case of 'Mendel', resistance was sourced from the turnip variety 'ECD-04'. Although 'ECD-04' contains three dominant genes, only one was stably transferred to 'Mendel' (Diederichsen et al., 2006). This gene was subsequently mapped to the locus on chromosome A03, which corresponds to the previously described and cloned gene *CRa* (Fredua-Agyeman & Rahman, 2016). The Mendel variety was used as a source of resistance for multiple oilseed rape breeding programs worldwide, but the breakdown of this resistance prompted the development of a second generation of clubroot-resistant varieties, with undisclosed genetic sources (Askarian et al., 2021; Diederichsen et al., 2014).

1.7. Use of qualitative and quantitative resistance in breeding

The exploitation of natural quantitative and qualitative resistance introduces unique advantages and challenges. The polygenic nature of quantitative resistance makes its use in breeding more complex, as to reach its full potential, it requires a coordinated transfer of multiple QTLs into the receptive genetic background. However, this polygenic character can enhance the durability of the resistance, as the selection pressure exerted by a single small-effect locus might not be very high, particularly when multiple modalities are involved (Pilet-Nayel et al., 2017). Furthermore, the quantitative effects are often non-specific, conferring resistance to a broader range of pathotypes. Despite these properties, erosion of

quantitative resistance has been observed in some systems (Caffier et al., 2016; Delmas et al., 2016; Montarry et al., 2012; Pilet-Nayel et al., 2017).

In contrast, introducing qualitative resistance, particularly monogenic, is relatively straightforward in simple marker-assisted selection scenarios. However, the risk of a rapid breakdown of such resistance is a significant concern. It is noteworthy, though, that durable, broad-spectrum qualitative resistances exist, as exemplified by the *mlo* gene that has been used to control powdery mildew in barley for over 40 years, although its mode of action is different than that of classical R genes (Büschges et al., 1997; Jørgensen, 1992).

An interesting and widely studied prospect is the 'pyramiding' of resistance genes, involving the combination of quantitative resistance with qualitative resistance. The *Rlm6* gene, which confers resistance to *Leptosphaeria maculans* in *B. napus*, serves as a compelling example - when paired with a quantitatively resistant background, the durability of the resistance increased by more than twofold, as compared to a susceptible background (Brun et al., 2010; Delourme et al., 2014). A related strategy involves stacking multiple race-specific major genes. This approach has been used in clubroot resistance breeding and extended the range of pathotypic specificity and durability (Matsumoto et al., 2012; Tonu et al., 2023). However, such schemes carry an inherent risk of the emergence and spread of multi-resistant pathotypes, as exemplified in the case of *L. maculans* resistance (Balesdent et al., 2022). Simulation studies have shown that the employment of spatial and temporal rotations of resistant varieties (Djidjou-Demasse et al., 2017; Rimbaud et al., 2018) could mitigate this risk.

Regardless of the mode of action and type of resistance, its implementation requires a methodology for identifying the responsible genetic loci and transferring them into the breeding material. The recent advancements in molecular biology techniques and genomics have greatly expanded the possibilities for crop improvement and accelerated research.

1.8. Genomic technologies bridge the gap between basic research and breeding

Agricultural breeding has significantly evolved in recent times, transitioning from phenotypic selection to the use of advanced genomic technologies. Historically, breeding involved intentional or inadvertent actions that promoted the relative fitness of plants with desirable traits, like yield, taste, ease of cultivation, and resistance to pests or diseases. With the advent of genetics in the early 20th century, more systematic approaches emerged, such as population and pedigree breeding. This led to a vast expansion of high-yielding, stress-tolerant crop varieties. Nevertheless, these approaches were constrained by a focus on phenotypes, which take time to develop fully and often are intercorrelated and polygenic (Breseghello & Coelho, 2013).

The development of molecular biology techniques transformed this landscape. Notably, introducing restriction enzyme and PCR-based assays allowed for the identification of sequence variants between individuals. This genotypic information was then used to create the linkage maps of cosegregating markers, subsequently establishing the link between phenotypic traits and specific polymorphisms in genetic sequences. These 'genetic markers' could, in turn, be used to fix a trait of interest in breeding material with limited phenotyping effort. Such marker-assisted selection (MAS) greatly accelerated the breeding process and played a pivotal role in modern practices. Initially, the markers were sparse, and their development was labor-intensive, which limited the throughput and resolution. However, further advancements introduced technologies that enabled accessible genome-wide genotyping of many individuals, leveraging the extensive genomic single nucleotide variability. It was achieved primarily through the development of SNP chips and sequencing technologies.

The first whole genome sequencing projects, based on Sanger's chain termination method, had low throughput and were extremely costly. It limited their application to human and model organisms, including, fortunately for the *Brassica* research, *Arabidopsis thaliana*. The emergence of second-, or next-generation sequencing (NGS), based on a massive simultaneous reading of short DNA fragments, revolutionized the field. Within three years of its introduction, sequencing costs dropped by a factor of 1000 (and ten-fold more since then), facilitating a rapid and broad emergence of genomic and transcriptomic projects (Kris A. Wetterstrand, 2019). *Brassica rapa* was the first sequenced, assembled, and annotated

Brassica genome (*Brassica rapa* Genome Sequencing Project Consortium, 2011), followed by *Brassica oleracea* (S. Liu et al., 2014) and *Brassica napus* (Chalhoub et al., 2014). Reference genomic sequences facilitated the examination of population variability through resequencing and helped expand our knowledge of the coding and regulatory context of the loci of interest. This, in turn, greatly facilitated functional research on genetic factors impacting phenotypes.

The reference sequences led to the development of the *Brassica* 60k genotyping chip, which spans approximately 50 thousand SNPs across A and C genomes (Clarke et al., 2016). This development significantly reduced the genotyping cost per sample in relation to resequencing. Although the sequencing cost can be significantly alleviated using reduced representation techniques, such as skim sequencing or genotyping by sequencing (GBS) (Deschamps et al., 2012; Golicz et al., 2015), working with sequencing data is not standardized and requires expert knowledge and significant computational resources. Employment of a fixed array of genotypic loci sacrificed genomic resolution but allowed straightforward analysis for predefined sites, enabling the identification of single nucleotide, and some structural variants (Mason et al., 2017). The brassica *Brassica* 60k, or its 15k subset, was subsequently used to create high-density genetic maps and precisely map traits, including clubroot resistance, within the crosses as well as genome-wide association studies (GWAS) (Fredua-Agyeman et al., 2020; Li et al., 2016).

Apart from allowing precise mapping of sources of mono- and oligogenic traits for use in marker-assisted selection and functional studies, the attainability of genome-wide coverage of genotypic information allows its utilization as a selection criterion for improving polygenic traits in genomic selection scenarios (Crossa et al., 2017). Such an approach could potentially be used in the case of quantitative resistance breeding (Poland & Rutkoski, 2016).

1.8.1. Long reads help resolve structural variability

The short-read, second-generation sequencing technology faces inherent limitations. The length of the Illumina-based sequence fragments leads to ambiguity in the case of repeated sequences. It is especially problematic in highly repetitive plant genomes, including *B. napus* (Chalhoub et al., 2014). The development of third-generation, long-read sequencing technologies, such as PacBio and Oxford Nanopore Technologies (ONT), addresses this

problem. They enabled more continuous and correct genome assemblies and better assessment of sequence polymorphism. It is especially evident in the case of medium-sized structural variation, which is widespread in *B. napus* and linked to trait variability but was obscured by the length of NGS reads (Chawla et al., 2021). Currently, the relative ease of obtaining high-quality long-read-based genomic assemblies leads to the pan-genomic era (Bayer et al., 2020; Golicz et al., 2016; Hurgobin et al., 2018; J.-M. Song et al., 2020). The graph-based representations of genomic variation within populations make it easier to dissect extensive structural variability, especially prominent in the case of disease resistance genes (Dolatabadian et al., 2020).

Another frontier in modern breeding is genetic engineering. Although traditional breeding methods employed random mutational breeding, the real revolution came with the development of CRISPR-Cas9 and its optimization for use in plants. This system allows for precise, transgene-free mutagenesis and homology-driven gene transfer, significantly reducing production time (Chen et al., 2019). Unfortunately, using CRISPR-based transgenesis outside basic research is illegal in the European Union. However, cisgenesis is currently under discussion; therefore, the regulatory landscape may change, enabling new opportunities (EFSA Panel on Genetically Modified Organisms (GMO) et al., 2022).

All these developments are exciting and offer huge prospects for basic research as well as the development of better genotypes. They are shaping the future of food production, challenged by climate change and population growth. Many of the presented technologies facilitated the research presented in this work.

2. Aims of the thesis

This Ph.D. thesis was nested within the project funded by National Science Centre grant 2016/22/M/NZ9/00604 led by Prof. Wojciech Karłowski - a joint effort of Adam Mickiewicz University Poznań, The Plant Breeding and Acclimatization Institute - National Research Institute Poznań, Institute of Plant Protection - National Research Institute Poznań and Justus-Liebig-University Gießen. The focus of the project was to elucidate the genetic bases of clubroot resistance in *Brassica napus* 'Tosca' with the aim to expand the understanding of this host-pathogen system, as well as implement molecular markers to facilitate the development of novel clubroot-resistant *Brassica napus* varieties.

The primary objective of this thesis was to identify the genetic loci associated with clubroot resistance in the *Brassica napus* cultivar 'Tosca'. The results of this effort were published in Frontiers in Plant Science (Kopec et al., 2021), and the article is included as a third chapter of this thesis. Identifying the resistance locus created an opportunity for further exploration and characterization of its gene content, as well as sequence and functional differences between the resistant and susceptible genotypes. These analyses included the identification of genetic polymorphism and clubroot-induced gene expression variability, enabling the recognition of candidate resistance factors. The identification of a potential resistance factor laid a solid foundation for the experimental evaluation of its role in clubroot resistance. Finally, the striking structural and sequence polymorphism between the studied genotypes, specifically concerning the potential resistance factor, prompted a broader comparative analysis of the resistance locus composition across a larger panel of Brassicaceae genomes.

3. Identification and characterization of clubroot resistance locus in *B. napus* 'Tosca'

3.1. Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca

Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM.

Frontiers in Plant Science 2021; doi: 10.3389/fpls.2021.639631

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 April 2021
doi: 10.3389/fpls.2021.639631

Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in *Brassica napus* cv. Tosca

*Piotr M. Kopec[1], Katarzyna Mikolajczyk[2], Ewa Jajor[3], Agnieszka Perek[3], Joanna Nowakowska[2], Christian Obermeier[4], Harmeet Singh Chawla[4,5], Marek Korbas[3], Iwona Bartkowiak-Broda[2] and Wojciech M. Karlowski[1]**

[1] Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University Poznan, Poznan, Poland, [2] Department of Genetics and Breeding of Oilseed Crops, Plant Breeding and Acclimatization Institute-National Research Institute, Poznan, Poland, [3] Institute of Plant Protection - National Research Institute, Poznan, Poland, [4] Department of Plant Breeding, Justus-Liebig-Universitaet Giessen, Giessen, Germany, [5] Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

OPEN ACCESS

Edited by:
Jacqueline Batley, University of Western Australia, Australia

Reviewed by:
Harsh Raman, New South Wales Department of Primary Industries, Australia
Maoteng Li, Huazhong University of Science and Technology, China

***Correspondence:**
Wojciech M. Karlowski
wmk@amu.edu.pl

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *09 December 2020*
Accepted: *09 March 2021*
Published: *08 April 2021*

Citation:
Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I and Karlowski WM (2021) Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca. Front. Plant Sci. 12:639631. doi: 10.3389/fpls.2021.639631

Clubroot, caused by *Plasmodiophora brassicae* infection, is a disease of growing importance in cruciferous crops, including oilseed rape (*Brassica napus*). The affected plants exhibit prominent galling of the roots that impairs their capacity for water and nutrient uptake, which leads to growth retardation, wilting, premature ripening, or death. Due to the scarcity of effective means of protection against the pathogen, breeding of resistant varieties remains a crucial component of disease control measures. The key aspect of the breeding process is the identification of genetic factors associated with variable response to the pathogen exposure. Although numerous clubroot resistance loci have been described in *Brassica* crops, continuous updates on the sources of resistance are necessary. Many of the resistance genes are pathotype-specific, moreover, resistance breakdowns have been reported. In this study, we characterize the clubroot resistance locus in the winter oilseed rape cultivar "Tosca." In a series of greenhouse experiments, we evaluate the disease severity of *P. brassicae*-challenged "Tosca"-derived population of doubled haploids, which we genotype with Brassica 60 K array and a selection of SSR/SCAR markers. We then construct a genetic map and narrow down the resistance locus to the 0.4 cM fragment on the A03 chromosome, corresponding to the region previously described as *Crr3*. Using Oxford Nanopore long-read genome resequencing and RNA-seq we review the composition of the locus and describe a duplication of TIR-NBS-LRR gene. Further, we explore the transcriptomic differences of the local genes between the clubroot resistant and susceptible, inoculated and control DH lines. We conclude that the duplicated TNL gene is a promising candidate for the resistance factor. This study provides valuable resources for clubroot resistance breeding programs and lays a foundation for further functional studies on clubroot resistance.

Keywords: ***Brassica napus*, *Plasmodiophora brassicae*, Oxford Nanopore, TNL, RNA-Seq, QTL, resistance, duplication**

INTRODUCTION

Plasmodiophora brassicae Wor., an obligate, soil-borne parasite of crucifers (Brassicaceae), is an agent responsible for clubroot disease. During the two-stage infection (Kageyama and Asano, 2009; Liu et al., 2020), the pathogen hijacks multiple nodes of host metabolism and induces hyperplasia and hypertrophy of the underground organs leading to a prominent galling. The galls act as a major physiological sink that supports the proliferation and development of the pathogen while reducing the fitness of the host (Malinowski et al., 2019). Deformations of the root system impair the plant's capacity for water and nutrient uptake, leading to growth retardation, wilting, and premature, non-optimal flowering (Korbas et al., 2009). Many important crops cultivated worldwide, including oilseed rape (*Brassica napus*), belong to the Brassicaceae (Dixon, 2007). Clubroot disease has been becoming a global problem of increasing economic impact in cruciferous crops and has been ranked under the top 10 most significant worldwide threats to oilseed rape production (Dixon, 2009; Zheng et al., 2020). An infection of oilseed rape was shown to cause up to 60% loss of yield at relatively low spore densities, and total yield failure at a higher pathogen pressure (Strehlow et al., 2015). Once introduced, *P. brassicae* is hard to eradicate. Resting spores can live in the soil for up to 20 years (Wallenhammar, 1996), and spread easily via, for example, dirt on farm equipment (Cao et al., 2009). Many protective measures against the pathogen, e.g., crop rotation or chemical control, are of limited efficiency (Hwang et al., 2014). Therefore, the breeding of resistant plant varieties remains a crucial component of clubroot control efforts.

The key aspect of the breeding process is the identification of genetic features associated with plant response to pathogen exposure. Numerous clubroot resistance loci were described in *Brassica* crops (Landry et al., 1992; Figdore et al., 1993; Voorrips and Visser, 1993; Grandclément and Thomas, 1996; Voorrips et al., 1997; Matsumoto et al., 1998; Moriguchi et al., 1999; Suwabe et al., 2003, 2006; Hirai et al., 2004; Laurens and Thomas, 2004; Piao et al., 2004; Rocherieux et al., 2004; Nomura et al., 2005; Saito et al., 2006; Sakamoto et al., 2008; Nagaoka et al., 2010; Kato et al., 2012; Chen et al., 2013; Chu et al., 2013, 2014; Hatakeyama et al., 2013, 2017; Pang et al., 2014, 2018; Zhang et al., 2014; Fredua-Agyeman and Rahman, 2016; Lee et al., 2016; Huang et al., 2017; Yu et al., 2017; Dakouri et al., 2018; Hirani et al., 2018; Nguyen et al., 2018; Peng et al., 2018; Laila et al., 2019; Choi et al., 2020; Farid et al., 2020; Karim et al., 2020) and are reviewed in (Neik et al., 2017; Lv et al., 2020).

Several genetic studies were performed in *B. napus*. Manzanares-Dauleux et al. (2000) described a major resistance gene *Pb-Bn1* on the A04 chromosome and two quantitative loci on A04 and C05 chromosomes. The resistance derived from the DH ECD-04 line (selected from *Brassica rapa* subsp. rapifera), which was utilized in many breeding programs for the development of clubroot-resistant cultivars, including winter oilseed rape "Mendel," was mapped to the *CRa/CRb* region on the A03 chromosome (Diederichsen and Sacristan, 1996; Diederichsen et al., 2006; Fredua-Agyeman and Rahman, 2016; Zhang et al., 2016). Werner et al. (2008) mapped 19 QTLs spread across 8 chromosomes. In addition, a couple of association studies were conducted on the *B. napus/P. brassicae* pathogenic model. Li et al. (2016) identified 9 loci, 7 of which were not described previously. Hejna et al. (2019) identified 2 major and 7 minor loci, with the most prominent peak overlapping the *CRa* region. Fredua-Agyeman et al. (2020) identified three genomic hotspots corresponding to *Crr3/CRk/CRd* and *CRa/CRb/CRb*^{Kato} regions on A03 and *Crr1* region on A08 in a GWAS study of 124 rutabaga accessions from Nordic countries.

Additionally, two resistance genes were cloned thus far: *CRa* (Ueno et al., 2012) and *Crr1* (Hatakeyama et al., 2013). Both genes belong to the TIR-NBS-LRR (TNL; Toll/interleukin-1 receptor-like – nucleotide-binding site – leucine-rich repeat) protein domain family, reported as a key component of effector-triggered immunity (DeYoung and Innes, 2006; McHale et al., 2006).

Despite a seemingly ample collection of resistance loci, continuous updates on the sources of resistance are necessary. *P. brassicae* shows pathogenic specialization, and the host's resistance genes often confer immunity to only subsets of pathotypes. Moreover, the breakdown of clubroot resistance in the case of some *P. brassicae* pathotypes has been repeatedly reported (Diederichsen et al., 2014; Strelkov et al., 2016).

In this study, we map the resistance locus of the Swedish resynthesis-derived winter-type oilseed rape cultivar "Tosca" (Happstadius et al., 2003; Diederichsen et al., 2009) to a small region on the A03 chromosome. Using the long-read Oxford Nanopore (ON) sequencing technology, we review the genomic structure of the locus in "Tosca" as well as in susceptible "BRH-1" breeding line. In addition, we perform an RNA-seq experiment to identify infection-induced differentially expressed genes. These data are subsequently linked to the genic composition of the resistance locus. Based on the results, we attribute the "Tosca" resistance phenotype to a locus constitutively expressing a duplicated TNL gene, located within the *Crr3* (Hirai et al., 2004) region, directly upstream of the region homologous to the *CRd* (Pang et al., 2018). This study provides valuable resources for clubroot-resistant rapeseed breeding programs and lays a foundation for further functional studies on clubroot resistance.

MATERIALS AND METHODS

Plant Material

A doubled haploid (DH) segregating population of 250 DH lines was developed by Plant Breeding Strzelce Ltd. (IHAR-PIB Group; division in Borowo) from a cross of a winter oilseed rape (*B. napus*) clubroot resistant cultivar "Tosca" and a susceptible BRH-1 breeding line, using isolated microspore culture technique as described in (Cegielska-Taras et al., 2002; Szała et al., 2020).

Pathogen Source, Preparation, and Plant Inoculation

Samples of *B. napus* root galls induced by *P. brassicae* were collected from infested oilseed rape fields in Lower Silesian Province in Poland. The inoculum for the greenhouse

experiments was prepared by isolating resting spores from the galls. The galls were blended, and the homogenate was filtered through a layer of gauze and centrifuged for 5 min at 3,500 rpm to obtain a clear suspension. Spore density was measured using a 0.1 mm deep, improved Neubauer counting chamber (Marienfeld-Superior) and a bright field microscope (Olympus BX 50). The density was adjusted to 1×10^8 spores/ml. For inoculation, each experimental pot containing five 1-week-old seedlings was watered with the spore suspension. The same batch of inoculum was used in all experiments. Additionally, to assess the *P. brassicae* pathotype, galls from 25 DH lines were collected and individually processed into a set of spore suspensions. *P. brassicae* pathotype of every suspension was classified using the Somé system (Some et al., 1996).

Experimental Design and Conditions

The greenhouse experiments were performed between April 2018 and August 2019 in the Research Centre of Quarantine, Invasive and Genetically Modified Organisms – Institute of Plant Protection National Research Institute. The experiment followed the principle of augmented design. The plants were grown in a series of 6 temporally successive blocks (batches). Each of the batches included around 60 test DH lines, augmented with 8 reference lines ("checks") – 6 phenotypically extreme DH genotypes that were selected from the first experiment and parental lines.

For every line, 15 plants were grown in 3 pots: 2 pots for treated (inoculated) and 1 pot for untreated control, 5 plants each. Pots were randomly distributed in 4 trays for treated and 2 trays for untreated control plants. Separate, fixed trays were used for treated and untreated plants to avoid water or soil-borne contamination. The soil pH value was 6.0. The temperature (±0.5°C) was set to 18°C/16°C day/night regime for the first 2 weeks of cultivation, and then elevated to 20°C/18°C. The photoperiod was set to a 14 h/10 h light/darkness scheme. The air humidity (±3%) was 60%. Soil humidity was kept in the range between 60 and 70%.

Despite the controlled experimental conditions, we observed a significant batch effect – seasonal phenotypic variability among the analyzed DH lines. Therefore, an additional experiment was carried out including more lines in one, common batch (242, including the checks) at the expense of the number of tested plants per line (5 instead of 10). Additionally, to promote the infection by *P. brassicae*, the temperature was elevated to 20°C for the first 2 weeks and 24°C/20°C for the next five.

Phenotyping and Phenotypic Data Analysis

After 7 weeks of growth (42 days after inoculation), the plants were phenotyped for classical underground morphological symptoms of clubroot disease. Each plant was removed from the ground and washed with water. The degree of infection (DOI) was evaluated on a 4-degree scale (Vigier et al., 1989), where 0 indicates a healthy root system, 1 refers to 1–10% of root system altered (small galls on lateral roots), 2 denotes 11–50% root system altered, and 3 describes 51–100% root system altered. The disease index (DI) for each genotype by batch was then calculated by obtaining the arithmetic mean of the DOI and rescaling it to the percent scale.

To obtain the DI over the entire experiment, adjusted for the batch effect (phenotypic variability between the greenhouse runs), the DOI data were fit into a linear mixed model using the lme4 library (Bates et al., 2015) for R (R Core Team, 2020):

$$P_{ij} = \mu + g_i + B_j + (gb)_{ij} + e_{ij} \quad (1)$$

where P_{ij} stands for the phenotype of the *i*th genotype in the *j*th batch, μ is the general mean of the experiment, g_i is the random effect of the *i*th genotype, B_j is the fixed effect of the *j*th batch, $(gb)_{ij}$ is the random effect of the interaction between the *i*th genotype and *j*th batch, and e_{ij} is the error term. Next, the conditional mode (Best Linear Unbiased Prediction; BLUP) of the genotype was obtained. The BLUP-DI values were used in a subsequent QTL mapping.

Heritability Estimation

Broad sense heritability (H^2) of the DOI was estimated after (Stahl et al., 2019) following the concept of (Piepho and Möhring, 2007), with the equation:

$$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + SE^2} \quad (2)$$

where σ_G^2 is the genetic variance, derived from a full random model (Eq. 3) and SE^2 is the squared standard error of the difference between the means, derived from a mixed model (Eq. 4). The analysis was conducted using the R packages lmerTest (Kuznetsova et al., 2016), lsmeans (Lenth, 2016), and lme4 (Bates et al., 2015).

$$P_{ij} = \mu + g_i + b_j + (gb)_{ij} + e_{ij} \quad (3)$$

where P_{ij} stands for the phenotype of the *i*th genotype in the *j*th batch, μ is the general mean of the experiment, g_i is the random effect of the *i*th genotype, b_j is the random effect of the *j*th batch, $(gb)_{ij}$ is the random effect of the interaction between the *i*th genotype and *j*th batch, and e_{ij} is the error term.

$$P_{ij} = \mu + G_i + B_j + (gb)_{ij} + e_{ij} \quad (4)$$

where P_{ij} stands for the phenotype of the *i*th genotype in the *j*th batch, μ is the general mean of the experiment, G_i is the fixed effect of the *i*th genotype, B_j is the fixed effect of the *j*th batch, $(gb)_{ij}$ is the random effect of the interaction between the *i*th genotype and *j*th batch, and e_{ij} is the error term.

To investigate the reliability of each of the batches, their influence on the H^2 was assessed by recalculating the H^2 with a leave-one-out approach.

Genotyping

The plants were genotyped using The *Brassica* 60 K *Illumina Infinium*™ SNP array (Clarke et al., 2016) and a set of SSR and SCAR markers of known clubroot resistance loci (**Supplementary Table 1**). For *Brassica* 60 K genotyping, the plant

material collected from young leaves was sent to the commercial service provider TraitGenetics in Gatersleben (Germany) for DNA isolation and further processing. For SSR/SCAR analysis, the DNA was extracted from young leaves using a modified CTAB method (Doyle and Doyle, 1990). PCR amplification products were visualized on a 1.5% agarose gel (SCAR) and using the ABI PRISM 3130 OXL capillary electrophoresis (SSR).

Filtering of Genotyping Data

To check for duplicates, the lines were clustered with complete linkage based on Jaccard's distance. Lines with <0.05 distance were regarded as duplicates, and only one of them (randomly selected) was used in further analyses. Lines with more than 0.02% of heterozygous calls were discarded. Homomorphic (>95%) markers and markers with distorted segregation patterns (1:3) were also removed from further analyses. Redundant markers were binned.

Genetic Map Construction and QTL Mapping

A genetic map was constructed using the R/qtl package (Broman et al., 2003). For ordering the markers, the R/TSPmap program was used (Monroe et al., 2017). QTL Mapping was conducted with Haley-Knott regression implemented in the scan1 function of the R/qtl2 (Broman et al., 2019) package. log10(p) significance cutoff was determined using a permutation test with n = 1,000.

Genome Sequencing of Parental Lines

Genomic DNA from parental lines was sequenced using ON technology. DNA from young leaves was extracted following a protocol described by Chawla et al. (2020). The libraries were prepared using the SQK-LSK109 kit, following the manufacturer's recommendations, and sequenced on R9.4.1 Flow Cells.

Genome Sequencing Data Analysis

The *B. napus* reference genomes used for the study were: Darmor-*bzh* v4.1 (Chalhoub et al., 2014), deposited on EnsemblPlants as AST_PRJEB5043_v1; Express 617 assembly v1 (Lee et al., 2020); reference pan-genome v0 (Song et al., 2020). Darmor-*bzh* genes within the mapped resistance locus were functionally classified using Pannzer2 (Törönen et al., 2018).

Base calling from ON signals was performed using Guppy, and raw reads were mapped to the reference genomes with minimap2 (Li, 2018) with -x map-ont parameters, and filtered for uniquely mapping reads with samtools (Li et al., 2009) using -q 60 option. Local SNV calling was performed using longshot (Edge and Bansal, 2019), with default parameters. SV calling was executed using sniffles (Sedlazeck et al., 2017), with –min_support 5 option. The potential effect of the variants differing parental accessions was determined with the SnpEff (Cingolani et al., 2012).

The reads overlapping the TNL gene on the Express 617 *B. napus* genome assembly were extracted from the raw sequence file, assembled using Redbean (wtdbg2; Ruan and Li, 2020), and polished once using the wtpoa-cns tool.

Transcriptome Sequencing

For transcriptomic experiments, one resistant and one susceptible DH line were selected. The roots of two biological replicates per line of infected and control plants were harvested on the day of phenotyping (7 weeks after inoculation), immediately frozen in liquid nitrogen, and stored at −80°C. The tissue was blended in liquid nitrogen using a mortar and pestle. The total RNA was extracted with the Qiagen Plant RNeasy kit. TruSeq mRNA strand-specific libraries were prepared and sequenced on Illumina NovaSeq600 in a 2 × 150 bp paired-end layout. Library preparation and sequencing were conducted by Macrogen.

Reconstruction of the TNL Genes-Encoded Transcripts Using RNA-Seq Reads

RNA-seq reads were pooled by DH line and mapped to the Express 617 genomic sequence assembly supplemented with the fragment containing the duplication as a pseudochromosome using STAR (Dobin et al., 2013) with the following parameters: –outFilterMismatchNoverLmax 0.1 –outFilterMismatchNoverReadLmax 0.1 –alignIntronMax 2000 –alignIntronMin 15 –outSAMprimaryFlag AllBestScore. Reads mapping to the pseudochromosome were then assembled using Trinity (Grabherr et al., 2011) with -genome_guided_bam –genome_guided_max_intron 2000 options. Assembled transcripts were re-mapped to the Express 617 reference sequence supplemented with a pseudochromosome with a minimap2 -x splice for validation. ORFs were predicted and translated using NCBI's ORFfinder[1]. For sequence comparison, the CDS and protein sequences were aligned with EMBL-EBI's Clustal Omega and EMBOSS Needle (Madeira et al., 2019). The sequence-based prediction of protein domains was carried out with InterProScan (Jones et al., 2014).

Analysis of Differential Gene Expression

Raw RNA-seq reads were trimmed using Trimmomatic (Bolger et al., 2014) with default options and mapped to the reference genome using STAR with the following parameters: –outFilterMismatchNoverLmax 0.1 –outFilterMismatchNoverReadLmax 0.1 –alignIntronMax 2000 –alignIntronMin 15. The fragments were counted using the featureCounts (Liao et al., 2014) program with *-s* 2 *-p -M* flags. The differential expression analysis was performed using limma (Ritchie et al., 2015; Law et al., 2016)/edgeR (Robinson et al., 2010) R packages, following the procedure described by Law et al. (2016). Raw counts were normalized via TMM, and log-CPM values were used for the DE analysis. The fit was processed with limma's treat() with lfc = 1 parameter, thus genes with fold-change significantly larger than 2 were deemed as differentially expressed. Gene Ontology enrichment analysis was conducted with g:Profiler (Raudvere et al., 2019). To assess the expression of the "Tosca" TNL copies, the reads were mapped to the reference with the fragment containing the duplication

[1] https://www.ncbi.nlm.nih.gov/orffinder/

attached as a pseudochromosome, and TPM values of TMM normalized counts were calculated.

Creation of Figures

All plots were generated with r/ggplot2 (Wickham, 2009). Figures were assembled with Inkscape[2].

RESULTS

Phenotyping of "Tosca" × "BRH-1" DH Population

To identify the locus harboring resistance to clubroot disease in the "Tosca" winter oilseed rape cultivar, a mapping population of 250 DH lines was developed in a cross with a susceptible, double-low line BRH-1. For phenotyping experiments, the lines were divided into 7 groups and tested separately in controlled environmental conditions. In addition to the tested DH lines, each experimental batch contained a set of referential "checks" – both parents and 6 DH lines used in all experiments (**Figure 1**, colored lines). These 6 lines were identified in the first experiment as showing contrasting phenotypes (resistant or susceptible) and no visible developmental abnormalities. In each of the experiments, 1-week-old seedlings were inoculated with a suspension of *P. brassicae* spores prepared from the pathotype-P3-dominant environmental sample. After a total of 7 weeks of growth, the plants were examined for infection-induced morphological pathologies of roots and evaluated on a 4-step severity-dependent scale. For every line, a percent scale DI was calculated.

Analysis of the first 6 batches clearly demonstrated that despite random grouping of the lines the DI distribution was not equal between the batches (Kruskal–Wallis test, $p < 10–6$), with mean DI ranging from 13.13 to 46.86% (**Figure 1A** and **Table 1**). Phenotypes of the 8 checks were shifting according to the batches' trend, showing that the observed differences were independent of the selection of the lines in each of the subsets (**Figure 1A**). The differences are well explained by seasonal changes, with a higher incidence observed in batches carried out during spring and summer, despite that the plants were grown in controlled greenhouse conditions.

Since the main goal of this experiment was to identify DH lines with the strongest resistant phenotype, we decided to repeat the tests using growing conditions that promote *P. brassicae* infection. Therefore, in the last 7th batch (**Figure 1B**), we have included most (240 out of 250) of the DH lines and increased the temperature by 2°C (for details see section "Materials and Methods"). In all 7 batches, the DI values followed a clear bimodal distribution, suggesting that the majority of the phenotypic effect is linked to a single locus. To adjust the DI for the batch effect for QTL mapping, the phenotyping data were fit to a linear mixed model and the BLUP of the genotypic effect (BLUP-DI) was obtained (**Figure 1C**). The estimated broad-sense heritability of the disease severity for the entire experiment was $H^2 = 0.729$.

TABLE 1 | Summary statistics of the DI of the seven batches and BLUP-DI.

Batch	Min	Max	Mean	Median	Q_1	Q_3	SD	H^2_{LOO}	ΔH^2
1	0.00	90.48	30.77	36.67	0.00	55.56	28.87	0.70	0.03
2	0.00	100.00	46.84	57.143	0.00	83.33	40.02	0.65	0.08
3	0.00	50.00	13.11	6.70	0.00	26.70	15.29	0.79	−0.03
4	0.00	62.96	18.89	16.67	5.36	30.56	14.79	0.77	−0.04
5	0.00	92.59	44.23	58.33	0.00	66.67	32.29	0.73	0.00
6	0.00	100.00	38.59	40.00	0.00	73.30	36.02	0.77	−0.02
7	0.00	100.00	63.84	100.00	13.30	100.00	43.97	0.52	0.21
BLUP-DI	−6.45	77.22	37.15	46.87	11.71	56.92	24.32		

H^2_{LOO} is the H^2 without given batch; $\Delta H^2 = H^2 - H^2_{LOO}$.

Identification of the Resistance Locus by Genetic Mapping

The mapping population was genotyped using the *Brassica* 60 k SNP array (Clarke et al., 2016) and a set of SSR and SCAR

[2]Harrington et al. http://www.inkscape.org/

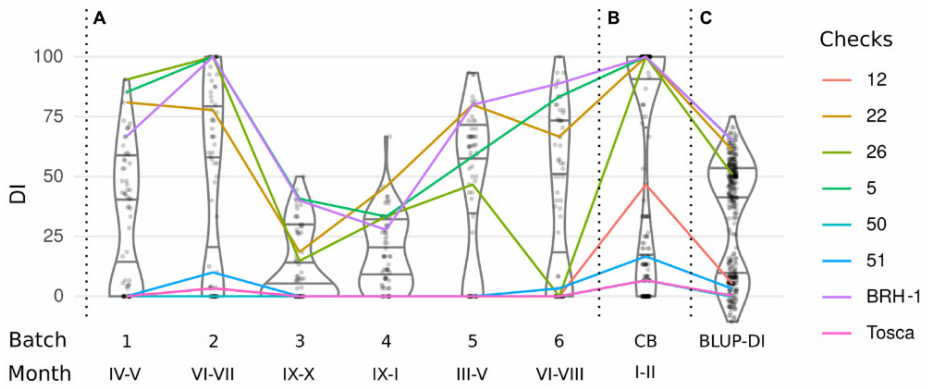


FIGURE 1 | Distribution of the Disease Index (DI) in **(A)** a series of phenotyping experiments, **(B)** joint phenotyping of 240 lines in conditions promoting pathogenesis, and **(C)** a cumulative estimation of the DI-BLUP. The color lines connect DI's of 8 checks tested in each experiment.

markers linked to various clubroot resistance *loci*. Markers showing segregation distortion or high heterozygosity were discarded from further analysis. Segregating "Failed" SNP calls were regarded as potential presence-absence variants (Gabur et al., 2018) and kept in the analysis. The constructed genetic map consisted of 1,406 bins of cosegregating markers distributed among 19 linkage groups corresponding to the 19 chromosomes of *B. napus*. The total length of the map was 1866.1 cM with an average spacing of 1.3 cM and a max spacing of 37.2 cM (**Supplementary Table 4**).

QTL mapping on BLUP-DI data revealed a single locus on the A03 chromosome (**Supplementary Table 5**). Bayes Credible Interval (BCI) for the QTL spanned 0.4 cM between 11.980 and 12.378 cM on the genetic map (**Figure 2A** and **Supplementary Table 5**). The same region, although with a larger BCI span, was detected in individual QTL mappings for every phenotyping batch (**Supplementary Table 6**). No evidence suggested the involvement of other loci affecting the trait. The locus exhibited a large effect, with a 45.65 difference between mean values of BLUP-DI for the bin of markers exhibiting the strongest correlation with the phenotype (**Figures 2B,C** and **Supplementary Table 5**). Recombination events in the proximity of the QTL were identified and compared with the phenotypes. This analysis revealed that the state of a single bin of markers, cosegregating with the representative Bn-A03-p15102212 marker at 11.980 cM was sufficient to explain the phenotype (**Figure 2D**).

To physically anchor the resistance locus, either probe or primer sequences (depending on the marker type) from within the bin with the strongest correlation with the trait were aligned to the Darmor-*bzh* 4.1 reference genome. The closest markers flanking the bin spanned a region of 91,088 bp on supercontig LK031800. However, three of the markers cosegregating with the peak marker mapped to a supercontig LK033659, suggesting that the reference genome has been misassembled.

To resolve this discrepancy, we have mapped both contigs (LK031800 and LK033659) to the recently published *B. napus* reference pangenome (Song et al., 2020) and Express 617 assemblies (Lee et al., 2020). As shown in **Figure 3**, both contigs map to the same region in the newer long-read-based genomic sequences. The LK031800 (2,713,116 bp) contig maps to two distinct distant parts of the reference sequence that are separated by region corresponding to LK033659 (51,625 bp) and a small

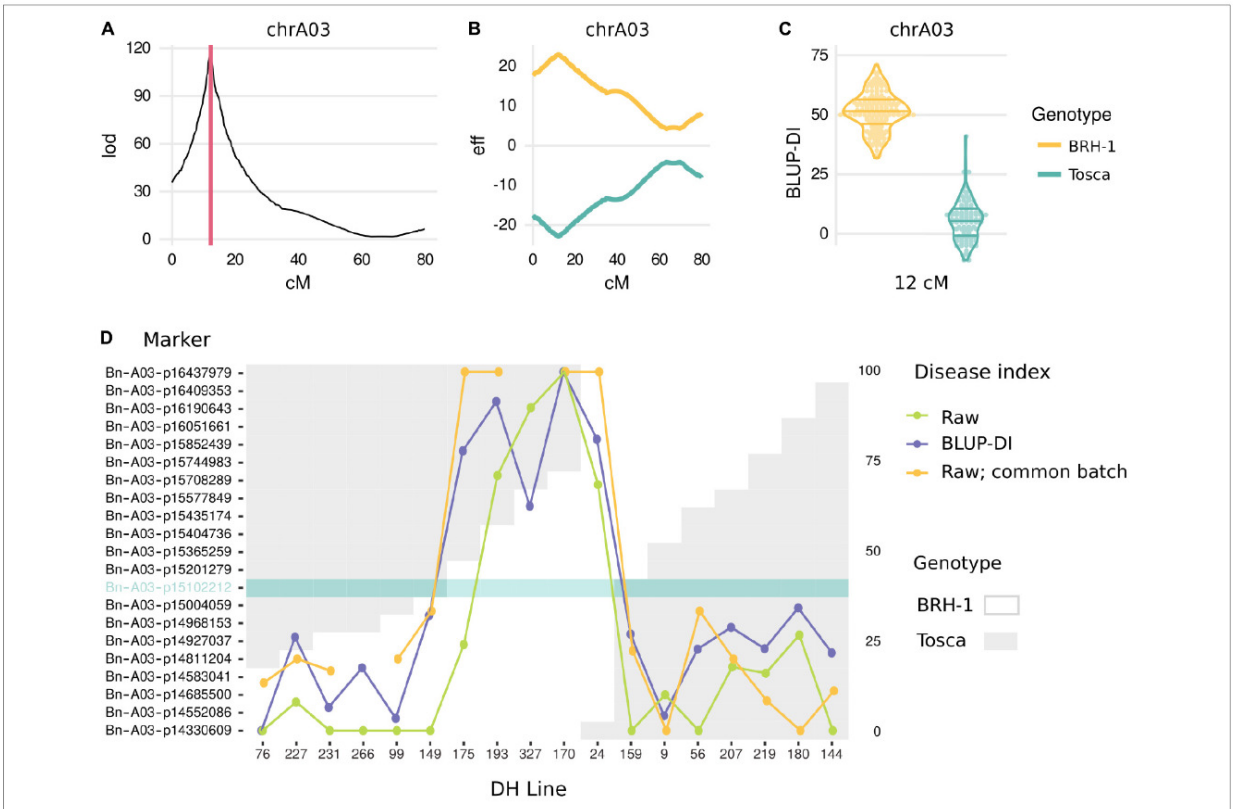


FIGURE 2 | Results of genetic mapping. **(A)** LOD score for QTL presence along the A03 chromosome genetic map. Red lines span Bayes Credible Intervals. **(B)** Effect of "BRH-1" (yellow) and "Tosca" (green) alleles along the A03 chromosome genetic map. **(C)** Phenotype (BLUP-DI) distribution of DH lines carrying "BRH-1" yellow and "Tosca" green alleles at the peak marker. **(D)** Analysis of recombinants. The plot shows a genotype (white: "BRH-1"-inherited, gray: "Tosca"-inherited) at markers surrounding the mapped locus. The phenotype for 18 DH lines recombining in the proximity of the locus is shown as a DI from one of the 1–6 batches, and, if available, 7th, common batch, as well as BLUP-DI. A marker solely explaining the phenotype is highlighted in teal.

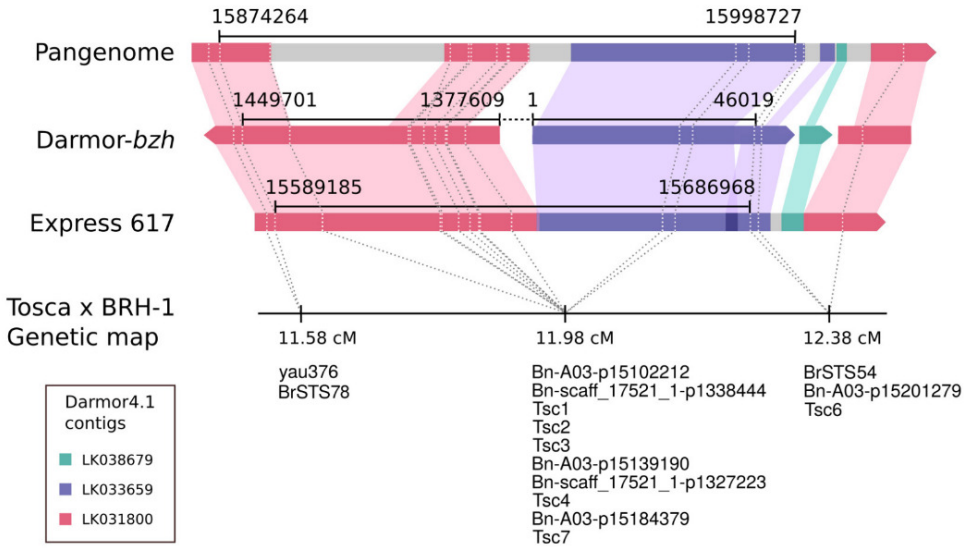


FIGURE 3 | Physical localization of the resistance locus (black line) in different assemblies of the *Brassica napus* genome defined by the bin of cosegregating markers at 11.98 cM. Colored bars represent Darmor-*bzh* 4.1 contigs. Gray color depicts sequences in the pangenome and Express 617 not covered by the Darmor-*bzh* 4.1 contigs.

contig LK038676 of 4,535 bp in size. For further verification of the Darmor-*bzh* being misassembled, a new set of 6 SCAR markers was designed upstream, within, and downstream of the LK033659 (Tsc, **Supplementary Table 2**). The PCR results confirmed the observed segregation pattern and were in perfect agreement with the primers' physical location. Summarizing, the locus defined by the bin of markers (with the representative marker Bn-A03-p15102212) is covered by both new reference *B. napus* assemblies as well as two contigs from Darmor-*bzh* 4.1, but in the latter case, one of the contigs (LK031800) is misassembled (**Figure 3**).

Concluding, the genetic factor of resistance to *P. brassicae* infection is located in the region covering 124,463 bp on the reference pangenome, 97,783 bp on the Express 617 assembly, and 118,111 bp on the Darmor-*bzh* 4.1 assembly. Genetic and physical evidence suggests that the mapped resistance locus falls within the region homologous to the *Crr3* locus (Hirai et al., 2004), directly upstream of the region homologous to the *CRd* (Pang et al., 2018; **Figure 4**). Accordingly, the resistance locus identified in this study is hereafter referred to as $Crr3^{Tsc}$. The $Crr3^{Tsc}$ contains 25 annotated protein-coding genes. The full list with functional descriptions is included in **Supplementary Table 7**.

Structural Variation Within the Resistance Locus Between "Tosca" and "BRH-1"

To explore in detail the properties of the region covering resistance in the "Tosca" genetic background, we have sequenced the genomes of the parental lines with ON technology. The reads mapped to the Express 617 reference genomic sequence consistently overlapped the resistance locus genomic region (97,783 bp) with an average read coverage of 16.24 for "Tosca" and 28.50 for "BRH-1," with 93.6 and 91.9% of positions covered with at least 5 reads, respectively (**Figure 5A** and **Supplementary Figure 1**).

The long-read mapping results showed differences between the "Tosca" and reference genome assembly. Interestingly, the read coverage depth of the fragment overlapping a TNL gene (BnaA03g29300D) was approximately doubled in comparison to the surrounding sequences in "Tosca," but not in "BRH-1" (**Figure 5A** and **Supplementary Figure 1**). Moreover, neither of the reads spanned the entire gene and both of its flanking regions, and many of them mapped twice to the gene. To further investigate these observations, we have *de novo* reassembled this fragment using exclusively long reads mapping to the gene. The new assembly, supported by 20× average coverage and multiple span-through reads, revealed a 7 kb duplication in the "Tosca", but not "BRH-1" genome (including a full copy of the TNL gene – described below; yellow box in **Figure 5B** and **Supplementary Figure 2**).

The duplication identified in "Tosca" was further confirmed using a pair of primers flanking the polymorphic site (TD1_F1/TD1_R, **Supplementary Table 2**) that generate different product lengths for "BRH-1," both duplicated "Tosca" TNL paralogs and their homeolog from the C genome in "Tosca" and "BRH-1" (C genome homeologs are nearly identical in both lines; **Supplementary Figure 3**). PCR reactions performed

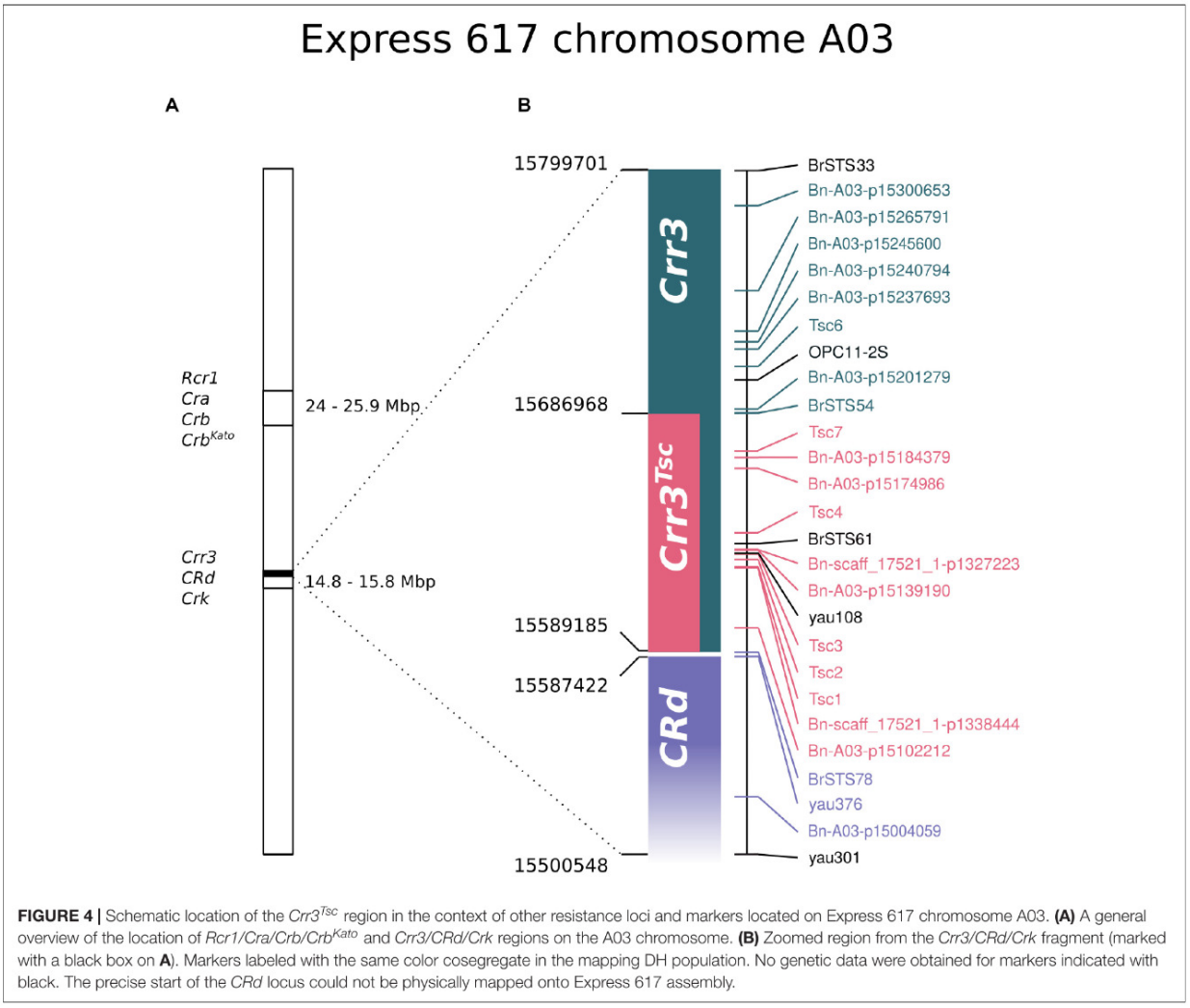


FIGURE 4 | Schematic location of the *Crr3*^{Tsc} region in the context of other resistance loci and markers located on Express 617 chromosome A03. **(A)** A general overview of the location of *Rcr1/Cra/Crb/Crb*^{Kato} and *Crr3/CRd/Crk* regions on the A03 chromosome. **(B)** Zoomed region from the *Crr3/CRd/Crk* fragment (marked with a black box on **A**). Markers labeled with the same color cosegregate in the mapping DH population. No genetic data were obtained for markers indicated with black. The precise start of the *CRd* locus could not be physically mapped onto Express 617 assembly.

on the parents and 11 DH lines with varying degrees of infection resistance revealed the expected pattern of bands, with both "Tosca"-specific alleles segregating with the resistance phenotype (**Figure 6**).

The duplicated region covers the entire TNL gene and a fragment homologous to STP6 **Figure 5B** and **Supplementary Figure 2**. In Darmor-*bzh*, the STP6 fragment is annotated as a distinct gene (BnaA03g29290D), while in Express 617 assembly, the TNL and STP genes are merged into a single entity (A03p030030.1_BnaEXP). Neither the TNL nor the STP fragment is annotated in the scaffoldA03 of the pangenome.

Our RNA-seq data analysis suggests that in both copies the TNL and STP fragment form a single transcription unit; however, the CDS terminates before reaching the STP fragment. The duplicates are separated by a 12 kbp spacer, containing two ∼1.2 kbp transcribed, spliced regions. Blast search of genomic and transcriptomic sequences of both transcribed fragments did not provide any conclusive results. Transcript variants of the transcribed regions contained fragmented ORFs (up to 117 aa in length) with limited similarity to known proteins.

Additionally, the duplicated genes (including the STP fragment) differ in their splicing structure, producing transcripts with 8 and 9 exons, respectively. They share more than 90% identity on the genomic sequence level, with most differences situated downstream of the TNL encoding ORF. On the transcript level, the similarity is 90.9% (4529/4982). The protein sequences are identical in 94.9% (1051/1108) and similar in 96.3 (1067/1108). The alignment has 13 gaps (1.2%), with 10 located at the very end.

Sequence Variation Within Resistance Locus

The location and molecular effect of variants differing between "Tosca" and "BRH-1" were assessed using SNPeff with reference to the Express 617 assembly. The region covering resistance

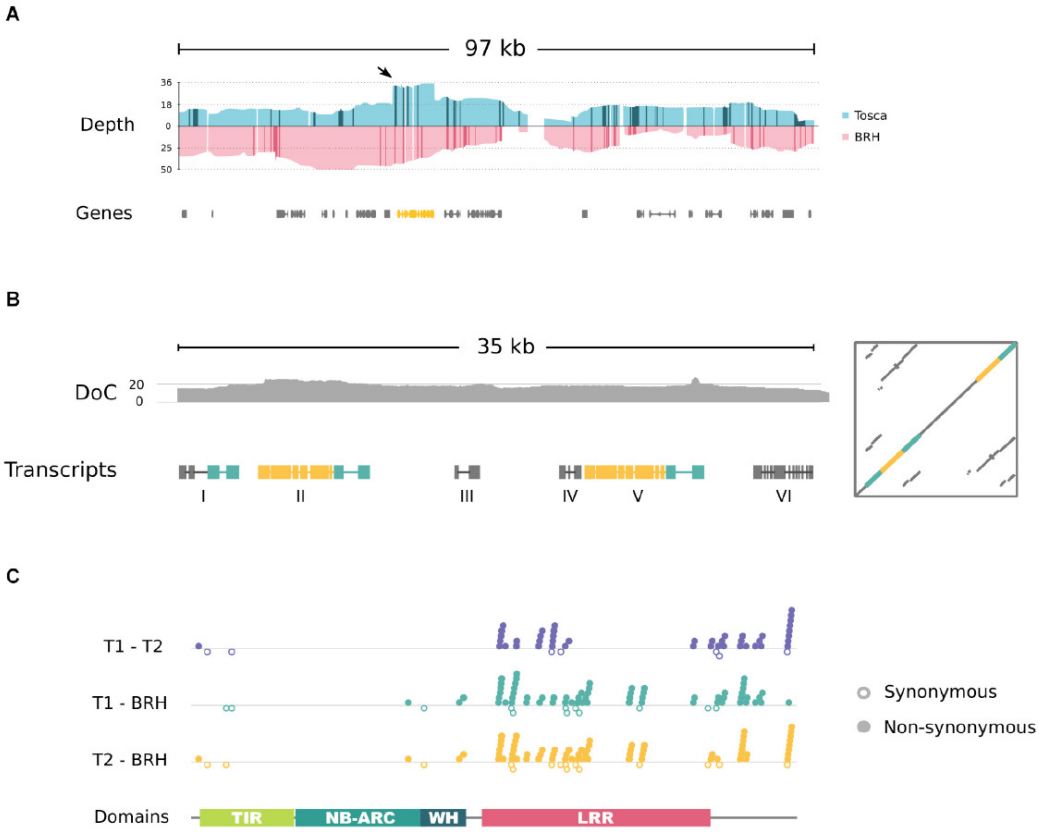


FIGURE 5 | (A) Simplified plot showing resistance locus coverage of ONT long reads aligned to the Express 617 reference genome. A sharp, flat peak (indicated by an arrow) in "Tosca" genomic data marks the location of the fragment containing the TNL gene (yellow). Darker colors on the coverage plot represent insertions larger than 10 bp. **(B)** Schematic overview of (i) depth of coverage of ON reads (DoC) and (ii) within the duplicated "Tosca" region. Yellow – a fragment corresponding to the BnaA03g29300D gene; green – homologous fragments of STP6 gene. **(C)** Pairwise alignment-based comparison of CDS sequence variability between TNL homologs in "Tosca" (T1,T2) and "BRH-1." Synonymous substitutions are marked as empty circles below the axis, non-synonymous are indicated by filled circles above.

to *P. brassicae* infection contains 1521 polymorphic sites: 1327 SNPs, 61 indels, 1 small duplication, and 132 mixed-type variants. The large, duplicated region, covering the TNL gene found in the "Tosca" genome, was omitted from the SNPeff analysis and evaluated separately.

Most of the detected variants are located within intergenic regions. Exonic and intronic polymorphisms account for 6.15 and 7% of the total sequence variability, respectively. The non-synonymous/synonymous substitution ratio is 0.51. The molecular effect, as defined by SNPeff, was high for 0.25%, moderate for 2.07%, and low for 4.62% of the variants. The remaining differences were classified as modifiers. Both genotypes showed the presence of insertions and deletions, located mainly in introns and intergenic regions. Sequences of insertions larger than 500 bp did not show similarity to any annotated genes. A more detailed, gene-oriented analysis of the SNP effect revealed that among protein-coding genes with detectable expression in at least one of the studied lines, only two genes, namely A03p030010.1_BnaEXP (Darmor-*bzh* BnaA03g29270D) and A03p030120.1_BnaEXP/A03p030120.2_BnaEXP (Darmor-*bzh* BnaA03g57410D), contain high-effect variants. A03p030120.1_BnaEXP/A03p030120.2_BnaEXP encodes a metallochaperone and carries a splice donor variant in the "Tosca" cultivar. Besides, this gene harbors the highest variability, with 29 missense and 18 synonymous differences between the lines. A03p030010.1_BnaEXP encodes a chaperonin and contains a premature stop codon in the "BRH-1" (**Supplementary Table 8**).

To reduce the effect of potential ambiguous mapping, the coding sequences of the duplicated "Tosca" TNL genes were compared based on a transcript assembly. The predicted CDS contained a large number of variants between the "BRH-1" gene and both copies from "Tosca" (**Figure 5C**). Interestingly, as noted before, the "Tosca" paralogs differ considerably at the sequence and gene structure levels (**Figure 5C**, T1-T2; **Supplementary Figure 2**).

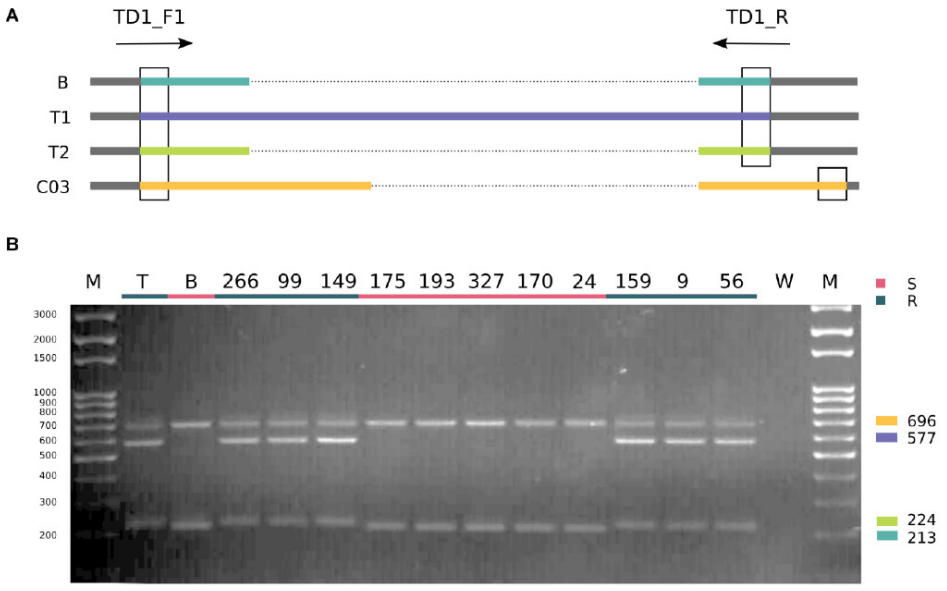


FIGURE 6 | Allele-discriminating PCR. The gel electropherogram **(B)** shows a result of PCR reaction designed as depicted on the schematic **(A)**. B, T1, and T2 stand for BnaA03g29300D copies from "BRH-1" and "Tosca," respectively. C03 represents a homoeologous region on the C03 chromosome, identical between the parental lines. Expected products are color-coded according to the schematic, and juxtaposed with the gel. Gel lanes: M – size marker, T – "Tosca," B – "BRH-1," W – water-containing control reaction, 266-56 – selected lines, recombining in the proximity of the resistance locus. Susceptible lines are underlined with red, resistant with teal. Both "Tosca" alleles (fragments 224 bp and 577 bp) segregated with the resistance.

The *C*-terminal coding fragments have different lengths (60 bp and 30 bp). The lack of 30 bp in the T1 gene results in a frameshift(s) and, consequently, in a different amino acid sequence at the *C*-end of the encoded protein. Apart from the *C*-terminal variance, the sequences differ with regard to 6 synonymous, 36 non-synonymous substitutions, and 9 in-frame deletions (6 and 3 bp long). 35 of the missense variants cluster within and in close vicinity to the LRR domain coding sequence, with the rest of the protein sequence differing only at one position near the *N*-terminus (**Figure 5C**).

The "BRH-1" TNL homologous gene has the same *C*-terminal composition as the "Tosca" paralog T2 (**Figure 5C**, T2-BRH). The BRH and T2 genes differ with regard to 1 in-frame deletion (3 bp), 12 synonymous, and 63 non-synonymous substitutions. On the other hand, the BRH and T1 genes, besides the *C*-terminus variance, differ with regard to 15 synonymous and 59 non-synonymous substitutions (**Figure 5C**, T1-BRH). Similar to the T1–T2 comparison, nearly all differences between "BRH-1" and "Tosca" gene copies are localized near the *C*-terminus and in the LRR domain coding sequence. The cDNA, CDS, and protein sequences of the "Tosca" and "BRH-1" genes are available in the **Supplementary Material**.

Differential Gene Expression Analysis of Resistant and Susceptible DH Lines

To further characterize the 25 genes located within the locus harboring resistance to clubroot disease, we have examined the differences in transcript levels between inoculated and non-inoculated control roots from resistant and susceptible DH lines in the context of the global pattern of differentially expressed genes.

RNA-seq comparison of transcript accumulation between non-inoculated control resistant versus susceptible plants showed a differential signal for 1,247 genes, with only one located within the resistance locus – the BnaA03g29270 gene (**Supplementary Tables 9,10**). This gene shows a significantly higher (logFC = 1.9, adjusted p-value $\leq$ 0.0017) expression level in the resistant line. The gene encodes for a homolog of an *Arabidopsis thaliana* chaperone protein (CCT3).

Subsequently, we have explored differentially expressed genes in roots 46 days after inoculation (DAI) with *P. brassicae*, using a resistant and susceptible line from the mapping population.

In the case of the resistant line, we have identified 111 genes that showed differential transcript accumulation after inoculation (91 up- and 20 down-regulated genes; **Supplementary Tables 9, 10**). 53 of these genes were differentially expressed only in the resistant line. Most of them fall into three general Gene Ontology classes: chitin metabolism, regulation of growth, and defense response (**Supplementary Table 11**). None of the differentially expressed genes identified in the resistant line was located within the resistance locus. The nearest gene showing a differential expression pattern – BnaA03g28780D (encoding Hevein-like preprotein, reported to be involved in the defense response against fungi and bacteria) – is located 200 kbp upstream from the locus.

Analysis of the inoculated susceptible line revealed a much high number (6821) of differentially expressed genes (**Supplementary Tables 9, 10**). Among them, 2,778 were up- and 4,043 were down-regulated. The Gene Ontology-based assignment showed a much more diverse spectrum of molecular functions, among others: oxidative stress response, carbohydrate metabolism, lignin metabolism, chitin metabolism, defense response, auxin signaling (**Supplementary Table 12**). The KEGG pathway enrichment analysis yielded significant hits for phenylpropanoid biosynthesis, biosynthesis of secondary metabolites, glutathione metabolism, stilbenoid, diarylheptanoid and gingerol biosynthesis, and flavonoid biosynthesis.

Four of the differentially expressed genes were located within the resistance locus: down-regulated germin-like protein (BnaA03g29240D – ortholog of AtGLP8, AT3G05930), up-regulated Sugar Transporter Protein (BnaA03g29310D – ortholog of AtSTP6, AT3G05960), down-regulated Fantastic Four protein (BnaA03g57340D – ortholog of AtFAF4, AT3G06020), up-regulated protein trichome birefringence-like (BnaA03g57390D – ortholog of AtTBL10, AT3G06080). None of them, however, were differentially expressed in the resistant line. Moreover, their expression levels were similar in resistant and susceptible control, non-inoculated plants.

The transcript levels of the TNL gene BnaA03g29300D remained unchanged for "BRH-1" and both "Tosca" copies in both lines after inoculation; however, the "BRH-1" and "Tosca" copies were expressed at relatively high levels in the control and the inoculated plants. The transcript levels of the two "Tosca" copies detected in the resistant line added up to twice the amount of the transcript level of one copy expressed in the susceptible line (**Supplementary Table 13**).

DISCUSSION

Based on genetic mapping of a population of 250 DH plants, we were able to identify a single locus conferring resistance to clubroot disease in the winter oilseed rape cultivar "Tosca." The "Tosca" resistance has a different background than the widely utilized "ECD-04," introgressed into the "Mendel" cultivar (Diederichsen and Sacristan, 1996; Diederichsen et al., 2006; Fredua-Agyeman and Rahman, 2016), which makes the source relevant in the *B. napus* breeding efforts.

The identified "Tosca" resistance locus, designated as $Crr3^{Tsc}$, in *B. napus* is located on the A03 chromosome within a previously described *Crr3* locus described in *B. rapa* (Hirai et al., 2004; Saito et al., 2006), which together with *CRk* (Sakamoto et al., 2008; Matsumoto et al., 2012), and *CRd* (Pang et al., 2018) forms a larger cluster of clubroot resistance genetic factors. This cluster has been recently spotted in a GWA study in a panel of *B. napus* ssp. *napobrassica* (rutabaga), which, like "Tosca", are of Nordic origin (Fredua-Agyeman et al., 2020). The region of ∼750 kbp identified in this GWA study was associated with resistance to *P. brassicae* pathotypes 2B and 8P (classified according to Canadian Clubroot Differential Set), which are subsets of Some's P2 pathotype (Strelkov et al., 2018). Here, we show that a region of ∼120 kbp of the $Crr3^{Tsc}$ locus explains the resistance to field isolates consisting of a mixture of pathotypes with the highest prevalence of P3.

The locus was anchored based on a single bin of 11 marker sequences to a region of 97,783 bp of the *B. napus* Express 617 genome assembly. Within the locus, we have identified 25 protein-coding genes. 13 of them were found to be constitutively expressed at the late stage of infection and 4 were found to be differentially expressed between contrasting susceptible and resistant lines of the mapping population. Some of these genes show a functional annotation that makes them interesting candidates to be involved in various stages of *P. brassicae* infection.

Resistance might be expressed constitutively or induced after the initial infection with the pathogen. The non-inoculated control plants showed significant differences in constitutive gene expression patterns, but only 1 out of 1247 was located within the mapped resistance locus. The differentially expressed gene BnaA03g29270D, a homolog of *Arabidopsis thaliana* chaperone protein CCT3, does not offer a direct and evident connection to the mechanism of plant resistance and is unlikely to be involved in resistance expression. The analysis of gene expression affected by the interaction with the pathogen at the later stage of infection provided more interesting candidates. For example, BnaA03g29310D gene which is a homolog of AtSTP6. STPs are monosaccharide/H + symporters that mediate the transport of monosaccharides from the apoplast into the cells (Büttner, 2010). We have observed a significant upregulation of STP6 in infected roots of susceptible, but not of resistant plants. STP family genes, namely STP8 and STP13, have previously been reported to be up-regulated upon clubroot infection in *A. thaliana* (Walerowski et al., 2018), while STP4, STP12, STP1 showed a differential expression pattern in an infected, clubroot-susceptible *Brassica oleracea* cultivar CS-JF1 (Zhang et al., 2019). However, as the expression was affected only in the susceptible line, this gene might have been up-regulated in response to a successful transformation of plant metabolism by the clubroot pathogen during the invasion of the roots. Thus, a potential resistance effect would have to be driven by the inhibition of the expression induction. A similar phenomenon might be responsible for the expression of the gene BnaA03g57390D, harbored within the resistance locus, encoding a homolog of Trichome Birefringence Like 10 (TBL10) protein. This gene was up-regulated in the roots of susceptible, infected plants, but remained constant in the resistant line. TBL proteins are modifiers of the cell wall (Bischoff et al., 2010; Yuan et al., 2016; Gao et al., 2017) and AtTBL10 was found to be involved in O-acetylation of pectin (Stranne et al., 2018). In many reports, pectin hypoacetylation has been linked to increased disease resistance (reviewed in Pauly and Ramírez, 2018), which may explain the lack of TBL10 induction in clubroot defense reaction. Another down-regulated gene from the resistance locus, BnaA03g57340D, is a member of the FANTASTIC FOUR (FAF) protein family. Its expression was down-regulated in susceptible, but not in resistant plants. Overexpression of FAF members was shown to inhibit root growth, which could be rescued by exogenous sucrose (Wahl et al., 2010). Thus,

FAF might perform a role in integrating auxin and sugar signaling during infection progression, allowing the pathogen to manipulate the physiological processes of susceptible plants for more efficient infection progression. Another differentially expressed gene from the resistance locus which has been described to be involved in disease resistance expression is BnaA03g29240D – a homolog of Germin-like protein 8 (GLP8). GLPs are well established as an important component of the biotic stress response (Dunwell et al., 2008; Ilyas et al., 2016). In *B. napus,* GLPs are involved in oxidative burst initiation during *Sclerotinia sclerotiorum* infection (Rietz et al., 2012). GLP5 was also found to have higher expression in a line of *B. rapa* carrying the *Rcr1* clubroot resistance gene (Song et al., 2016). In our study, GLP8 expression was highly reduced in the roots of susceptible plants, though it remained constant in the resistant plants.

Because the expression analysis did not reveal a clear, dominant candidate gene located in the resistance locus that could be responsible for the resistance in "Tosca" we have further explored the properties of the mapped genomic fragment by sequencing the parental cultivars with Oxford Nanopore technology. Detailed analysis of the sequencing data showed a high level of polymorphism on a single nucleotide as well as a larger scale. Despite a large number of differences between the parental lines, most of them were located within the non-coding (genic and non-genic) regions. Additionally, the biological impact of the majority of the polymorphisms was predicted to be low in most of the 25 genes. However, the mapping of the ON reads revealed one striking difference – a large duplication in "Tosca" that covered a full copy of the BnaA03g29300D gene. The duplicated gene contains TIR, NB-ARC, and LRR domains, thus belonging to the TNL subclass of NLR genes. Many of these genes are known to be involved directly or indirectly in the recognition of pathogen effector molecules and initiation of downstream defense responses (reviewed in: Dubey and Singh, 2018; de Araújo et al., 2019). As these genes are known to be involved in the very early stages of signaling cascades, they potentially could be differentially expressed in the early stages of the infection process. The identification of a recent copy of the TNL genes to some extent conforms with this presumption. Assuming that recently duplicated genes retain their original function, we may speculate that the effect of enhanced resistance to pathogen infection in "Tosca" is linked with cumulatively elevated expression (2 times) of two copies of the TNL genes. We cannot, however, exclude an alternative possibility that the new copy of the gene acquired new specificity toward the particular *P. brassicae* pathotypes or that both copies are involved in a more complex resistance initiation (see later). So far, two clubroot resistance genes have been cloned, *Crr1a* (Hatakeyama et al., 2013) and *Cra/CRb* (Ueno et al., 2012; Hatakeyama et al., 2017), both encoding TNL proteins.

The identification of the genomic fragment corresponding to the region defined by the peak marker for the resistance locus was not straightforward using the Darmor-*bzh* 4.1 genome assembly. The coverage of the region was not complete and one of the three contigs mapping to this fragment was misassembled. The locus, however, was correctly placed on the long-read-based reference pangenome and Express 617 assemblies. The Darmor-*bzh* 4.1 reference assembly was constructed prior to the advance of long-read sequencing technologies mainly based on Illumina short-read sequencing and is thus highly fragmented (Lee et al., 2020). We have to note that between the submission and publication of this article, an upgraded, Oxford Nanopore-based version of Darmor-bzh genome (v10) was published, in which the region in question is assembled in agreement with the results of our study and other long-read assemblies (Rousseau-Gueutin et al., 2020). Moreover, it has been shown that in *B. napus* more than 50% of known RGA copies do not occur in a single reference genotype assembly but require analysis of pangenome assemblies for detection (Dolatabadian et al., 2020). Thus, our analysis represents a typical example of how the use of long-read sequencing technology and pangenome sequence assemblies allows more efficient dissection of plant disease resistance loci.

Frequent duplications and clusterization of resistance-related NLR genes are a well-established phenomenon (reviewed in de Araújo et al., 2019; van Wersch and Li, 2019). BnaA03g29300D is flanked by homologous STP6 sequences. This configuration may have served as a foundation for homology-dependent duplication events, for example, by unequal crossing-over. Remarkably, both TNL gene copies seem to be functional, i.e., neither underwent pseudogenization. Both are constitutively expressed in 7-week-old plants and the transcripts contain full-length ORF's. Importantly, the copies harbor a large proportion of polymorphic, non-synonymous sites observed between "Tosca" and "BRH-1," nearly all of which lay within the pattern-recognizing LRR domain, implicating a strong positive selection acting on this domain. Positive selection promoting rapid sequence changes in the NLR genes, especially within LRR domains has already been frequently reported (Bergelson et al., 2001; Mondragón-Palomino et al., 2002; Yang et al., 2013; Karasov et al., 2014; Gao et al., 2018; Han, 2019). LRR domains are reported to be the major factors determining recognition specificity (reviewed in Padmanabhan et al., 2009); therefore, the apparent differences in the amino acid sequence domain may be responsible for the capability of the "Tosca" to sense the *P. brassicae* elicitors and induce the downstream defense response. Remarkably, the two tandem paralogs in the "Tosca" genome themselves differ significantly in the amino acid sequence of their LRR domains. The differences may allow for a broader range of elicitor recognition or the coordination of a more complex response to infection. The TNL proteins are known to engage in functional homo- and heterodimerization (Williams et al., 2014) and various modes of entanglement in the defense response initiation, which often involves clustered genes (reviewed in de Araújo et al., 2019; van Wersch and Li, 2019). Nonetheless, in the case of previously described clubroot resistance locus *Cra/Crb/CRb*^{kato}, which consists of at least six tandemly repeated NLR genes, a majority of the resistance effect is attributed to a single gene, with residual, unexplained effect (Hatakeyama et al., 2017).

Furthermore, the BnaA03g29300D gene is a homolog of *B. rapa* Bra001175, which has been shown to have a higher

level of expression in the clubroot-resistant, compared to the susceptible genotype of the *CRd*-carrying line during the early stages of the infection process, at day 13 after inoculation (Pang et al., 2018). In our study, using RNA-seq data, when each of the duplicated genes was tested separately, we could not find statistical differences in BnaA03g29300D expression between resistant and non-resistant inoculated lines and no evidence of induction after inoculation. However, both duplicated genes are expressed at a similar, relatively high level in the roots of "Tosca"-background plants, showing a cumulative 2-times higher transcript accumulation before infection compared to the "BRH-1"-derived, single copy line.

In summary, the search for the genetic background of resistance to *P. brassicae* infection in *B. napus* cv. Tosca revealed a complex picture of genomic and transcriptomic changes. Based on genetic mapping, structural genomics, expression analyses, and functional annotation, we conclude that the TNL gene (BnaA03g29300D) duplication is most likely to be involved in the resistance. Certainly, further experimental tests, including a gene knockout and functional complementation must be conducted to confirm the role of this gene, and/or its duplication, in the resistance against *P. brassicae*.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI SRA, BioProject number: PRJNA685314.

AUTHOR CONTRIBUTIONS

KM, CO, and WK conceived the research. EJ and AP performed and MK supervised the phenotyping experiments. KM and JN performed and IB-B supervised the SSR and SCAR genotyping experiments and production of the DH line seeds. PK collected, processed, and analyzed the phenotyping, genotyping, and sequencing data and performed association analyses. HSC generated a part of Oxford Nanopore sequencing data. WK supervised and coordinated the whole project. PK and WK wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by NCN 2016/22/M/NZ9/00604 and POWR.03.02.00-00-I022/16.

ACKNOWLEDGMENTS

The authors would like to thank Andreas Stahl for help with heritability estimation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.639631/full#supplementary-material

REFERENCES

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bergelson, J., Kreitman, M., Stahl, E. A., and Tian, D. (2001). Evolutionary dynamics of plant R-genes. *Science* 292, 2281–2285. doi: 10.1126/science.1061337

Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., et al. (2010). TRICHOME BIREFRINGENCE and its homolog AT5G01360 encode plant-specific DUF231 proteins required for cellulose biosynthesis in *Arabidopsis*. *Plant Physiol.* 153, 590–602. doi: 10.1104/pp.110.153320

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

Broman, K. W., Gatti, D. M., Simecek, P., Furlotte, N. A., Prins, P., Sen, Ś, et al. (2019). R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics* 211, 495–502. doi: 10.1534/genetics.118.301595

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Büttner, M. (2010). The *Arabidopsis* sugar transporter (AtSTP) family: an update. *Plant Biol.* 12(Suppl. 1), 35–41. doi: 10.1111/j.1438-8677.2010.00383.x

Cao, T., Manolii, V. P., Strelkov, S. E., Hwang, S.-F., and Howard, R. J. (2009). Virulence and spread of *Plasmodiophora brassicae* [clubroot] in Alberta, Canada. *Can. J. Plant Pathol.* 31, 321–329. doi: 10.1080/07060660909507606

Cegielska-Taras, T., Tykarska, T., Szała, L., Kuraś, L., and Krzymański, J. (2002). Direct plant development from microspore-derived embryos of winter oilseed rape *Brassica napus* L. ssp. Oleifera (DC.) Metzger. *Euphytica* 124, 341–347.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953.

Chawla, H. S., Lee, H., Gabur, I., Vollrath, P., Tamilselvan-Nattar-Amutha, S., Obermeier, C., et al. (2020). Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. *Plant Biotechnol. J.* 19, 240–250. doi: 10.1111/pbi.13456

Chen, J., Jing, J., Zhan, Z., Zhang, T., Zhang, C., and Piao, Z. (2013). Identification of novel QTLs for isolate-specific partial resistance to *Plasmodiophora brassicae* in *Brassica rapa*. *PLoS One* 8:e85307. doi: 10.1371/journal.pone.0085307

Choi, S. R., Oh, S. H., Chhapekar, S. S., Dhandapani, V., Lee, C. Y., Rameneni, J. J., et al. (2020). Quantitative trait locus mapping of clubroot resistance and *Plasmodiophora brassicae* pathotype banglim-specific marker development in *Brassica rapa*. *Int. J. Mol. Sci.* 21:4157. doi: 10.3390/ijms21114157

Chu, M., Song, T., Falk, K. C., Zhang, X., Liu, X., Chang, A., et al. (2014). Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. *BMC Genomics* 15:1166. doi: 10.1186/1471-2164-15-1166

Chu, M., Yu, F., Falk, K. C., Liu, X., Zhang, X., Chang, A., et al. (2013). Identification of the Clubroot Resistance Gene Rpb1 and introgression of the resistance into canola breeding lines using a marker-assisted approach. *Acta Hortic.* 2013, 599–605. doi: 10.17660/actahortic.2013.1005.74

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.4161/fly.19695

Clarke, W. E., Higgins, E. E., 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, 1887–1899. doi: 10.1007/s00122-016-2746-7

Dakouri, A., Zhang, X., Peng, G., Falk, K. C., Gossen, B. D., Strelkov, S. E., et al. (2018). Analysis of genome-wide variants through bulked segregant RNA sequencing reveals a major gene for resistance to *Plasmodiophora brassicae* in *Brassica oleracea*. *Sci. Rep.* 8:17657. doi: 10.1038/s41598-018-36187-5

de Araújo, A. C., Fonseca, F. C. A., Cotta, M. G., Alves, G. S. C., and Miller, R. N. G. (2019). Plant NLR receptor proteins and their potential in the development of durable genetic resistance to biotic stresses. *Biotechnol. Res. Innov.* 3, 80–94.

DeYoung, B. J., and Innes, R. W. (2006). Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat. Immunol.* 7, 1243–1249. doi: 10.1038/ni1410

Diederichsen, E., Beckmann, J., Schondelmeier, J., and Dreyer, F. (2006). Genetics of clubroot resistance in *Brassica napus* 'Mendel'. *Acta Hortic.* 706, 307–312. doi: 10.17660/actahortic.2006.706.35

Diederichsen, E., Frauen, M., Linders, E. G. A., Hatakeyama, K., and Hirai, M. (2009). Status and perspectives of clubroot resistance breeding in crucifer crops. *J. Plant Growth Regul.* 28, 265–281. doi: 10.1007/s00344-009-9100-0

Diederichsen, E., Frauen, M., and Ludwig-Müller, J. (2014). Clubroot disease management challenges from a German perspective. *Can. J. Plant Pathol.* 36, 85–98. doi: 10.1080/07060661.2013.861871

Diederichsen, E., and Sacristan, M. D. (1996). Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to Plasmodiophora brassicae Wor. *Plant Breed.* 115, 5–10. doi: 10.1111/j.1439-0523.1996.tb00862.x

Dixon, G. R. (2007). *Vegetable Brassicas and Related Crucifers*. Wallingford: CABI.

Dixon, G. R. (2009). The occurrence and economic impact of *Plasmodiophora brassicae* and clubroot disease. *J. Plant Growth Regul.* 28, 194–202. doi: 10.1007/s00344-009-9090-y

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Dolatabadian, A., Bayer, P. E., Tirnaz, S., Hurgobin, B., Edwards, D., and Batley, J. (2020). Characterization of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. *Plant Biotechnol. J.* 18, 969–982. doi: 10.1111/pbi.13262

Doyle, J. J., and Doyle, J. L. (1990). Isolation of plant DNA from fesh tissue. *Focus* 12:1053.

Dubey, N., and Singh, K. (2018). "Role of NBS-LRR proteins in plant defense," in *Molecular Aspects of Plant-Pathogen Interaction*, eds A. Singh and I. Singh (Singapore: Springer), 115–138. doi: 10.1007/978-981-10-7371-7_5

Dunwell, J. M., George Gibbings, J., Mahmood, T., and Saqlan Naqvi, S. M. (2008). Germin and germin-like proteins: evolution, structure, and function. *Crit. Rev. Plant Sci.* 27, 342–375. doi: 10.1080/07352680802333938

Edge, P., and Bansal, V. (2019). Longshot enables accurate variant calling in diploid genomes from single-molecule long read sequencing. *Nat. Commun.* 10:4660.

Farid, M., Yang, R.-C., Kebede, B., and Rahman, H. (2020). Evaluation of *Brassica oleracea* accessions for resistance to *Plasmodiophora brassicae* and identification of genomic regions associated with resistance. *Genome* 63, 91–101. doi: 10.1139/gen-2019-0098

Figdore, S. S., Ferreira, M. E., Slocum, M. K., and Williams, P. H. (1993). Association of RFLP markers with trait loci affecting clubroot resistance and morphological characters in *Brassica oleracea* L. *Euphytica* 69, 33–44. doi: 10.1007/bf00021723

Fredua-Agyeman, R., and Rahman, H. (2016). Mapping of the clubroot disease resistance in spring *Brassica napus* canola introgressed from European winter canola cv. "Mendel.". *Euphytica* 211, 201–213. doi: 10.1007/s10681-016-1730-2

Fredua-Agyeman, R., Yu, Z., Hwang, S.-F., and Strelkov, S. E. (2020). Genome-wide mapping of loci associated with resistance to clubroot in *Brassica napus* ssp. napobrassica (Rutabaga) Accessions From Nordic Countries. *Front. Plant Sci.* 11:742. doi: 10.3389/fpls.2020.00742

Gabur, I., Chawla, H. S., Liu, X., Kumar, V., Faure, S., von Tiedemann, A., et al. (2018). Finding invisible quantitative trait loci with missing data. *Plant Biotechnol. J.* 16, 2102–2112. doi: 10.1111/pbi.12942

Gao, Y., He, C., Zhang, D., Liu, X., Xu, Z., Tian, Y., et al. (2017). Two trichome birefringence-like proteins mediate xylan acetylation, which is essential for leaf blight resistance in rice. *Plant Physiol.* 173, 470–481. doi: 10.1104/pp.16.01618

Gao, Y., Wang, W., Zhang, T., Gong, Z., Zhao, H., and Han, G.-Z. (2018). Out of water: the origin and early diversification of plant R-genes. *Plant Physiol.* 177, 82–89. doi: 10.1104/pp.18.00185

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Grandclément, C., and Thomas, G. (1996). Detection and analysis of QTLs based on RAPD markers for polygenic resistance to *Plasmodiophora brassicae* woron in *Brassica oleracea* L. *Theor. Appl. Genet.* 93, 86–90. doi: 10.1007/s001220050251

Han, G.-Z. (2019). Origin and evolution of the plant immune system. *New Phytol.* 222, 70–83. doi: 10.1111/nph.15596

Happstadius, I., Ljungberg, A., Kristiansson, B., and Dixelius, C. (2003). Identification of *Brassica oleracea* germplasm with improved resistance to *Verticillium wilt*. *Plant Breed.* 122, 30–34. doi: 10.1046/j.1439-0523.2003.00774.x

Hatakeyama, K., Niwa, T., Kato, T., Ohara, T., Kakizaki, T., and Matsumoto, S. (2017). The tandem repeated organization of NB-LRR genes in the clubroot-resistant CRb locus in *Brassica rapa* L. *Mol. Genet. Genomics* 292, 397–405. doi: 10.1007/s00438-016-1281-1

Hatakeyama, K., Suwabe, K., Tomita, R. N., Kato, T., Nunome, T., Fukuoka, H., et al. (2013). Identification and characterization of Crr1a, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *PLoS One* 8:e54745. doi: 10.1371/journal.pone.0054745

Hejna, O., Havlickova, L., He, Z., Bancroft, I., and Curn, V. (2019). Analysing the genetic architecture of clubroot resistance variation in *Brassica napus* by associative transcriptomics. *Mol. Breed.* 39:112.

Hirai, M., Harada, T., Kubo, N., Tsukada, M., Suwabe, K., and Matsumoto, S. (2004). A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. *Theor. Appl. Genet.* 108, 639–643. doi: 10.1007/s00122-003-1475-x

Hirani, A. H., Gao, F., Liu, J., Fu, G., Wu, C., McVetty, P. B. E., et al. (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. doi: 10.3389/fpls.2018.01628

Huang, Z., Peng, G., Liu, X., Deora, A., Falk, K. C., Gossen, B. D., et al. (2017). Fine mapping of a clubroot resistance gene in chinese cabbage using SNP markers identified from bulked segregant RNA sequencing. *Front. Plant Sci.* 8:1448. doi: 10.3389/fpls.2017.01448

Hwang, S.-F., Howard, R. J., Strelkov, S. E., Gossen, B. D., and Peng, G. (2014). Management of clubroot (*Plasmodiophora brassicae*) on canola (*Brassica napus*) in western Canada. *Can. J. Plant Pathol.* 36, 49–65. doi: 10.1080/07060661.2013.863806

Ilyas, M., Rasheed, A., and Mahmood, T. (2016). Functional characterization of germin and germin-like protein genes in various plant species using transgenic approaches. *Biotechnol. Lett.* 38, 1405–1421. doi: 10.1007/s10529-016-2129-9

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031

Kageyama, K., and Asano, T. (2009). Life Cycle of *Plasmodiophora brassicae*. *J. Plant Growth Regul.* 28, 203–211. doi: 10.1007/s00344-009-9101-z

Karasov, T. L., Horton, M. W., and Bergelson, J. (2014). Genomic variability as a driver of plant–pathogen coevolution? *Curr. Opin. Plant Biol.* 18, 24–30. doi: 10.1016/j.pbi.2013.12.003

Karim, M. M., Dakouri, A., Zhang, Y., Chen, Q., Peng, G., Strelkov, S. E., et al. (2020). Two clubroot-resistance genes, Rcr3 and Rcr9wa, mapped in *Brassica rapa* using bulk segregant RNA sequencing. *Int. J. Mol. Sci.* 21:5033. doi: 10.3390/ijms21145033

Kato, T., Hatakeyama, K., Fukino, N., and 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, 282–287. doi: 10.1270/jsbbs.62.282

Korbas, M., Jajor, E., and Budka, A. (2009). Clubroot (*Plasmodiophora brassicae*) - a threat for oilseed rape. *J. Plant Protect. Res.* 49, 446–451. doi: 10.2478/v10045-009-0071-78

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2016). *lmerTest: Tests in Linear Mixed Effects Models. R Package version 2.0–30.*

Laila, R., Park, J.-I., Robin, A. H. K., Natarajan, S., Vijayakumar, H., Shirasawa, K., et al. (2019). Mapping of a novel clubroot resistance QTL using ddRAD-seq in

Chinese cabbage (*Brassica rapa* L.). *BMC Plant Biol.* 19:13. doi: 10.1186/s12870-018-1615-8

Landry, B. S., Hubert, N., Crete, R., Chang, M. S., Lincoln, S. E., and 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. doi: 10.1139/g92-061

Laurens, F., and Thomas, G. (2004). Inheritance of resistance to clubroot (*Plasmodiophora brassicae* Wor.) *in Kale (Brassica oleracea* ssp. Acephala). *Hereditas* 119, 253–262. doi: 10.1111/j.1601-5223.1993.00253.x

Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., et al. (2016). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research* 5:ISCB Comm J-1408. doi: 10.12688/f1000research.9005.3

Lee, H., Chawla, H. S., Obermeier, C., Dreyer, F., Abbadi, A., and Snowdon, R. (2020). Chromosome-scale assembly of winter oilseed rape *Brassica napus*. *Front. Plant Sci.* 11:496. doi: 10.3389/fpls.2020.00496

Lee, J., Izzah, N. K., Choi, B.-S., Joh, H. J., Lee, S.-C., Perumal, S., et al. (2016). Genotyping-by-sequencing map permits identification of clubroot resistance QTLs and revision of the reference genome assembly in cabbage (*Brassica oleracea* L.). *DNA Res.* 23, 29–41.

Lenth, V. R. (2016). Least-squares means: the R package lsmeans. *J. Statist. Softw.* 69, 1–33.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). 1000 genome project data processing subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, L., Luo, Y., Chen, B., Xu, K., Zhang, F., Li, H., et al. (2016). A genome-wide association study reveals new loci for resistance to clubroot disease in *Brassica napus*. *Front. Plant Sci.* 7:1483. doi: 10.3389/fpls.2016.01483

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Liu, L., Qin, L., Zhou, Z., Wilhelmina, G. H., Liu, S., and Wei, Y. D. (2020). Refining the life cycle of *Plasmodiophora brassicae*. *Phytopathology* 110, 1704–1712. doi: 10.1094/phyto-02-20-0029-r

Lv, H., Fang, Z., Yang, L., Zhang, Y., and Wang, Y. (2020). An update on the arsenal: mining resistance genes for disease management of *Brassica* crops in the genomic era. *Hortic. Res.* 7:34. doi: 10.1038/s41438-020-0257-9

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641.

Malinowski, R., Truman, W., and Blicharz, S. (2019). Genius architect or clever thief—how *Plasmodiophora brassicae* reprograms host development to establish a pathogen-oriented physiological sink. *Mol. Plant Microbe Interact.* 32, 1259–1266. doi: 10.1094/mpmi-03-19-0069-cr

Manzanares-Dauleux, M. J., Delourme, R., Baron, F., and Thomas, G. (2000). Mapping of one major gene and of QTLs involved in resistance to clubroot in *Brassica napus*. *Theor. Appl. Genet.* 101, 885–891. doi: 10.1007/s001220051557

Matsumoto, E., Ueno, H., Aruga, D., Sakamoto, K., and Hayashida, N. (2012). Accumulation of three clubroot resistance genes through marker-assisted selection in chinese cabbage (*Brassica rapa* ssp. pekinensis). *J. Jpn. Soc. Hortic. Sci.* 81, 184–190. doi: 10.2503/jjshs1.81.184

Matsumoto, E., Yasui, C., Ohi, M., and Tsukada, M. (1998). Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Euphytica* 104:79. doi: 10.1023/A:1018370418201

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

Mondragón-Palomino, M., Meyers, B. C., Michelmore, R. W., and Gaut, B. S. (2002). Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res.* 12, 1305–1315. doi: 10.1101/gr.159402

Monroe, J. G., Grey Monroe, J., Allen, Z. A., Tanger, P., Mullen, J. L., Lovell, J. T., et al. (2017). TSPmap, a tool making use of traveling salesperson problem solvers in the efficient and accurate construction of high-density genetic linkage maps. *BioData Mining* 10:38. doi: 10.1186/s13040-017-0158-0

Moriguchi, K., Kimizuka-Takagi, C., Ishii, K., and Nomura, K. (1999). A genetic map based on RAPD, RFLP, isozyme, morphological markers and QTL analysis for clubroot resistance in *Brassica oleracea*. *Breed. Sci.* 49, 257–265. doi: 10.1270/jsbbs.49.257

Nagaoka, T., Doullah, M. A. U., Matsumoto, S., Kawasaki, S., Ishikawa, T., Hori, H., et al. (2010). Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. *Theor. Appl. Genet.* 120, 1335–1346. doi: 10.1007/s00122-010-1259-z

Neik, T. X., Barbetti, M. J., and Batley, J. (2017). Current status and challenges in identifying disease resistance genes in *Brassica napus*. *Front. Plant Sci.* 8:1788. doi: 10.3389/fpls.2017.01788

Nguyen, M. L., Monakhos, G. F., Komakhin, R. A., and Monakhos, S. G. (2018). The new clubroot resistance locus is located on chromosome A05 in Chinese Cabbage (*Brassica rapa* L.). *Russ. J. Genet.* 54, 296–304. doi: 10.1134/s1022795418030080

Nomura, K., Minegishi, Y., Kimizuka-Takagi, C., Fujioka, T., Moriguchi, K., Shishido, R., et al. (2005). Evaluation of F2 and F3 plants introgressed with QTLs for clubroot resistance in cabbage developed by using SCAR markers. *Plant Breed.* 124, 371–375. doi: 10.1111/j.1439-0523.2005.01105.x

Padmanabhan, M., Cournoyer, P., and Dinesh-Kumar, S. P. (2009). The leucine-rich repeat domain in plant innate immunity: a wealth of possibilities. *Cell. Microbiol.* 11, 191–198. doi: 10.1111/j.1462-5822.2008.01260.x

Pang, W., Fu, P., Li, X., Zhan, Z., Yu, S., and Piao, Z. (2018). Identification and mapping of the clubroot resistance gene CRd in chinese cabbage (*Brassica rapa* ssp. pekinensis). *Front. Plant Sci.* 9:653. doi: 10.3389/fpls.2018.00653

Pang, W., Liang, S., Li, X., Li, P., Yu, S., Lim, Y. P., et al. (2014). Genetic detection of clubroot resistance loci in a new population of *Brassica rapa*. *Hortic. Environ. Biotechnol.* 55, 540–547. doi: 10.1007/s13580-014-0079-5

Pauly, M., and Ramírez, V. (2018). New insights into wall polysaccharide O-Acetylation. *Front. Plant Sci.* 9:1210. doi: 10.3389/fpls.2018.01210

Peng, L., Zhou, L., Li, Q., Wei, D., Ren, X., Song, H., et al. (2018). Identification of quantitative trait loci for clubroot resistance in *Brassica oleracea* with the use of *Brassica* SNP microarray. *Front. Plant Sci.* 9:822. doi: 10.3389/fpls.2018.00822

Piao, Z. Y., Deng, Y. Q., Choi, S. R., Park, Y. J., and Lim, Y. P. (2004). SCAR and CAPS mapping of CRb, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *Theor. Appl. Genet.* 108, 1458–1465. doi: 10.1007/s00122-003-1577-5

Piepho, H. P., and Möhring, J. (2007). Computing heritability and selection response from unbalanced plant breeding trials. *Genetics* 177, 1881–1888. doi: 10.1534/genetics.107.074229

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47, W191–W198.

Rietz, S., Bernsdorff, F. E. M., and Cai, D. (2012). Members of the germin-like protein family in *Brassica napus* are candidates for the initiation of an oxidative burst that impedes pathogenesis of *Sclerotinia sclerotiorum*. *J. Exp. Bot.* 63, 5507–5519. doi: 10.1093/jxb/ers203

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rocherieux, J., Glory, P., Giboulot, A., Boury, S., Barbeyron, G., Thomas, G., et al. (2004). Isolate-specific and broad-spectrum QTLs are involved in the control of clubroot in *Brassica oleracea*. *Theor. Appl. Genet.* 108, 1555–1563. doi: 10.1007/s00122-003-1580-x

Rousseau-Gueutin, M., Belser, C., Da Silva, C., Richard, G., Istace, B., Cruaud, C., et al. (2020). Long-read assembly of the *Brassica napus* reference genome Darmor-bzh. *GigaScience* 9:giaa137. doi: 10.1093/gigascience/giaa137

Ruan, J., and Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. *Nat. Methods* 17, 155–158. doi: 10.1038/s41592-019-0669-3

Saito, M., Kubo, N., Matsumoto, S., Suwabe, K., Tsukada, M., and Hirai, M. (2006). Fine mapping of the clubroot resistance gene, Crr3, in *Brassica rapa*. *Theor. Appl. Genet.* 114, 81–91. doi: 10.1007/s00122-006-0412-1

Sakamoto, K., Saito, A., Hayashida, N., Taguchi, G., and 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. doi: 10.1007/s00122-008-0817-0

Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A., et al. (2017). Accurate detection of complex structural variations using single molecule sequencing. *Biorxiv* [Preprint]. doi: 10.1101/169557

Some, A., Manzanares, M. J., Laurens, F., Baron, F., Thomas, G., and 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. doi: 10.1046/j.1365-3059.1996.d01-155.x

Song, J.-M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., et al. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nat. Plants* 6, 34–45. doi: 10.1038/s41477-019-0577-7

Song, T., Chu, M., Lahlali, R., Yu, F., and Peng, G. (2016). Shotgun label-free proteomic analysis of clubroot (*Plasmodiophora brassicae*) resistance conferred by the gene Rcr1 in *Brassica rapa*. *Front. Plant Sci.* 7:1013. doi: 10.3389/fpls.2016.01013

Stahl, A., Vollrath, P., Samans, B., Frisch, M., Wittkop, B., and Snowdon, R. J. (2019). Effect of breeding on nitrogen use efficiency-associated traits in oilseed rape. *J. Exp. Bot.* 70, 1969–1986. doi: 10.1093/jxb/erz044

Stranne, M., Ren, Y., Fimognari, L., Birdseye, D., Yan, J., Bardor, M., et al. (2018). TBL10 is required for O -acetylation of pectic rhamnogalacturonan-I in *Arabidopsis thaliana*. *Plant J.* 96, 772–785. doi: 10.1111/tpj.14067

Strehlow, B., de Mol, F., and Struck, C. (2015). Risk potential of clubroot disease on winter oilseed rape. *Plant Dis.* 99, 667–675. doi: 10.1094/pdis-05-14-0482-re

Strelkov, S. E., Hwang, S.-F., Manolii, V. P., Cao, T., and Feindel, D. (2016). Emergence of new virulence phenotypes of *Plasmodiophora brassicae* on canola (*Brassica napus*) in Alberta, Canada. *Eur. J. Plant Pathol.* 145, 517–529. doi: 10.1007/s10658-016-0888-8

Strelkov, S. E., Hwang, S.-F., Manolii, V. P., Cao, T., Fredua-Agyeman, R., Harding, M. W., et al. (2018). Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. *Can. J. Plant Pathol.* 40, 284–298. doi: 10.1080/07060661.2018.1459851

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., et al. (2003). Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *Theor. Appl. Genet.* 107, 997–1002. doi: 10.1007/s00122-003-1309-x

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., et al. (2006). Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173, 309–319. doi: 10.1534/genetics.104.038968

Szała, L., Sosnowska, K., and Cegielska-Taras, T. (2020). Induced chromosome doubling in microspores and regenerated haploid plants of *Brassica napus*. *Acta Biol. Cracoviensia Ser. Bot.* 62, 23–31.

Törönen, P., Medlar, A., and Holm, L. (2018). PANNZER2: a rapid functional annotation web server. *Nucleic Acids Res.* 46, W84–W88. doi: 10.1093/nar/gky350

Ueno, H., Matsumoto, E., Aruga, D., Kitagawa, S., Matsumura, H., and Hayashida, N. (2012). Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. *Plant Mol. Biol.* 80, 621–629. doi: 10.1007/s11103-012-9971-5

van Wersch, S., and Li, X. (2019). Stronger when together: clustering of Plant NLR Disease resistance Genes. *Trends Plant Sci.* 24, 688–699. doi: 10.1016/j.tplants.2019.05.005

Vigier, B., Chiang, M. S., and Hume, D. J. (1989). Source of resistance to clubroot (*Plasmodiophora brassicae* Wor.) in triazine-resistant spring canola (rapeseed). *Can. Plant Dis. Surv.* 69, 113–115.

Voorrips, R. E., Jongerius, M. C., and Kanne, H. J. (1997). Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. *Theor. Appl. Genet.* 94, 75–82. doi: 10.1007/s001220050384

Voorrips, R. E., and Visser, D. L. (1993). Examination of resistance to clubroot in accessions of *Brassica oleracea* using a glasshouse seedling test. *Netherlands J. Plant Pathol.* 99, 269–276. doi: 10.1007/bf01974308

Wahl, V., Brand, L. H., Guo, Y.-L., and Schmid, M. (2010). The FANTASTIC FOUR proteins influence shoot meristem size in *Arabidopsis thaliana*. *BMC Plant Biol.* 10:285. doi: 10.1186/1471-2229-10-285

Walerowski, P., Gündel, A., Yahaya, N., Truman, W., Sobczak, M., Olszak, M., et al. (2018). Clubroot disease stimulates early steps of phloem differentiation and recruits sweet sucrose transporters within developing galls. *Plant Cell* 30, 3058–3073. doi: 10.1105/tpc.18.00283

Wallenhammar, A. C. (1996). Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathol.* 45, 710–719. doi: 10.1046/j.1365-3059.1996.d01-173.x

Werner, S., Diederichsen, E., Frauen, M., Schondelmaier, J., and Jung, C. (2008). Genetic mapping of clubroot resistance genes in oilseed rape. *Theor. Appl. Genet.* 116, 363–372. doi: 10.1007/s00122-007-0674-2

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Berlin: Springer Science & Business Media.

Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., et al. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* 344, 299–303. doi: 10.1126/science.1247357

Yang, S., Li, J., Zhang, X., Zhang, Q., Huang, J., Chen, J.-Q., et al. (2013). Rapidly evolving R genes in diverse grass species confer resistance to rice blast disease. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18572–18577. doi: 10.1073/pnas.1318211110

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

Yuan, Y., Teng, Q., Zhong, R., Haghighat, M., Richardson, E. A., and Ye, Z.-H. (2016). Mutations of *Arabidopsis* TBL32 and TBL33 affect xylan acetylation and secondary wall deposition. *PLoS One* 11:e0146460. doi: 10.1371/journal.pone.0146460

Zhang, H., Feng, J., Hwang, S.-F., Strelkov, S. E., Falak, I., Huang, X., et al. (2016). Mapping of clubroot (*Plasmodiophora brassicae*) resistance in canola (*Brassica napus*). *Plant Pathol.* 65, 435–440. doi: 10.1111/ppa.12422

Zhang, T., Zhao, Z., Zhang, C., Pang, W., Choi, S. R., Lim, Y. P., et al. (2014). Fine genetic and physical mapping of the CRb gene conferring resistance to clubroot disease in *Brassica rapa*. *Mol. Breed.* 34, 1173–1183. doi: 10.1007/s11032-014-0108-1

Zhang, W., Wang, S., Yu, F., Tang, J., Yu, L., Wang, H., et al. (2019). Genome-wide identification and expression profiling of sugar transporter protein (STP) family genes in cabbage (*Brassica oleracea* var. capitata L.) reveals their involvement in clubroot disease responses. *Genes* 10:71. doi: 10.3390/genes10010071

Zheng, X., Koopmann, B., Ulber, B., and von Tiedemann, A. (2020). A global survey on diseases and pests in oilseed rape—current challenges and innovative strategies of control. *Front. Agron.* 2:908. doi: 10.3389/fagro.2020.590908

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Kopec, Mikolajczyk, Jajor, Perek, Nowakowska, Obermeier, Chawla, Korbas, Bartkowiak-Broda and Karlowski. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

3.2. Supplementary figures

Figure S1. IGV screenshot of clubroot resistance locus on Express 617 assembly. Genomic ONT reads of Tosca and BRH-1 cultivars mapped to the reference and gene annotation track visible. TNL gene marked with an arrow.

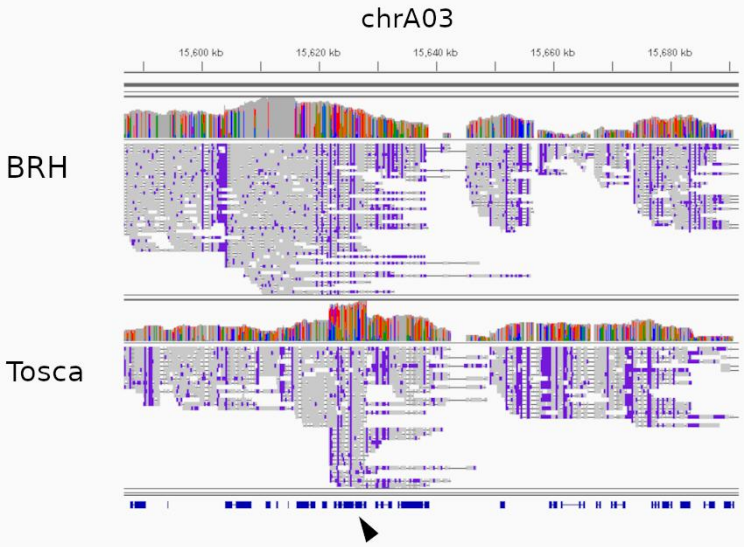


Figure S2. Fragment of assembled Tosca region with TNL gene copies marked with red.
In tracks Illumina RNA-seq reads from Tosca background DH line and ONT reads from Tosca.

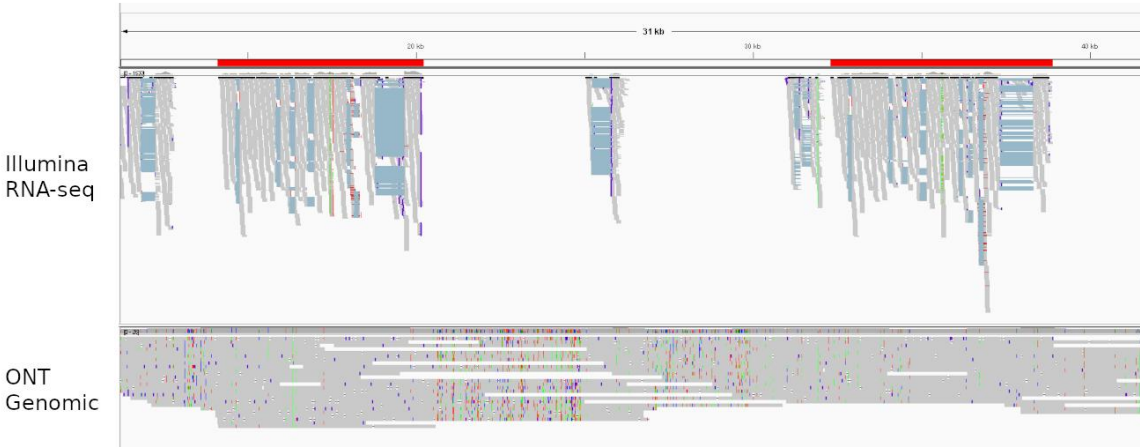
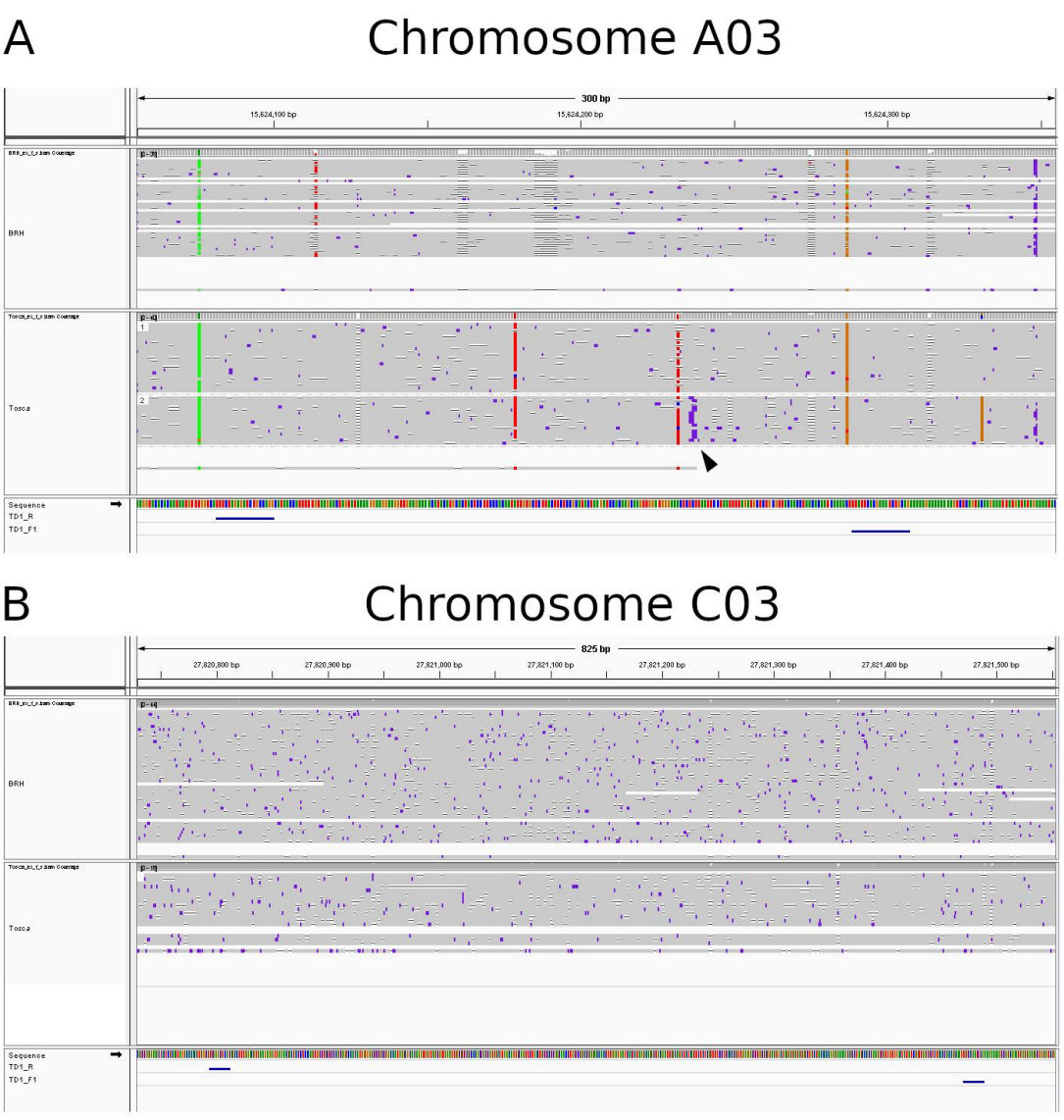


Figure S3. (A) Region covered by TD1_R TD1_F1 primers on chromosome A03 of Express 617 assembly. ONT reads show differences between the BRH-1 and Tosca lines. In case of Tosca two clusters of reads, corresponding to each gene copy are visible, with ~350bp insertion marked with an arrow.
(B) Homoeologous C03 region - no visible differences between Tosca and BRH-1.



3.3. Author's contribution statement

Author's Statement

I declare that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, is a part of my doctoral thesis.

I was involved in the phenotyping and genotyping of the mapping population. I have processed the RNA-seq samples, and generated part of the Nanopore sequencing data. I have collected, processed, analyzed, and integrated the phenotyping, genotyping, and sequencing data. These analyses included genetic map construction, QTL mapping, local gene content and sequence polymorphism evaluation, gene expression study, and genetic marker design. I have prepared all the figures and co-wrote the article.

Poznań, 25.06.2023

Piotr Kopeć

Supervisor: Prof. dr hab. Wojciech Karłowski

3.4. Co-author's contribution statements

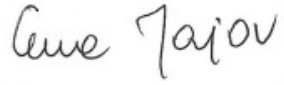
Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Katarzyna Mikolajczyk
Affiliation	Plant Breeding and Acclimatization Institute – National Research Institute, Poznań Division, Department of Oilseed Crops
Contribution	Conceiving the research, performing the SSR and SCAR genotyping experiments and production of the DH line seeds.
Date	13.06.2023
Signature	

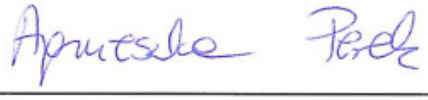
Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Ewa Jajor
Affiliation	**INSTITUTE OF PLANT PROTECTION – NATIONAL RESEARCH INSTITUTE**
Contribution	Continuation of research, performed the phenotyping experiments.
Date	20.06.2023
Signature	

Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Agnieszka Perek
Affiliation	Institute of Plant Protection – National Research Institute
Contribution	Performing the phenotyping experiments
Date	14.06.2023
Signature	

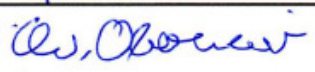
Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Joanna Nowakowska
Affiliation	Plant Breeding and Acclimatization Institute – National Research Institute, Poznań Division, Department of Oilseed Crops
Contribution	Performing the SSR and SCAR genotyping experiments and production of the DH line seeds.
Date	14.06.2023
Signature	Joanna Nowakowska

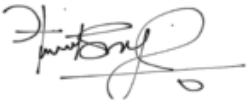
Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Christian Obermeier
Affiliation	Justus Liebig University, Department of Plant Breeding, Heinrich-Buff-Ring 26, Giessen, Germany
Contribution	Conceived the research
Date	19.6.2023
Signature	

Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Harmeet Singh Chawla
Affiliation	University of Manitoba
Contribution	Generated a part of Oxford Nanopore sequencing data
Date	14-06-2023
Signature	

Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Marek Korbas
Affiliation	**INSTITUTE OF PLANT PROTECTION – NATIONAL RESEARCH INSTITUTE**
Contribution	Continuation of research, supervised the phenotyping experiments.
Date	20.06.2023
Signature	

Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Iwona Bartkowiak-Broda
Affiliation	Plant Breeding and Acclimatization Institute – National Research Institute, Poznań Division, Department of Oilseed Crops
Contribution	Supervising the SSR and SCAR genotyping experiments and production of the DH line seeds.
Date	14.06.2023
Signature	

Authorship Contribution Statement

I declare that I am aware that the article **"Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca"** Kopec PM, Mikolajczyk K, Jajor E, Perek A, Nowakowska J, Obermeier C, Chawla HS, Korbas M, Bartkowiak-Broda I, Karlowski WM. Front Plant Sci. 2021 Apr 8;12:639631. doi: 10.3389/fpls.2021.639631, has been included in the doctoral thesis of Piotr Kopeć.

Co-Author's Name	Wojciech Karłowski
Affiliation	Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University Poznań
Contribution	Conceived the research, supervised and coordinated the whole project, and co-wrote the manuscript.
Date	30.06.2023
Signature	

3.5. Comments to 'Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in Brassica napus cv. Tosca'

At the outset of this work, the genetic bases of resistance of *B. napus* 'Tosca' were not publicly known. However, during the course of this study, the results of parallel work by the Monsanto company were published as patent US10415101B2. In the patent documentation, the resistance locus was mapped to a 4 cM region, which includes the 0.4 cM interval determined in our study. Following the submission of our manuscript, another patent, by BASF, was published under US20210017529A1. This document addressed functional aspects of the genetic basis of 'Tosca'-derived resistance. The implications of this publication in regard to this work are discussed in the next chapter.

Shortly after the publication of our manuscript, the results of a parallel Canadian study were made publicly available, wherein the resistance locus was mapped to the 5 cM region, consistent with our study (Fredua-Agyeman et al., 2021). Interestingly, due to the susceptibility to the most common Canadian *P. brassicae* pathotypes, 'Tosca' was deemed unsuitable as the sole source of resistance in local breeding programs. This underscores the necessity of considering the pathogen's population structure in resistance breeding efforts.

Importantly, all of the studies agree on the genetic locus conferring the resistance, and none of the published studies, apart from this work, acknowledged the existence of the TNL-family gene duplication.

Furthermore, although not directly featuring 'Tosca', but relevant to this work, the 1 Mb resistance locus homologous to *Crr3*, has been identified in *B. napus*, with resistance introgressed from the rutabaga variety 'Polycross'. The authors based their candidate gene assessment solely on the results of differential expression analysis, suggesting that the resistance might be a consequence of *STP6* downregulation in resistant plants - an observation consistent with our study. The TNL-family gene was stably expressed, with no significant differences between mock and inoculated plants at both 7 and 14 DAI, in susceptible and resistant plants (Z. Wang et al., 2022).

4. Functional study of the duplicated TNL gene

4.1. Introduction

The identification of a tandemly duplicated TNL gene within a narrow locus associated with clubroot resistance inspired further investigation into the functional role of these genes. At this point, it was essential to determine whether the effect could be attributed to any single paralog, both paralogs acting independently, or a synergistic interaction between them.

For this purpose, transgenic *Arabidopsis thaliana* lines were developed, each expressing a single copy or both copies simultaneously. These transformants were subsequently inoculated with *P. brassicae* spores, and the severity of their responses was measured. The *A. thaliana* Col-0 ecotype was selected for this study due to its susceptibility to the *P. brassicae* isolate in question, as established in preliminary tests.

The close evolutionary relatedness between *A. thaliana* and *Brassica* species facilitates the transferability of functional analyses. Moreover, this approach has proven effective in previous studies, notably in the case of two cloned and characterized TNL-family clubroot resistance genes, *Crr1a* and *CRa*/*CRb,* derived from *B. rapa* (Hatakeyama et al., 2013, 2017). In these studies, gain-of-function Col-0 mutants exhibited a resistance with similar pathotypic specificity to the donor cultivars. Therefore, the benefits of working with *A. thaliana*, such as short generation time, abundant seed production, a lower degree of genomic multiplication, and ease of transformation, can be leveraged without compromising the validity of the results.

For my experiment, a synthetic CDS was used, circumventing the need to clone the gene. Several factors influenced the choice of this strategy: cloning of genomic fragments would require preserving the gene structure, and generating a single-insertion double transformant could be constrained by the size of the construct. Although cDNA cloning could resolve those issues, it relies on error-prone reverse transcriptase. Furthermore, the highly

similar 5' ends of the paralogs and the ambiguous 3' structure of the genes, due to the adjacent *STP6*-like fragment, were anticipated to pose additional challenges.

The synthetic CDS approach also facilitated the inclusion of a protein marker tag. Monomeric fluorescent proteins (FPs): mNeonGreen and mScarlet were selected for this purpose (Bindels et al., 2017; Shaner et al., 2013). To avoid potential disruption of protein function - a common issue with traditional fusion proteins - the TNLs were separated from the tags using a P2A element from the 2A family. This short sequence, derived from a porcine teschovirus, prevents the formation of a single internal peptide while allowing for downstream translation (Luke et al., 2015). Consequently, a fluorescence signal in the C-tagged translation unit serves as an indicator of a successful production of the upstream native protein.

To validate functional fluorescence, two additional transformant pools, expressing only the FPs, were created. Since FPs constitute another xenogenic element with a potential, albeit unlikely, effect on the infection process, they were included in the subsequent phenotypic experiment.

While the functional analysis was being conducted, a parallel, independent study from BASF, that assessed the genetic basis of 'Tosca'-derived resistance, was published under the patent US20210017529A1. In light of these new findings, a comparative analysis and synthesis of the relevant results was undertaken.

The results described in this and subsequent chapters have not yet been published in a scientific literature.

4.2. Materials and Methods

4.2.1. Nomenclature

The duplicated TNL gene is hereafter referred to as *crT*. The 'Tosca'-derived copies are numbered in accordance with (Kopec et al., 2021), i.e., T1 becomes *crT1,* and T2 becomes *crT2*. Collectively the 'Tosca'-derived copies are referred to as *crTn*. To simplify the description, the corresponding constructs and transformant plants are referred to as crT1, crT1, crT1,2 (collectively crTn), and mS, mNG (collectively mFP) for FP controls, whenever possible, as described in section 4.2.3.

4.2.2. CDS synthesis

The coding sequences of *crT1* and *crT2* genes were used as described in the supplementary information of (Kopec et al., 2021). The synthesis of *crT1*-P2A-*mNeonGreen* and *crT2*-P2A-*mScarlet* fragments and subcloning into the pUC57-Simple vector were ordered from GenScript. The sequence was validated using Sanger sequencing by the service provider.

4.2.3. Construct assembly and validation

For this experiment, five constructs were created (Fig 1), with expression driven by *A. thaliana* ubiquitin 10 (AT4G05320) promoter (ubq) and terminated by *Agrobacterium tumefaciens* nopaline synthase terminator (nos):

crTn:

crT1 ubq::*crT1*-P2A-*mNeonGreen*::nos, for the expression of *crT1* gene with codon-optimized mNeonGreen protein tag;

crT2 ubq::*crT2*-P2A-*mScarlet*::nos, for the expression of *crT2* gene with codon-optimized mScarlet protein tag;

crT1,2 ubq::*crT1*-P2A-*mNeonGreen*::Nos-ubq::*crT2*-P2A-*mScarlet*::nos, for expression of both genes, with corresponding FP tags;

mFP:

mS ubq::*mScarlet*::nos, and

mNG ubq::*mNeonGreen*::nos as controls.

All the cassettes were cloned into a pFGC5941-like binary vector (hereafter referred to as pFGC). This vector carries kanamycin and glufosinate resistance genes, which enable the selection of transformed bacteria and plants, respectively. The assembly of constructs crT1 and crT2 was carried out using circular polymerase extension cloning (CPEC) (Quan & Tian, 2009), while the remaining constructs were assembled via homology-based in vivo assembly in *Escherichia coli* DH5α (García-Nafría et al., 2016). All fragments were amplified with Phusion™ Plus high-fidelity polymerase (Thermo Scientific™), with the specific primers listed in Table 1. The pFGC backbone and nos terminator were transferred from pFGC-I2Cas9 plasmid (https://www.addgene.org/173158/), while the ubiquitin promoter with the 5'UTR was amplified from *A. thaliana* Col-0 genomic DNA.

Following successful insert-flanking-PCR and restriction analysis, selected clones were further validated with Sanger sequencing at the local Faculty of Biology, Adam Mickiewicz University sequencing facility. The sequencing spaned plasmid-insert boundaries in all cases, the nos-ubq junction in crT1,2, and almost the entire inserts for constructs carrying individual TNLs.

Table 1. Primers used throughout the study

Name	**Sequence (5' →3')**	**Use**
EPL1B_R2	cacgccgaaataaacgacca	Insert size check, Sanger sequencing
M13F_47	cgccagggttttcccagtcacgac	
pFGC_F	ttactagatcgcatgggaattctgagatttttcaaatcagtg	pFGC backbone amplification
pFGC_U_R	actcgtcgacagcttggcactggccgtc	
UBQ_R	agagagccatctgttaatcagaaaaactcagat	ubq promoter amplification
UBQ_F	gtgccaagctgtcgacgagtcagtaataaacg	
NosT_R	attcccatgcgatctagtaacatagatgacac	nos terminator amplification
NosT_T1_F	ttataaataagatcgttcaaacatttgg	
NosT_T2_F	atacaagtaagatcgttcaaacatttgg	

T1_P2A_mNG_U_F	tgattaacagatggctctctcattagcttc	*crTn*_P2A_FP fragments amplification
T1_P2A_mNG_R	ttgaacgatcttatttataaagctcatccatccc	
T2_P2A_mS_R	ttgaacgatcttacttgtataactcgtccatg	
T2U_F	gtcatctatgttactagatcgtcgacgagtcagtaataaac	crT1,2 construct assembly
T2U_R	gtcatctatgttactagatcgcatgggaattctgagattt	
T1U_R	tttattactgactcgtcgacgatctagtaacatagatgacaccg	
T1U_F	gtcatctatgttactagatcgcatgggaattctgagattt	
mNG_circ_F	tgattaacagatggtcagcaaagggga	mNG construct assembly
mNG_circ_R	tgctgaccatctgttaatcagaaaaactcagat	
mS_circ_R	tactgaccatctgttaatcagaaaaactcagat	mS construct assembly
mS_circ_F	tgattaacagatggtcagcaagggtgag	
ubq_seq1	ctgggtttgatcgttagatatcatc	Sanger Sequencing
Tsc_ss1	gtgatttctatctagatctggtg	
Tsc_ss2	cgccaatgatgttttggataaac	
Tsc_ss3	ggctgcacctgaaggttt	
Tsc_ss4	cttcgactatttgcctcctaa	
Tsc_ss5	ggtatctggatctctcggat	
Tsc_ss6	gcaccgtactactggaac	
Tsc_ss7	cttaagcaggcaggtgatg	Sanger Sequencing, genotyping, RT-PCR
mNGss1_R	gggcaccaggatccaag	
mSss1_R	caagagaaagggagtggc	
AM3_F	gtcgctagcaagacacttaaga	genotyping, RT-PCR
AM3_R	gtgtgccgagaaatatacgaac	
Actin_F	ccggtattgtgctggattct	
Actin_R	aatttcccgctctgctgttg	

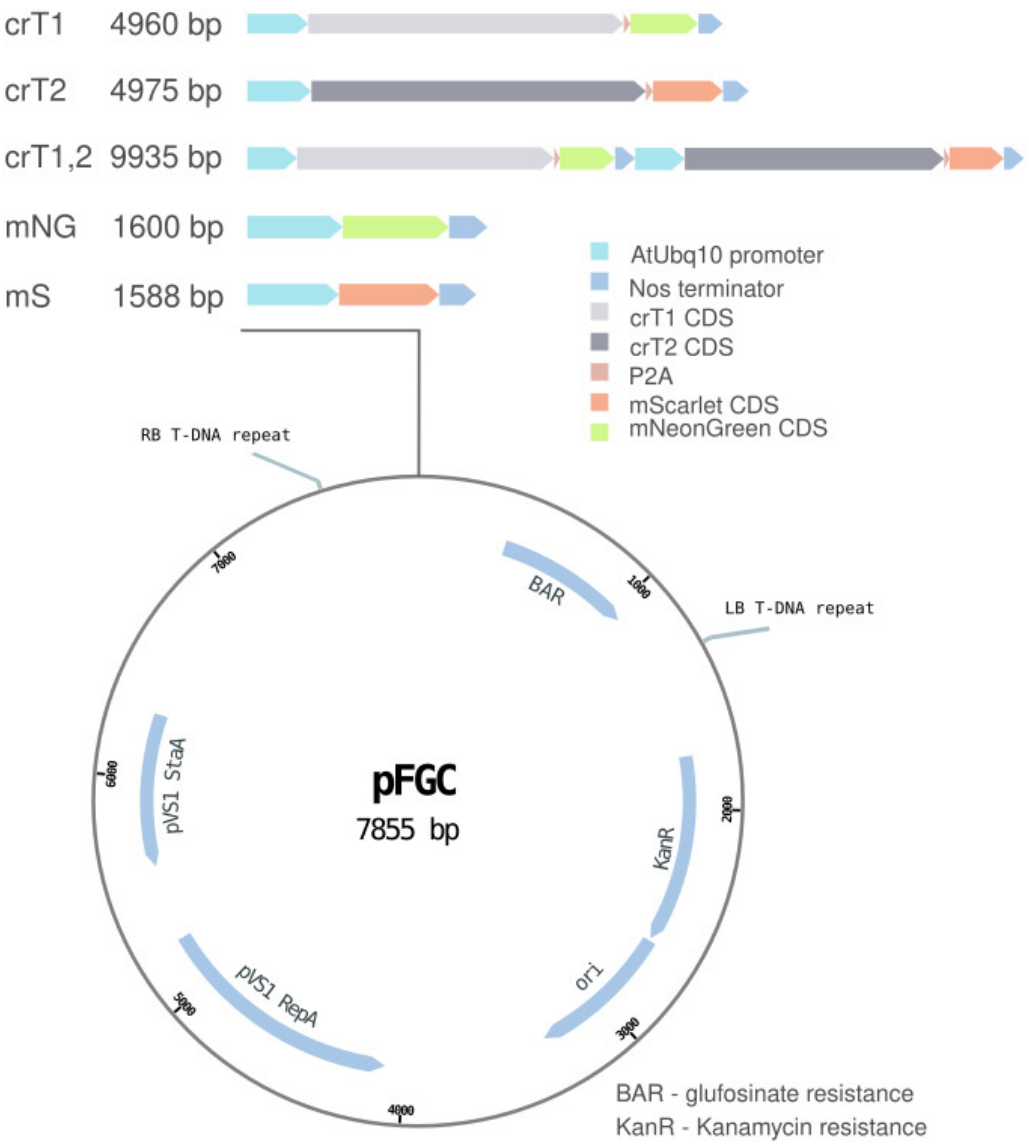


Figure 1. Vector design for *A. tumefaciens*-mediated *Arabidopsis* transformation

4.2.4. Agrobacterium-mediated *Arabidopsis* transformation

Agrobacterium tumefaciens GV3101 was transformed with plasmids via the freeze-thaw method (Weigel & Glazebrook, 2006), and transformants were then selected on LB agar plates containing 20 ug/ml rifampicin, 50 ug/ml kanamycin, and 40 ug/ml gentamicin. The presence and identity of the inserts were assessed through colony PCR with backbone-specific EPL1B_R2 and M13F_47 primers flanking the insertion site.

Subsequently, *A. thaliana* Col-0 plants were transformed using the floral-dip method (Clough & Bent, 1998). The plants were grown until stage 6.00 (Boyes et al., 2001). The primary bolt was trimmed to promote the growth of the auxiliary bolts. As the first flower opened, the buds were immersed in a suspension of *A. tumefaciens* in 5% sucrose and 0.04% Silwet Gold (UPL) and placed in a dark, humid box for 12 hours. After 6 and 12 days, newly emerged inflorescences were gently bathed with fresh dipping suspension using a pipette, following the protocol by (Narusaka et al., 2010), and the plants were subsequently subjected to another period of dark, humid treatment.

4.2.5. Transformant selection

The transformant T1 seedlings were initially selected based on glufosinate resistance. T0-derived seeds were sown onto coco-coir bedding. After germination, the seedlings were sprayed with a Basta® (BASF) solution containing a final glufosinate-ammonium concentration of 60 mg/l every 3-4 days. After two weeks, most of the plantlets had either died or become chlorotic, exhibiting severely retarded growth. Randomly selected healthy plantlets were transferred to soil for further propagation, and the presence of the transgenes was confirmed through PCR using Phire™ Plant Direct PCR protocol (Thermo Scientific™), with backbone-specific M13F_47 forward primer and reverse primer UBQ_R in the ubq promoter, multiplexed with actin-specific primers serving as control (Actin_F, Actin_R).

4.2.6. Expression validation

Transformants were assessed for protein expression based on fluorescence under Zeiss Axioskop 2 plus epifluorescence microscope illuminated with HBO100 100W Mercury short-arc lamp, with appropriate filters for mScarlet and mNeonGreen emission detection. mRNA expression was evaluated through RT-PCR. Total RNA was extracted from leaves and roots using Norgen Total RNA Purification Kit, treated with TURBO DNAse, and reverse-transcribed with Takara Superscript, using a mix of polyA and *crTn*-specific primer Tsc_ss6. A 35-cycle PCR with Phire™ polymerase was performed, using primers spanning the P2A-FP boundary, Tsc_ss7 forward and mNGss1_R or mSss1_R for *crT1* and *crT2,* respectively. Actin CDS primers (Actin_F, Actin_R) were used for positive RNA control, and primers spanning intergenic region on chromosome 4 (AM3_F, AM3_R) were used to control for genomic DNA contamination.

4.2.7. Phenotype assessment

Seeds from glufosinate-selected transformants, alongside the wild-type controls, were sown directly onto the soil and stratified for two days at 4°C. Following germination, the plant density was adjusted to approx 30-40 plantlets per pot. The highly clubroot-susceptible *B. rapa pekinensis* cv. 'Granaat' was included as an external control for positive infection. Two-week-old seedlings were inoculated and grown under conditions similar to those described in (Kopec et al., 2021). Eight weeks after inoculation, the plants were assessed for disease symptoms on a 0-3 scale, with 0 representing no symptoms, 1 representing small nodules on lateral roots, 2 representing larger galls on lateral roots, swollen main root and/or hypocotyl, and 3 representing severely galled roots and hypocotyl. The phenotyping was conducted in cooperation with dr Ewa Jajor from the Institute of Plant Protection - National Research Institute in Poznań.

4.3. Results

4.3.1. Validation of transformation

The initial selection of T1 seedlings based on glufosinate resistance was successful, leading to severe stunting in most seedlings (Fig 2 A). Selected healthy seedlings were genotyped for the transgene presence, with over 95% true-positive rate (Fig 2 B). Validated transformants, along with a few glufosinate-selected lines, were chosen for further propagation.

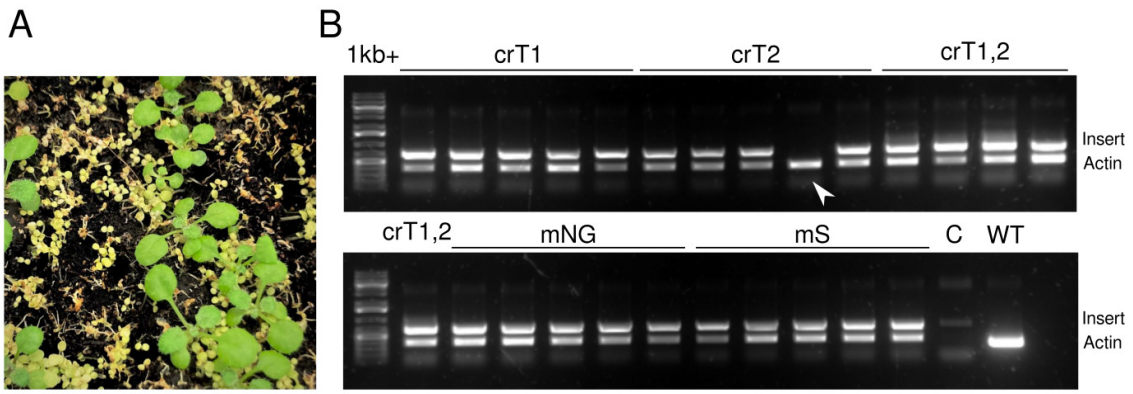


Figure 2. Transformant selection. **A)** Example result of glufosinate selection: healthy putative transformants are easily distinguishable from stunted WT seedlings. **B)** PCR validation of transformants, with actin-specific primers included for reaction control. A glufosinate-selected false positive is marked with an arrow. C - positive control (diluted crT1 plasmid carrying- *E.coli* colony)

4.3.2. Expression validation

The design of the cassettes included fluorescent protein tags, therefore the etiolated T2 seedlings were screened for fluorescence under an epifluorescence microscope. The control mS and mNG lines displayed a readily detectable signal throughout the entire plantlet (Fig 3 A). Unfortunately, no fluorescence was detected in the crT1, crT2, and crT1,2 plants, making direct confirmation of protein production unfeasible. To verify whether the transgenes were being expressed, an RT-PCR was performed (Fig 3 B). All tested plants exhibited an expected band pattern, indicating a successful transgene transcription.

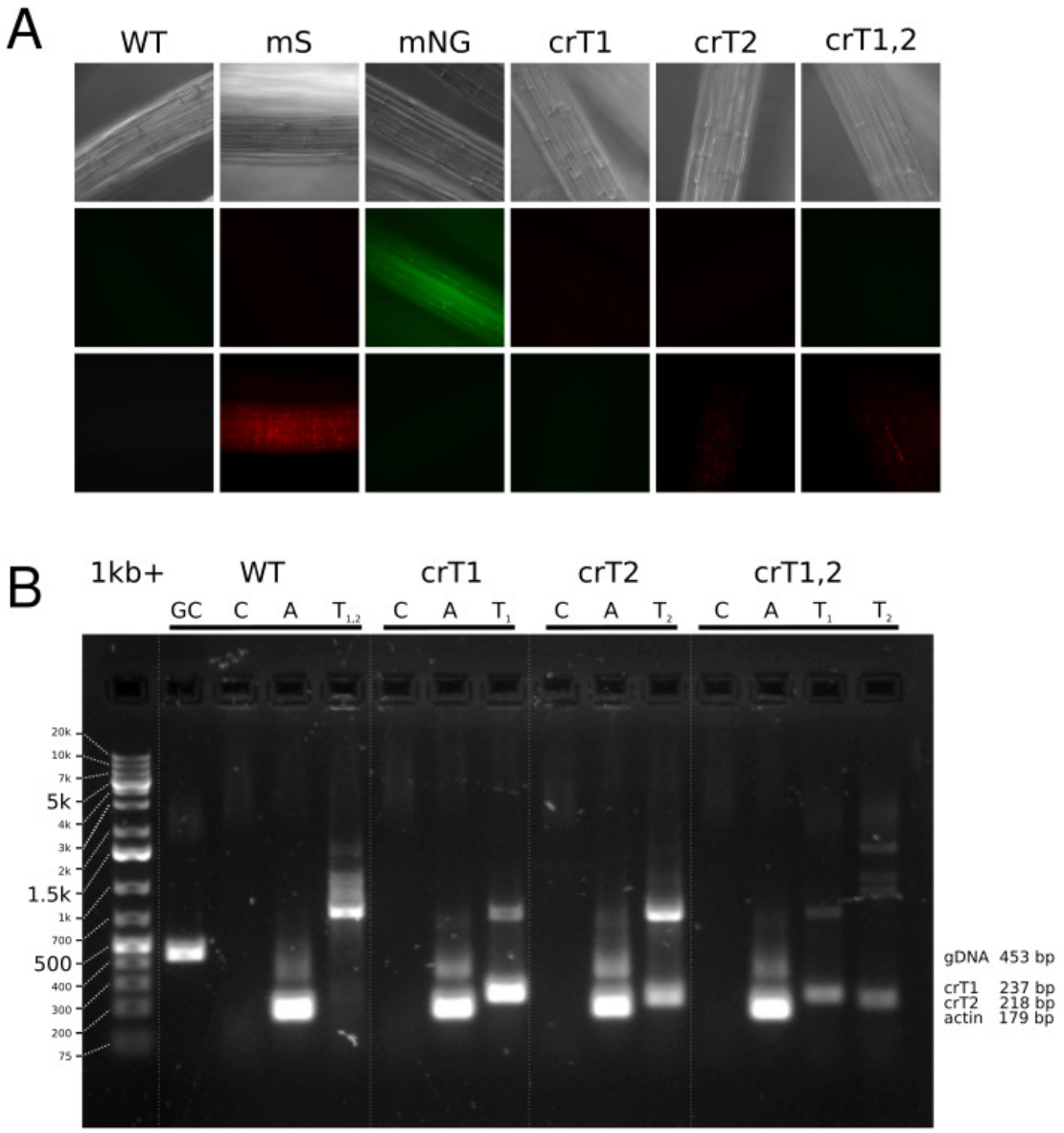


Figure 3. Transgene expression validation. **A)** Fluorescence assessment in stems of etiolated seedlings. No apparent signal is seen in crTn plants, as compared to FP controls. **B)** RT-PCR expression validation. C - gDNA contamination control reaction, with primers flanking the intergenic region. Positive reaction run on genomic DNA described as GC, with expected 453 bp product; A - actin mRNA, 179 bp; T1, T2 - common forward and *crT1* or *crT2*-specific reverse primers; expected products lengths 237 bp and 218 bp, respectively.

4.3.3. Phenotyping of transgenic plants

The preliminary phenotyping experiment was set up prior to the validation of transformation and expression, using glufosinate-selected lines. The results of this experiment support the hypothesis that *crT* plays a significant role in clubroot resistance [Fig 4]. The majority of the crTn plants showed no signs of the disease, in contrast to control transformants, and, to a lesser extent, wild-type plants. The effect did not depend on the *crTn* combinations - both single and double transformants demonstrated high resistance. ANOVA determined significant between-group differences with F = 11.32 and FDR adj. P-value < 6.21e-09, and the post hoc Tukey test showed significant differences between all *crT*-carrying pools, and mS and mNG controls. It is worth mentioning that the offspring of single transformants is expected to revert back to WT in ¼ of cases; however, the affected crTn plants were not genotyped to investigate this possibility. No statistically significant differences between WT and either pool were observed.

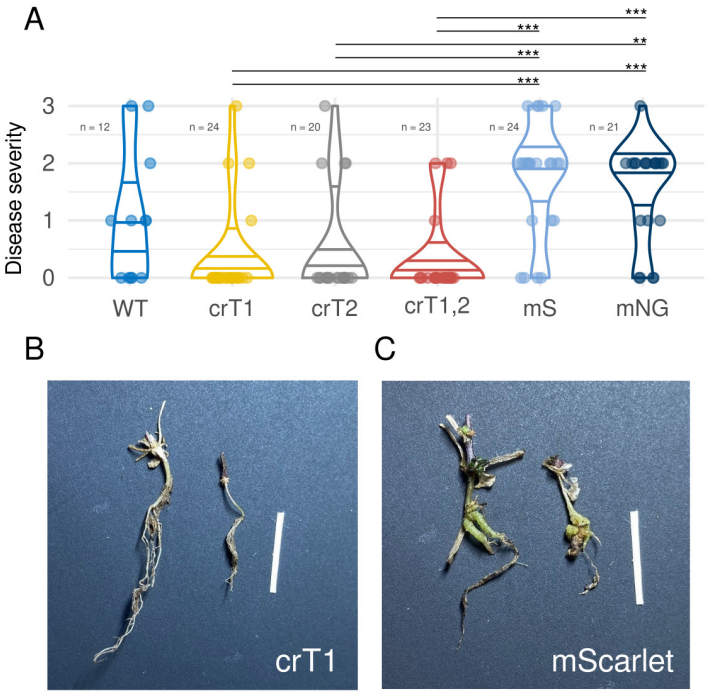


Figure 4. Preliminary experiment phenotyping results. **A)** Phenotype distributions across tested samples. Samples with significantly different means (Tukey test with FDR adjustment) are linked above, ** P<0.001, *** P<0.0001. **B)** Example healthy and **C)** affected root systems from test and control transformants. The white bar marks 1 cm.

4.3.4. Comparison with the results of the independent functional study

Concurrently with this functional study, an independent patent publication confirmed the function of the 'Tosca'-derived resistance allele, referred to as *CRT*, through transgenic expression in the susceptible *B. napus* background (US20210017529A1). The disclosed protein sequence of *CRT* shares a 90.2% pairwise identity with *crT1*. The only source of disparity is the C-terminal part of the proteins. Up to the 1089th position, the sequences are identical, then they diverge at eight positions terminal to *crT1*, and the *CRT* sequence extends for additional 110 AAs (Fig 5).

```
CRT             1001 fcavtddykirsykkdgcllvldyqmsqiplemnfdgldlkihidycrsa   1050
                     ||||||||||||||||||||||||||||||||||||||||||||||||||||
crT1            1001 FCAVTDDYKIRSYKKDGCLLVLDYQMSQIPLEMNFDGLDLKIHIDYCRSA   1050

CRT             1051 kikgwgirileedcssadnrlgypnilphvfeadecneadfvnkllivlg   1100
                     ||||||||||||||||||||||||||||||||||||....:.::
crT1            1051 KIKGWGIRILEEDCSSADNRLGYPNILPHVFEADECNEAGECGRQMM---   1097

CRT             1101 cnaehlklgikrdshptclwilvcgeavplflseiapaqlrgglnivfql   1150

crT1            1098 ---------------------------------------------------   1097

CRT             1151 mvtigilianlvnyftatvhpngwrialggaaiptvilvfgsliicetpt   1200

crT1            1098 ---------------------------------------------------   1097

CRT             1201 sfierkc   1207

crT1            1098 -------   1097
```

Figure 5. Alignment of the C-terminal ends of the *CRT* and *crT1* protein sequences, highlighting the divergence between these two sequences. The remaining parts of the sequences are identical.

Unfortunately, the coding sequence of the *CRT* was not disclosed. Therefore, the available genomic and protein sequences were aligned to the previously assembled 'Tosca' resistance locus sequence and examined using the IGV [Fig 6]. The *CRT* genomic sequence closely matches the genomic *crT1*, with just 12 mismatches (compared to 477 for *crT2*).

Interestingly, the C-terminal difference might be a consequence of alternative splicing near the 3' end. The *CRT* C-terminal AA sequence can be traced and assembled based on the given genomic sequence, but it requires a splicing pattern not supported by any of our RNA-seq reads, which were sourced from roots and leaves of both infected and control plants. However, it is plausible, as the putative *CRT* splice sites do follow a canonical GT-AG pattern in a manner consistent with the disclosed AA sequence. Furthermore, it is important to remember that the genes were expressed in different *Brassica napus* background genotypes - a 'Tosca' x 'BRH-1' derived DH line in the case of *crT1*, and an undisclosed line in the case of *CRT*.

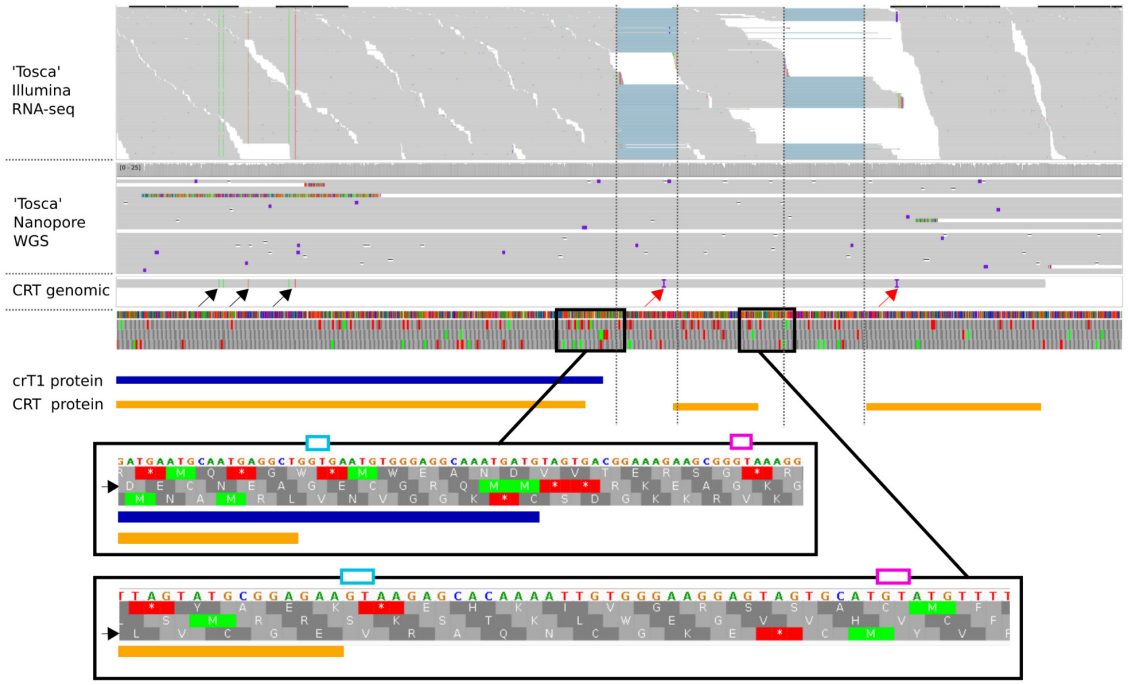


Figure 6. Illustration of the comparison of the C-terminal-coding *crT1* genomic sequence with the *CRT* sequences disclosed in the patent US20210017529A1. The data are aligned in reference to the fragment of assembled resistance locus from 'Tosca'. The tracks depict the following:

1. **'Tosca' Illumina RNA-seq**: Presents alignment of RNA-seq reads from a DH line carrying a resistant allele. Black vertical lines guide the observed splice sites into the subsequent tracks.
2. **'Tosca' Nanopore WGS:** Depicts the alignment of Nanopore sequencing data from 'Tosca', which was used as the base for assembly.
3. **CRT genomic:** This shows the alignment of the patent *CRT* genomic sequence. Dark arrows indicate a cluster of SNVs, evident also in RNA-seq data. These SNVs, as well as other variants within the coding region, were considered during the CDS design process. Variants beyond the *crT1* stop codon are indicated by red arrows.
4. **crT1 protein and CRT protein:** Depict the protein alignments of *crT1* from this study and the patent *CRT* to the genomic sequence.

Black boxes at the bottom enlarge the sequence context of the protein alignment, with the coding phase marked with an arrow. GT donor sites that align with RNA-seq data are marked with pink boxes, while putative alternative sites, consistent with published *CRT* AA sequence, are marked with blue boxes.

Significantly, the patent publication does not mention duplication. It does, however, provide genomic and AA sequences of experimentally validated susceptible alleles, none of which share a high identity with *crT2* (82.2 and 73.5% on the protein level, scattered throughout the sequence), and the genomic sequence of one of them is identical to *B. napus* 'Express 617' A03 orthologue.

4.4. Discussion

In light of the results achieved in this study, and those published in patent US20210017529A1, it can be concluded that the expression of specific variants of the *crT* gene present in the $Crr3^{Tsc}$ locus confers resistance to clubroot. This discovery enriches our repertoire of functionally confirmed TNL-family clubroot resistance genes.

Contrary to the patent publication, our study acknowledges the existence of two paralogous copies of the TNL gene present at the locus, directly stemming from the results of our previous research. Therefore, our results not only agree with but also extend upon the published patent by expanding the spectrum of known resistance alleles.

Significantly, the resistance effect is observed regardless of the transgene configuration, and both duplicates appear functional. This contrasts with the previously described clubroot resistance locus *CRa*/*CRb*, where only one of the locally clustered TNL genes conferred resistance (Hatakeyama et al., 2017).

This apparent functional redundancy prompts intriguing questions regarding the role of duplication. Is there a specific advantage associated with possessing two copies, and if so, what is it? Duplications can alleviate the selective pressure on a single copy, paving the way for neofunctionalization or subfunctionalization. One possibility is that each allele, alone or in combination with the other, responds to different effectors, thereby expanding the pathotypic specificity. Conversely, the response could be amplified if multiple avirulence factors, recognized by different alleles, coexist in a single race. Dual TNL systems have been described in plants (Eitas & Dangl, 2010). For example, *RPS4* and *RRS1* in *A. thaliana* function as a dual system that confers resistance to three distinct pathogens (Narusaka et al., 2009). Some evidence of functional synergy exists in the case of known clubroot resistance

loci. For example, the joint presence of *Crr1b* and *Crr2* is necessary for resistance to certain pathotypes (Hatakeyama, 2015; Suwabe et al., 2003).

Another potential benefit of the duplication could be that a simple increase in the receptor dosage might enhance the response. A dose-dependent efficacy was noted in the case of transgenic expression of *Crr1a* in susceptible *B. rapa* background, and natively the gene is incompletely dominant, suggesting either an additive effect or negative regulation by the susceptible allele (Hatakeyama et al., 2013). Interestingly the effect was not seen in transgenic *A. thaliana*, in which case the gene behaved in a clearly dominant way.

The question of dosage effect could likely be assessed with quantitatively inducible promoters (Zuo et al., 2000). However, this would likely require some optimization, as overexpression of NLRs can trigger autoimmunity (Freh et al., 2022). This phenomenon was demonstrated in studies on, among others, *RPS6* and *VICTR* (Ariga et al., 2017; Gloggnitzer et al., 2014), the closest *crT* homologs in *Arabidopsis thaliana*. Indeed the expression levels of NLRs are tightly regulated, and interfering with this regulation sometimes proves lethal (Freh et al., 2022; Gloggnitzer et al., 2014; Lai & Eulgem, 2018; Lapin et al., 2022). Nonetheless, the *crTn* expression in our experiment was driven by a moderately strong ubq promoter, and no obvious signs of autoimmunity were noted, although it was not systematically measured. Similarly, the *CRT* expression in patent publication was driven by a strong 35S promoter, and the plants were at least viable.

The question arises whether the single-copy expression level is not already optimal to bind the effector at a concentration low enough for the downstream defense response to be sufficient. One observation that could support this idea is the absence of expression upregulation of *crTn* during infection in our RNA-seq data, hinting that the high expression level might not be a primary factor influencing the response. However, this measurement was done quite late after the inoculation, with two copies present. A more focused investigation is needed to address this matter appropriately.

Another tempting speculation is that an elevated dose could hypothetically be more advantageous if the evolution of clubroot specificity were stepwise and the progenitor of the current strong-effect alleles had lower or no affinity to clubroot-specific molecules. This question could potentially be illuminated by examining the evolutionary history of the locus

and functionally evaluating a broader range of alleles, preferably within a broader panel of pathotypes.

However, it is also conceivable that the genes are truly redundant, with one copy gradually losing function altogether due to random mutations and genetic drift. Nevertheless, a relatively large non-synonymous sequence variability between the paralogs, without evidence of nonfunctionalization, suggests this may not be the case here.

4.4.1. Signs of alternative splicing of TNL allele

The single allele validated in the patent US20210017529A1 is nearly identical to *crT1* described in this study. The twelve mismatches observed between the genomic sequences are artifacts of the assembly process that relied on error-prone Nanopore reads. The *CRT* sequence in the patent is the correct one. However, for the *crTn* CDS design, Illumina-based RNA-seq data was considered and used to carefully knit and correct the coding sequences - a process that appears to have been successful, given the 100% identity in the overlapping part.

Protein sequences of *crT1* and *CRT* differ in the C-terminus, with *CRT* being longer. Although the patent publication does not define how the coding sequence was determined, gathered evidence suggests that the observed difference might be due to alternative splicing of the gene at the 3' end. There are known examples of functionally relevant alternative splicing of TNL genes, such as *A. thaliana RPS4*, which requires the presence of both - full-length and alternative transcripts for defense response (X.-C. Zhang & Gassmann, 2003). Our analysis suggests that in this case, the resistance is conferred regardless of the isoform present, at least within the respective backgrounds and pathotype combinations. However, a more direct confirmation of this hypothesis would be necessary.

This result partially contrasts with the *Crr1a*, in which variability at the C-terminal distinguishes resistant from susceptible versions. However, for this gene, these differences stemmed from allelic sequence variability, not alternative splicing of a single allele (Hatakeyama et al., 2022).

4.4.2. Study limitations

Given that this study was conducted in a heterologous system, a degree of caution is necessary for the interpretation of the results. It is notable that in previous studies on *Crr1*, certain discrepancies, discussed earlier, were identified between the effects of the transgene in *A. thaliana* gain of function and *B. rapa* complementation. Nonetheless, they did not affect primary conclusions regarding conferring the resistance.

A more robust approach for such a study could involve the creation of two single and double knockouts in 'Tosca', and two single and double allele substitutions in 'BRH-1'. It would likely yield more definitive results and could potentially shed some light on the putative quantitative aspect, regulatory interplay, and the possible existence of residual resistance factors within the locus.

However, the implementation of this procedure would require a substantial investment of time and financial resources to optimize the transformation protocols and expression systems, with the added complication of the vernalization requirement in winter-type genotypes. Given that functional analysis was not initially planned within the scope of this work, the heterologous expression in *A. thaliana,* proved to be a sufficient approach. It provided convincing and valuable results, particularly when considered alongside published patent data generated in *B. napus*.

One puzzling aspect of the experiment was the low proportion of strongly affected WT lines. It can be partially attributed to the fact that the phenotyping was conducted considerably late after inoculation. As a result, many plants with disintegrated root systems, which may represent the end-stage of the disease, were not scored at all. This could also account for the relatively low proportion of mNG and mS plants with the highest disease index. As the experiment was performed on the offspring of glufosinate-selected plants, without additional transformation validation, a re-evaluation using confirmed lines and a shorter growth period is currently underway.

Another issue encountered was the absence of detectable fluorescence. An important question arises whether the fluorescence is nonexistent or just very weak. It is intriguing to speculate if mechanisms responsible for controlling the TNL levels, such as RNA-mediated

silencing, might be contributing to the apparent lack of fluorescence of the protein tag. These mechanisms would naturally reduce the levels of co-expressed proteins. Another potential problem might be P2A cleavage efficiency, which can be context-dependent (Luke et al., 2015). The absence of fluorescence was not thoroughly investigated since it did not undermine the main conclusions resulting from the experiment.

4.4.3. Future directions

Future studies could expand upon this research by examining the disease response of each gene combination across a broader panel of *P. brassicae* pathotypes. Uncovering any differential signals could pave the way for identifying specific avirulence factors.

The potential avirulence factors could also be identified in a two-hybrid screen of *P. brassicae* transcriptome, or secretome, with *crTn* used as baits (Fields & Song, 1989). Identification of specific avirulence factors could subsequently provide a basis for researching the molecular mechanisms behind the response and facilitate uncovering of the neglected pathogen-oriented side of the story.

In addition to the functional analysis of the current state of the locus, tracing back the evolutionary history of $Crr3^{Tsc}$ could provide valuable insights and help elucidate different scenarios regarding the role of its duplication in clubroot disease. The final chapter of this work initiates a preliminary analysis in this regard.

5. Comparative analysis of the *Crr3^{Tsc}* homologous *loci*

5.1. Introduction

Among the genomes examined thus far, *B. napus* 'Tosca' is the only one in which a duplication of the *crT* gene has been identified. Our previous findings established that the expression of specific copies of the *crT* gene confers resistance to clubroot. Importantly, the resistance effect is observed regardless of the transgene configuration, suggesting that both versions are functional. This apparent functional redundancy provokes intriguing questions about the role of duplication in resistance. Is it functionally relevant now, and if so, in what context? If yes, is it due to the dosage, broader pathotypic specificity, or some other mechanism? Or perhaps the duplication facilitated the evolution of the resistance? While further functional studies may provide additional insights, a complementary approach involves a comparative analysis of the locus across the Brassicaceae genomes. Such analysis would allow us to answer the question of whether the duplication is specific to the 'Tosca' or other clubroot-resistant genotypes. If the duplication is specific only to clubroot-resistant lineage, this would hint at its relevance in the emergence or maintenance of the resistance. In such a case, identification of the genome carrying the most closely related singular allele would provide a good candidate for experimental studies on the emergence of clubroot resistance. On the other hand, identification of clubroot-susceptible genotypes hosting the duplication could pave the way for determining the molecular bases of the resistance by comparison of closely related but functionally divergent alleles of the *crT*.

Intriguingly, a preliminary examination of the genomic fragment homologous to the *Crr3^{Tsc}* resistance locus in *A. thaliana,* reveals an absence of the *crT* gene [Fig 1]. This observation at the very onset of the exploration hints at a complex evolutionary history of the locus that may extend beyond a duplication event. Therefore, a broader comparative analysis could provide additional insight into the evolution of resistance and highlight further directions for experimental evaluation.

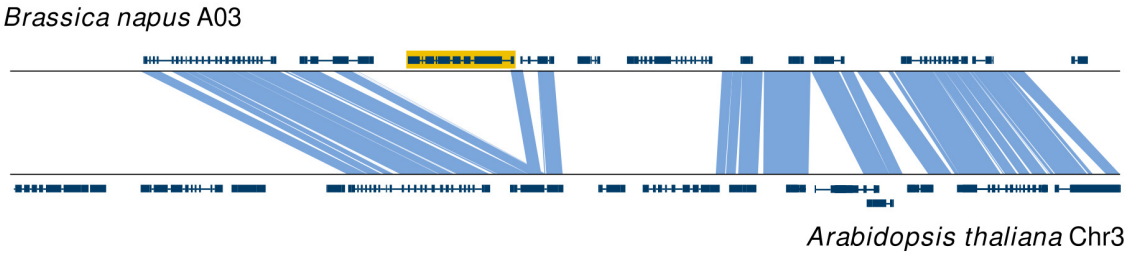


Figure 1. Comparison of synteny between a segment of the *B. napus* Darmor-*bzh* A03 locus homologous to *Crr3*^{Tsc} and a corresponding region on *A. thaliana* Col-0 chromosome 3, conducted in Ensembl. The *crT* gene is highlighted in yellow, situated between *STP6* and related *STP6*-like fragments.

In this chapter, a brief exploratory, comparative analysis of the locus, with a focus on the resistance-conferring *crT* gene is conducted. Firstly, syntenic genomic blocks in a selection of Brassicaceae genomes are identified, and their ancestry is traced back in relation to the *Brassica* lineage polyploidization events. The genic content of the blocks is then assessed, and an analysis of gene presence-absence dynamics is undertaken, with particular attention given to the copy number variation of the *crT* gene.

5.2. Methods

5.2.1. Genomes considered

Table 1 below lists the assemblies considered in the analyses. The last three columns describe which part of the analysis included the particular assembly.

Table 1. List of assemblies used for the analysis. PAV - gene presence-absence variation analysis; TNL CNV - analysis of the copy number variation of *crT*; Phylo - species phylogenetic analysis.

Species	**Genotype**	**Assembly**	**PAV**	**TNL CNV**	**Phylo**
Arabidopsis thaliana	Col-0	GCA_000001735.1 (TAIR10)	yes	NA	yes
Arabidopsis lyrata	MN47	GCA_000004255.1	yes	NA	yes
Arabis alpina	Pajares	GCA_000733195.1	no	NA	yes
Thlaspi arvense	MN106-Ref	GCA_911865555.2 https://doi.org/10.1111/pbi.13775	yes	NA	yes
Raphanus sativus	WK10039	GCA_000801105.3 10.1007/s00122-016-2708-0	yes	NA	yes
Brassica nigra	NI100	https://doi.org/10.1038/s41477-020-0735-y	yes	NA	yes
Brassica oleracea	TO1000	GCF_000695525.1	yes	yes	yes
	OX-heart_923	GCA_018177695.1	no	yes	no
	HDEM	GCA_900416815.2 https://www.nature.com/articles/s41477-018-0289-4	yes	yes	yes
Brassica rapa	Chiifu-401-42	GCF_000309985.2	yes	yes	yes
	Z1	GCA_900412535.3	yes	yes	yes
	PC-fu	GCA_025215035.1	no	yes	no

	SCU_BraROA	GCA_017639395.1	no	yes	no
	CT001	GCA_008629595.1	no	yes	no
	FPsc	GCA_003434825.1	yes	yes	yes
	ECD04	https://doi.org/10.1111/pbi.13827	no	yes	no
Brassica napus	Da-Ae	GCF_020379485.1	yes	yes	yes
	Zs11	GCF_000686985.2	no	yes	no
	GH06	GCA_026770255.1	no	yes	no
	Zy821	GCA_026770265.1	no	yes	no
	Darmor-*bzh*	https://doi.org/10.1093/gigascience/giaa137	yes	yes	yes
	Zs11	https://doi.org/10.1038/s41477-019-0577-7	no	yes	yes
	Zheyou73		no	yes	yes
	Westar		no	yes	yes
	Tapidor		no	yes	yes
	Shengli		no	yes	yes
	QuintaA		no	yes	yes
	No2127		no	yes	yes
	GanganF73		no	yes	yes

5.2.2. Detection of syntenic genomic blocks

Genomic and CDS sequences of *Crr3*^{Tsc} locus gene models based on *B. napus* Darmor-*bzh* 4.1 assembly, as defined in Kopec et al. 2021, were used as queries for blastn and tblastx searches, respectively, against studied genomes (Camacho et al., 2009). Hits were then filtered based on a minimal aligned length of 100 and an E-value of less than 0.001. The results were screened for blocks of hits ordered collinearly with the reference gene order.

5.2.3. Comparison with Ancestral Brassica Karyotype

Chromosome coloring by ancestry was done in reference to TAIR10 gene order on Ancestral Brassica Karyotype (ABK) map by (Murat et al., 2015). This process involved identifying the orthologs of TAIR10 genes included in the ABK among the annotated protein sequences of the genomes being studied, using Orthofinder (Emms & Kelly, 2019). The chromosomal positions of these orthologs were subsequently marked and colored according to the ABK chromosomal affiliation of the corresponding *Arabidopsis* ortholog that uniquely mapped to it, using the ggplot2 R package. The map was manually refined with Inkscape (Inkscape Project. (2020); https://inkscape.org) to enhance visual clarity.

5.2.4. Analysis of gene content and presence-absence variation

The term 'gene' used in the analysis encompasses both genes and annotated pseudogenes. For the selected, annotated genomes, the proteins and annotated pseudogenes from each syntenic block were examined for the best blastp or blastx hit against the *A. thaliana* reference proteome. The regions considered were flanked by orthologs of AT3G06035 and AT3G05900. Flanking orthologs were classified as absent only if the subsequent reference ortholog (beyond AT3G06035 or AT3G05900) was found further up or downstream. Genes present in one genome only were omitted from the analysis. Similarly, any additional structural variation, such as inversions or duplications, was disregarded at this stage.

The resulting binary presence-absence matrix was a foundation of further analysis in the R environment. The genomic blocks were clustered using a complete-linkage method based on Jaccard distance, using the 'dist' and 'hclust' functions from the R *stats* package. The matrix was subsequently visualized, using the *ggplot2* package. The gene (column) order was based on a consensus gene order among the blocks, while the order of the blocks (rows) was based on the results of the clustering.

5.2.5. TNL copy number variation

In all *B. napus*, *B. oleracea,* and *B. rapa* accessions, blocks located on chromosomes A03 and C03 were visually inspected for signs of TNL duplication, using blast results and Integrative Genomics Viewer (IGV) (Robinson et al., 2011).

5.2.6. Phylogenetic analysis

A phylogenetic tree was constructed based on annotated protein sequences using the OrthoFinder package. The tree was subsequently visualized with the online tool ITOL (itol.embl.de).

5.2.7. Sequence comparison of *B. napus* 'Tosca' and *B. rapa ssp. chinensis* 'PC-fu' TNL duplication

Orthologous genomic sequences flanked by specific sequence markers were extracted from the genome of *B. rapa ssp. chinensis* 'PC-fu', and assembled region from *B. napus* 'Tosca'. Sequences were globally aligned using the modified Needleman-Wunsch algorithm as implemented in Emboss Stretcher (Madeira et al., 2022), using default parameters. The number of gaps and mismatches, which respectively represent indel and SNV variability, was averaged in a rolling window of length 1000 and step 50 using a custom R function. Resistance locus genes, sourced from Darmor-*bzh* v4.1 annotation, were aligned with tblastx, to mark the gene positions. The alignment and phylogeny of the protein sequences were performed with Emboss Clustal Omega (Madeira et al., 2022) using the *Brassica rapa* (Brapa_1.0) reference A03 ortholog *Bra001175* as an outgroup. The alignment was visualized in *ggplot2*. The dot-plot was done in Emboss Dotmatcher, with window 1000 and threshold 500, and adjusted visually in Inkscape.

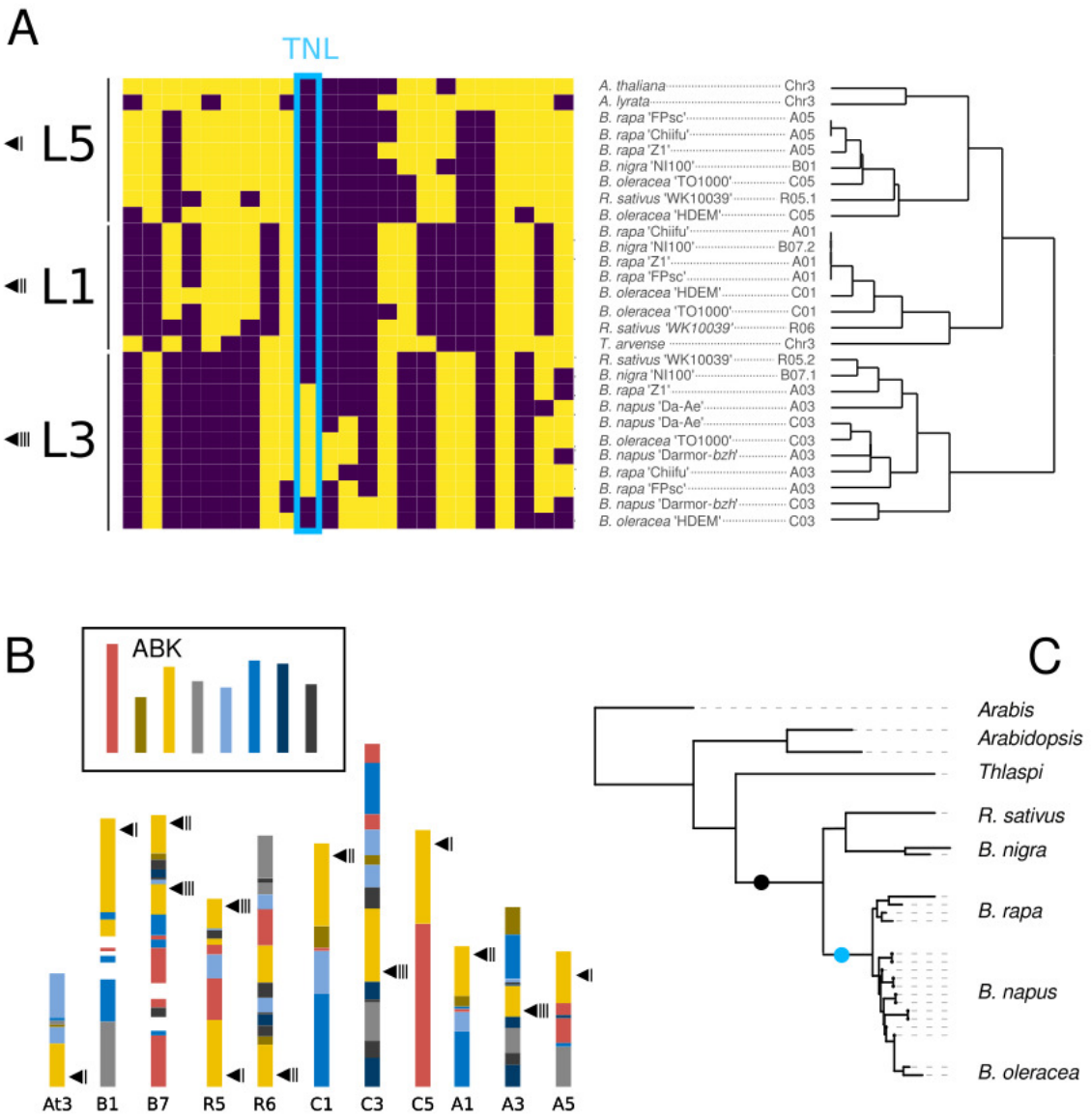


Figure 2. Comparative analysis of *Crr3*^{Tsc}-like genomic blocks across selected Brassicaceae chromosomes. **A)** The presence (yellow) - absence (aubergine) matrix of genes organized in columns. The location of *crT* gene is highlighted with the blue box. The blocks (rows) are arranged according to the hierarchical clustering based on genic content, as depicted on the dendrogram to the right. Block identities on the tree correspond to the names of the chromosomes hosting them. In the case of *B. nigra* and *R. sativus*, where two blocks are present on a single chromosome, the block IDs are suffixed as .1 or .2. Three distinct lineages, likely corresponding to the subgenomes involved in the whole-genome triplication (WGT) event, are indicated on the left side of the matrix. They are annotated with markers corresponding to the chromosomal positions in panel B. The *crT* gene is present only in the L3 lineage, but absent in *R.*

sativus and *B. nigra*. Consequently, within this genic context, the gene can only be found on chromosomes A03, and most of C03, although it is absent in some genotypes, as illustrated by *B.napus* 'Darmor-bzh' and *B. oleracea* 'HDEM'. **B)** Representative chromosomes hosting the *Crr3*^{Tsc}-like blocks among the Brassicaceae. Chromosomal regions are colored by ancestry, in reference to the Ancestral Brassica Karyotype (ABK, depicted in the inset). All *Crr3*^{Tsc}-like blocks are parts of larger segments that descend from ABK chromosome 3, marked in yellow. **C)** The phylogenetic tree of assessed genomes, with a black dot indicating the approximate WGT event, and a blue dot marking the most likely branch for the *crT* insertion.

To study the evolutionary history of the *Crr3*^{Tsc} resistance locus, selected Brassicaceae genomes were examined for syntenic genomic blocks [Fig 2 A]. Single genomic blocks were found on chromosome 3 for *A. thaliana*, *A. lyrata,* and *T. arvense.* All other considered species had three well-preserved blocks within their genomes. While additional, fragmented signals of synteny were detected, they were too divergent to be useful for this analysis.

The genic content of these blocks displayed widespread presence-absence variability, with only one gene, *GLP8*, consistently found in all blocks. When these blocks were clustered based on a gene presence-absence matrix, three distinct lineages emerged, each containing at most one block per genome. The lineages are subsequently referred to as L5, L1, and L3, after the names of C and A chromosomes hosting blocks belonging to the lineages. The origin of these blocks can be traced back to chromosome 3 of the Ancestral Brassica Karyotype [Fig 2 B].

Interestingly, the *crT* gene appears to be present only in the L3 lineage, but absent in *B. nigra* and *R. sativus*. Not all C03 chromosomes harbor the insertion. Among studied genomes, it is absent in *B. oleracea* 'HDEM' and 'OX-heart_923', as well as *B. napus* Darmor-*bzh,* making a reference *B. napus* genome a rather unique example. On the other hand, the presence of the insertion seems to be universal on chromosome A03.

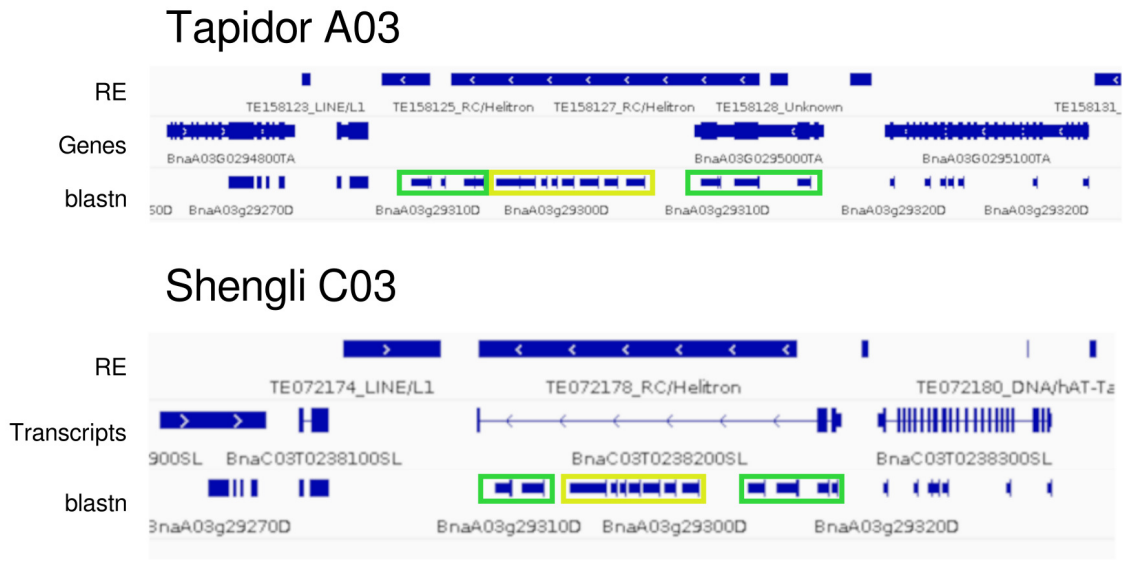


Figure 3. Uncertain annotation of a region homologous to *crT*. IGV tracks represent: RE - repetitive elements, including transposons; blastn - regions homologous to *crT* genes, based on Darmor-*bzh* v4 gene models. Regions of homology to STP6 are in green, TNL in yellow.
Top: An example from the Tapidor A03 chromosome, where the gene is not annotated.
Bottom: An illustration of a questionable annotation from the Shengli C03 chromosome. Exons are annotated in the TNL-flanking regions homologous to *STP6*, while the entire region homologous to the TNL is spanned by a putative intron. This annotation could potentially be explained by the mapping of RNA-seq reads derived from the *STP6*-like fragments, coupled with relatively lax constraints on insert and intron size, particularly in the case of absent TNL expression.
Similar occurrences are commonplace throughout the studied annotation and may potentially be a consequence of pseudogenization. As a result, these instances warrant careful evaluation.

Among the studied A03 and C03 chromosomes harboring the insertion, the gene (or a homologous region), was nearly universally present in a single copy. In multiple *B. napus* genomes, the gene was either not annotated, despite the presence of the insertion, or the quality of the annotated gene structure was questionable [Fig 3]. Interestingly, another genome, *B. rapa ssp. chinensis* (Pak Choi) cultivar 'PC-fu' was found to carry the duplicated *crT*.

Pairwise alignment of the 'Tosca' and 'PC-fu' regions harboring the duplication revealed a very similar structure and high sequence identity, with the most pronounced differences consisting of several larger indels [Fig 4 A, B]. The region located between the *crT* copies seems to be quite conserved in comparison to the genes themselves, which exhibit significant differences. The majority of polymorphic sites within the orthologous TNLs are located towards the 3' end, hosting the LRR domain, and towards the *STP6*-like fragment.

Interestingly, multiple sequence alignment of homologous *crT* protein sequences revealed that the in-paralogous sequences within the genomes bear a higher similarity than the orthologous sequences between them [Fig 4 C].

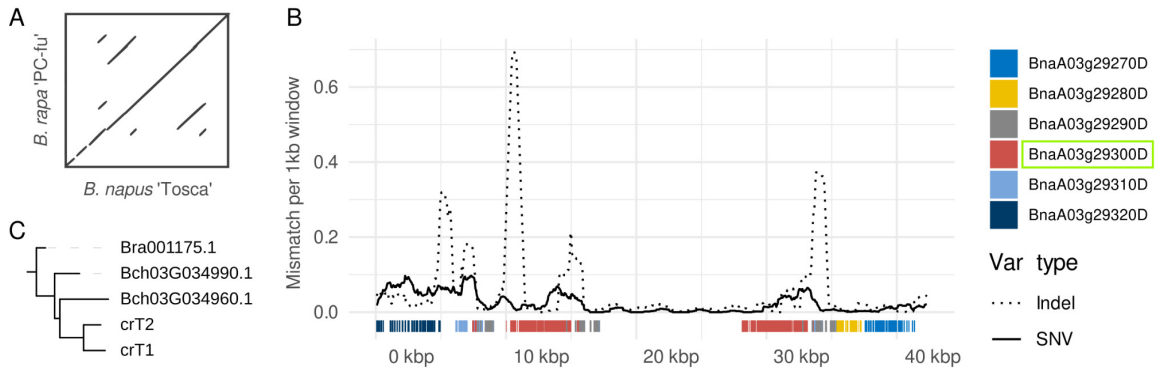


Figure 4. Comparison of *crT* duplication regions in *B. napus* 'Tosca' and *B. rapa ssp. chinensis* 'PC-fu'. **A)** Dot plot alignment, showing high structural similarity between the sequences, with few indels and substitutions. *crT* duplication and *STP6*-like fragment triplication are clearly visible. **B)** Pairwise alignment mismatches due to indel (dotted) and single nucleotide variation (solid) averaged in a 1 kb window with a 50 bp step. The positions of genes based on Darmor-*bzh* v4.1 annotated proteins are included, with the *crT* gene colored brown. **C)** Protein-sequence-based phylogeny of *crT* homologs, with ortholog from *B. rapa* 'Chiifu-401-42' included as an outgroup. Intriguingly, the in-paralogs appear less divergent than orthologs.

5.4. Discussion

5.4.1. Dynamic gene composition of the *Crr3^{Tsc}*-like genomic blocks

The comparative analysis of the *Crr3^{Tsc}* clubroot resistance locus reveals traces of a remarkably complex history. Genomic blocks, syntenic to the resistance locus, exhibit dynamic composition across the Brassicaceae species studied. The presence of three distinct lineages within the *Brassica* clade aligns with the whole genome triplication event that occurred between 13 and 17 million years ago (S. Liu et al., 2014; X. Wang et al., 2011). These lineages likely correspond to ohnologous regions originating from the amalgamated subgenomes. The differences observed between the lineages reflect the post-ploidy fractionation of the genomes, which include extensive gene losses and divergence of retained paralogues (S. Liu et al., 2014). Intriguingly, the *GLP8* gene is invariably present across the lineages. This hints at a selective advantage of copy retention, marking it an interesting target for future studies.

5.4.2. TNL gene is a relatively new element within the locus

The TNL-family gene, *crT*, is found exclusively in blocks from the L3 clade, and only in species derived from the C-genome. This indicates that the gene was inserted into the locus after the WGT event and after *R.sativus* - *B.oleracea* split around 6.5 million years ago, but before the divergence of the A-genome clade, including *B. rapa,* around 5 million years ago [Fig 2 C] (Cheng et al., 2017; S. Liu et al., 2014). Given the limited number of genomes analyzed, these timing-related results require careful interpretation. More definitive conclusions would benefit from expanding the analysis with more data.

The absence of the insertion on certain C03, but not A03 chromosomes, implies that the insertion was present in a lineage that gave rise to the A clade but was not universally fixed on the C03. Alternatively, the C03 insertion could have been secondarily lost. Consequently, the C genomes without the TNL on C03, along with the *Raphanus* and *B. nigra* clades, appear as promising candidates for identifying the ancestral locus of the gene. Tracing back the lineage of *crT* to before the insertion could allow us to assess whether the gene is associated with any other known clubroot resistance loci and therefore hint at the timing of the emergence of the resistance. Previous studies have shown that both of the functionally characterized resistance genes, *CRa* and *Crr1*, likely descend from the same

pre-WGT locus, suggesting that the clubroot specificity of that gene predates the loci (Z. Yang et al., 2022).

5.4.3. The potential mechanism of insertion

Transposon activity offers a plausible explanation for the insertion and the subsequent emergence of the novel resistance locus. The genomic fragment hosting the TNL, along with the adjacent *STP6*-like fragment, is annotated as a Helitron family transposable element in certain *B. napus* assemblies, as seen in Figure 3. Helitrons are a class of DNA transposons utilizing an unusual mode of transposition involving rolling-circle replication. They are known for their tendency to capture host genes or their fragments during relocation (Kapitonov & Jurka, 2007). Moreover, the *STP6*-like fragment alone is scattered throughout the genomes, in some cases forming clusters of proximally multiplicated copies, and is consistently annotated as a transposable element. Similar observations can be made regarding a fragment of the TNL gene. Although the fascinating possibility of transposon involvement warrants further investigation, it provides a plausible explanation for various features of the locus, such as the identity of *STP6*-like fragment, the insertion mechanism, and the peculiar arrangement with homologous *STP6* regions flanking the gene, which in turn could have paved the way for duplication via unequal crossing over.

The involvement of the transposons in the emergence of the novel resistance locus would complement the history of the loci harboring C*Ra* and *Crr1*, for which prevalent nonfunctional alleles are speculated to be a consequence of transposon-driven loss of function (Z. Yang et al., 2022).

5.4.4. Hypothetical role of the insertion in clubroot resistance

An intriguing point of speculation is whether the insertion event itself could catalyze the enhancement or emergence of clubroot resistance. The insertion took place in immense proximity to *STP6*, which in our study and the results of (Z. Wang et al., 2022), showed significant upregulation in the roots of clubroot-challenged, susceptible plants. It prompts an intriguing question: Could this upregulation influence the expression of the TNL gene as well? Some indications suggest this could be the case. For instance, in *Brassica rapa* R-o-18 (Ensembl SCU_BraROA_2.3), the *STP6* and *crT* genes are annotated together as a chimeric gene A03p036470.1_BraROA, with *STP6* located at the 5' end. Such configuration could, in principle, allow for direct induction of *crT* expression by *P. brassicae* infection.

While the upregulation of TNLs is not necessarily required to confer the resistance, it is likely that it could trigger it. Previous studies have demonstrated that upregulation of TNLs can lead to their auto-activation, and there is compelling evidence that the downstream response mirrors signaling triggered by an actual effector (Freh et al., 2022; Heidrich et al., 2013). While continuous overexpression can result in autoimmunity, in this hypothetical scenario, the TNL activation would depend on pathogen-triggered *STP6* induction. Consequently, this situation could allow the functional TNL to be evolutionarily preserved, facilitating the subsequent emergence of allelic specificity. This hypothesis is very speculative, and further studies are necessary to illuminate it.

5.4.5. The duplicated *crT* allele is uncommon among studied *Brassica* genomes

Within the C and A genome species that carry the *crT* insertion, the gene is present in a single copy in nearly all cases. This observation calls for a more careful reexamination of the assemblies along the source reads, as the tandemly duplicated regions might have collapsed during the assembly process. However, a preliminary inspection suggests this was not the case for the *B. napus* genomes included in the analysis. Furthermore, a possibility that requires further exploration is that some single-copy loci may have resulted from secondary deletion. Nevertheless, the relative underrepresentation of the duplicated allele underscores the possibility that the duplication is indeed specific to the clubroot-resistant lineage. Unfortunately, the availability of genomic sequences of clubroot-resistant varieties is currently limited. To my knowledge, the only genotype with qualitative resistance included in the analysis, besides 'Tosca', is the turnip variety 'ECD04', whose resistance depends on different genetic determinants. Sequencing more clubroot-resistant varieties, in particular the remaining known clubroot-resistant turnips, would certainly clarify this issue.

The only studied genome with duplicated *crT*, besides 'Tosca', was *Brassica rapa* 'PC-fu'. An essential question is whether the duplication in 'PC-fu' is associated with clubroot resistance. Unfortunately, the data reviewed thus far are insufficient to provide an answer. To the best of my knowledge, no published study has assessed the clubroot susceptibility of 'PC-fu'. However, a recent study explored the clubroot susceptibility of a large panel of 121 *B. rapa* ssp. *chinensis* accessions, showing a diverse range of disease severity, providing an opportunity to get some insight on that matter (Miao et al., 2023). While the study seems to use internal accession names, the availability of Illumina sequencing data for all lines

provides an opportunity to identify 'PC-fu', or closely related genotypes, if they are included in the set.

5.4.6. Homologous exchanges are involved in the evolution of *crT* paralogs

The proximal duplication in 'PC-fu' has a very similar structure to that of 'Tosca', with a nearly identical linker sequence, suggesting that observed states descend from the same event, rather than independent duplications. Remarkably, the in-paralogs share more similarities than orthologues, which is not expected in classical duplication followed by divergence scenarios. Such a phenomenon is a known occurrence within TNL clusters and can be a consequence of gene conversion or unequal crossing-over events.

Gene conversions are an important component of R gene evolution, which shuffle the LRR domains, potentially changing their binding specificity but at the same time homogenizing them within the genomes (Kuang et al., 2004). An intriguing question to be answered is whether the gene conversion within this particular locus is an ongoing process or an instance of concerted evolution - a consequence of a selective sweep following the homogenization towards an advantageous allele.

5.4.7. Potential widespread pseudogenization of *crT*

Examination of *B. napus crT* insertions revealed a large proportion of unannotated or uncertainly annotated alleles. If these alleles do not code canonically functional TNL i.e. underwent pseudogenization or encode fragmented proteins, it would introduce another layer of variability. Nonfunctionalisation might be a consequence of a 'fitness cost' involved with the presence of functional TNLs in the absence of a compatible pathogen (Barragan & Weigel, 2021; Tian et al., 2003). Apart from being functionally absent, TNL pseudogenes are thought to serve as a potential reservoir of useful variation for the generation of recombination-driven diversity in a similar manner to antibodies (Barragan & Weigel, 2021; Michelmore & Meyers, 1998). Moreover, truncated NLRs have been shown to be involved in pathogen recognition (Nishimura et al., 2017). Whether such processes apply in this case remains to be determined.

Another possibility is that these observations can be attributed to technical issues encountered during the automated annotation of this complex region. To interpret these results, an in-depth inspection of the sequencing, in particular transcriptomic data, is required, and the available gene models will need to be assessed and possibly manually corrected.

In conclusion, despite its relatively recent emergence, the resistance locus shows interesting evolutionary properties, further supporting the necessity of looking at the genomes from a 'pan-perspective'. In general, the *crT* evolutionary dynamics align with the birth and death model of R gene evolution, which encompasses copy number variation and functional diversification driven by random point mutations and homologous recombination under varying selection pressures (Michelmore & Meyers, 1998), which lead to rapid turnover of the genes. Similar evolutionary dynamics can be seen in other adaptive classes of genes, like mammalian olfactory receptors, immunoglobulins, or MHC genes (Nei & Rooney, 2005).

5.4.8. Conclusions and future directions

The comparative analysis revealed that the *crT* duplication is not widespread throughout the *Brassica* lineage; however, it does not provide answers regarding the role of the duplication in clubroot resistance. While our experimental results show that the present-day 'Tosca' alleles are dispensable in conferring resistance to the studied pathotype, the questions regarding their pathotypic specificity spectra and the potential role of duplication in the emergence of the resistance remain open.

The results of the analysis set the firm ground for further studies. Expanding the analysis with a broader set of A03 and C03 haplotypes would undoubtedly provide valuable insights into the evolutionary dynamics of this resistance locus and provide a panel of alleles for further functional studies. Although the availability of long-read assemblies is limited, not all of them were included in the analysis, and their number will undoubtedly grow. Moreover, the volume of accessible short-read genomic data for *B. napus* and its progenitors is quite large, and while the length of the reads might cloud the precise structure of the locus, the presence of the duplication could be assessed by leveraging mapping depth and variant frequency information.

Considering the 'A'-genome source of 'Tosca' resistance and the presence of duplicated alleles in 'PC-fu', the *B. rapa* clade seems like a promising space for exploration. Especially good targets for investigation are genotypes carrying the *Crr3* resistance locus, like *B. rapa* ssp. *rapifera* 'Milan White'. When this accession was tested with the allele-determining marker described in Kopec et al. (2021, Fig 5), it yielded a double product, consistent with the presence of duplication. Other valuable candidates are sources of resistance mapped to the region corresponding to *Crr3*, such as *B. rapa ssp. rapifera* 'Debra', a donor of *CRk*, and *B. rapa ssp. pekinensis* line '85-74' carrying *CRd*. *CRk* has not been fine-mapped, however, due to its different background and varying pathotypic specificity, it was considered distinct from *Crr3 (Sakamoto et al., 2008)*. Similarly, *CRd* was determined to possess a pathotypic specificity different from *Crr3*. *CRd* was postulated to be physically distinct from *Crr3*, to which it is directly adjacent, although no data were shown to support this claim. On the other hand, the *Bra001175*, an orthologue of *crT*, while located outside of the *CRd*, was proposed as one of the candidate genes (Pang et al., 2018). Therefore, it seems important to determine if any of the genes are actually equivalent and ascertain whether the pathotypic variability is not predominantly due to allelic variation.

Besides the duplicated allele, identifying singular alleles sharing a high proportion of genetic information with uncovered duplicates could be very valuable. This task might be complicated by the significant evolutionary dynamics of TNLs. Nonetheless, more known examples could provide an answer as to whether the clubroot resistance emerged before or after insertion and before or after the duplication, and provide additional insights into the evolution of resistance. Moreover, this could potentially enable us to discover novel resistance alleles, or closely related susceptible alleles, which could prove very beneficial for determining molecular bases of recognition specificity or other aspects of the response.

6. Conclusions

In this work, the genetic bases of resistance to *Plasmodiophora brassicae* infection in *Brassica napus* 'Tosca' have been elucidated. The results of this study provide valuable resources for direct incorporation into resistance breeding schemes and open up many new possibilities for further basic and applicational research.

The primary goal of the study, the identification of the locus harboring the resistance factor, was achieved. The evaluation of sequence polymorphism within the $Crr3^{Tsc}$ locus between clubroot-resistant and susceptible lines facilitated the development of molecular markers applicable to the breeding of resistant varieties. Analysis of local gene content, sequence and structural variability, and differential gene expression patterns allowed for the identification of a candidate resistance gene, *crT*.

This gene belongs to the classic R gene TNL family and is proximally duplicated in 'Tosca' accession, while only one copy is present in the susceptible line. Both 'Tosca' copies are highly polymorphic in the pattern-recognizing LRR domain compared to their susceptible-line counterpart. We have demonstrated that the ectopic expression of either of the 'Tosca'-derived *crT* copies is sufficient to confer resistance in *Arabidopsis thaliana*. This discovery broadens the spectrum of functionally confirmed TNL family resistance genes linked to a specific pathogen, thereby opening up vast possibilities for further research. Notably, the apparent functional redundancy hints at the prospect of variable pathotypic specificity of the copies.

Finally, a comparative analysis of the $Crr3^{Tsc}$ locus structure among a broader panel of Brassicaceae genomes has indicated the high evolutionary dynamics of the region, particularly concerning the *crT* gene. Remarkably, the analysis uncovered that the gene is a relatively new element in this genetic neighborhood, and its appearance is likely linked to transposon activity. Moreover, evidence of widespread pseudogenization and copy number variation extending beyond 'Tosca', prompts further research into the evolution of this resistance factor and the functional relevance of its discovered variability.

7. References

Adachi, H., Derevnina, L., & Kamoun, S. (2019). NLR singletons, pairs, and networks: evolution, assembly, and regulation of the intracellular immunoreceptor circuitry of plants. *Current Opinion in Plant Biology*, *50*, 121–131.

Ahmed, A., Munir, S., He, P., Li, Y., He, P., Yixin, W., & He, Y. (2020). Biocontrol arsenals of bacterial endophyte: An imminent triumph against clubroot disease. *Microbiological Research*, *241*, 126565.

Aist, J. R., & Williams, P. H. (1971). The cytology and kinetics of cabbage root hair penetration by *Plasmodiophora brassicae*. *Canadian Journal of Botany. Journal Canadien de Botanique*, *49*(11), 2023–2034.

Ariga, H., Katori, T., Tsuchimatsu, T., Hirase, T., Tajima, Y., Parker, J. E., Alcázar, R., Koornneef, M., Hoekenga, O., Lipka, A. E., Gore, M. A., Sakakibara, H., Kojima, M., Kobayashi, Y., Iuchi, S., Kobayashi, M., Shinozaki, K., Sakata, Y., Hayashi, T., … Taji, T. (2017). NLR locus-mediated trade-off between abiotic and biotic stress adaptation in Arabidopsis. *Nature Plants*, *3*, 17072.

Askarian, H., Akhavan, A., Manolii, V. P., Cao, T., Hwang, S.-F., & Strelkov, S. E. (2021). Virulence Spectrum of Single-Spore and Field Isolates of *Plasmodiophora brassicae* Able to Overcome Resistance in Canola (*Brassica napus*). *Plant Disease*, *105*(1), 43–52.

Ayers, G. W., & Lelacheur, K. E. (1972). Genetics of resistance in rutabaga to two races of *Plasmodiophora brassicae*. *Canadian Journal of Plant Science. Revue Canadienne de Phytotechnie*, *52*(6), 897–900.

Balesdent, M.-H., Gautier, A., Plissonneau, C., Le Meur, L., Loiseau, A., Leflon, M., Carpezat, J., Pinochet, X., & Rouxel, T. (2022). Twenty Years of *Leptosphaeria maculans* Population Survey in France Suggests Pyramiding *Rlm3* and *Rlm7* in Rapeseed Is a Risky Resistance Management Strategy. *Phytopathology*, PHYTO04220108R.

Barragan, A. C., & Weigel, D. (2021). Plant NLR diversity: the known unknowns of pan-NLRomes. *The Plant Cell*, *33*(4), 814–831.

Bayer, P. E., Golicz, A. A., Scheben, A., Batley, J., & Edwards, D. (2020). Plant pan-genomes are the new reference. *Nature Plants*, *6*(8), 914–920.

Bindels, D. S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K. E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M. A., & Gadella, T. W. J., Jr. (2017). mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nature Methods*, *14*(1), 53–56.

Bjornson, M., & Zipfel, C. (2021). Plant immunity: Crosstalk between plant immune receptors. *Current Biology: CB*, *31*(12), R796–R798.

Boutrot, F., & Zipfel, C. (2017). Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annual Review of Phytopathology*, *55*, 257–286.

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., & Görlach, J. (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *The Plant Cell*, *13*(7), 1499–1510.

Bradshaw, J. E., Gemmell, D. J., & Wilson, R. N. (1997). Transfer of resistance to clubroot (*Plasmodiophora brassicae*) to swedes *Brassica napus* L. var. *napobrassica* Peterm) from *B. rapa*. *The Annals of Applied Biology*, *130*(2), 337–348.

Breseghello, F., & Coelho, A. S. G. (2013). Traditional and modern plant breeding methods with examples in rice (*Oryza sativa* L.). *Journal of Agricultural and Food Chemistry*, *61*(35), 8277–8286.

Brun, H., Chèvre, A.-M., Fitt, B. D. L., Powers, S., Besnard, A.-L., Ermel, M., Huteau, V., Marquer, B., Eber, F., Renard, M., & Andrivon, D. (2010). Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in B*rassica napus*. *The New Phytologist*, *185*(1), 285–299.

Buczacki, S. T., Toxopeus, H., Mattusch, P., Johnston, T. D., Dixon, G. R., & Hobolth, L. A. (1975). Study of physiologic specialization in P*lasmodiophora brassicae*: Proposals for attempted rationalization through an international approach. *Transactions of the British Mycological Society*, *65*(2), 295–303.

Büschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J., Töpsch, S., Vos, P., Salamini, F., & Schulze-Lefert, P. (1997). The barley *Mlo* gene: a novel control element of plant pathogen resistance. *Cell*, *88*(5), 695–705.

Caffier, V., Le Cam, B., Al Rifaï, M., Bellanger, M.-N., Comby, M., Denancé, C., Didelot, F., Expert, P., Kerdraon, T., Lemarquand, A., Ravon, E., & Durel, C.-E. (2016). Slow erosion of a quantitative apple resistance to *Venturia inaequalis* based on an isolate-specific Quantitative Trait Locus. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *44*, 541–548.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421.

Cao, T., Manolii, V. P., Strelkov, S. E., Hwang, S.-F., & Howard, R. J. (2009). Virulence and spread of *Plasmodiophora brassicae*[clubroot] in Alberta, Canada. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *31*(3), 321–329.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corréa, M., Da Silva, C., Just, J., Falentin, C., Koh, C. S., Le Clainche, I., Bernard, M., Bento, P., Noel, B., … Wincker, P. (2014). Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science*, *345*(6199), 950–953.

Chang, A., Lamara, M., Wei, Y., Hu, H., Parkin, I. A. P., Gossen, B. D., Peng, G., & Yu, F. (2019). Clubroot resistance gene *Rcr6* in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa*. *BMC Plant Biology*, *19*(1), 224.

Chawla, H. S., Lee, H., Gabur, I., Vollrath, P., Tamilselvan-Nattar-Amutha, S., Obermeier, C., Schiessl, S. V., Song, J.-M., Liu, K., Guo, L., Parkin, I. A. P., & Snowdon, R. J. (2021). Long-read sequencing reveals widespread intragenic structural variants in a recent allopolyploid crop plant. *Plant Biotechnology Journal*, *19*(2), 240–250.

Cheng, F., Liang, J., Cai, C., Cai, X., Wu, J., & Wang, X. (2017). Genome sequencing supports a multi-vertex model for Brassiceae species. *Current Opinion in Plant Biology*, *36*, 79–87.

Chen, K., Wang, Y., Zhang, R., Zhang, H., & Gao, C. (2019). CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annual Review of Plant Biology*, *70*, 667–697.

Clarke, W. E., Higgins, E. E., Plieske, J., Wieseke, R., Sidebottom, C., Khedikar, Y., Batley, J., Edwards, D., Meng, J., Li, R., Lawley, C. T., Pauquet, J., Laga, B., Cheung, W., Iniguez-Luy, F., Dyrszka, E., Rae, S., Stich, B., Snowdon, R. J., … Parkin, I. A. P. (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. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *129*(10), 1887–1899.

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology*, *16*(6), 735–743.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., & Varshney, R. K. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in Plant Science*, *22*(11), 961–975.

Czubatka-Bieńkowska, A., Kaczmarek, J., Marzec-Schmidt, K., Nieróbca, A., Czajka, A., & Jędryczka, M. (2020). Country-Wide qPCR Based Assessment of *Plasmodiophora brassicae* Spread in Agricultural Soils and Recommendations for the Cultivation of Brassicaceae Crops in Poland. *Pathogens*, *9*(12). https://doi.org/10.3390/pathogens9121070

Dakouri, A., Zhang, X., Peng, G., Falk, K. C., Gossen, B. D., Strelkov, S. E., & Yu, F. (2018). Analysis of genome-wide variants through bulked segregant RNA sequencing reveals a major gene for resistance to *Plasmodiophora brassicae* in *Brassica oleracea*. *Scientific Reports*, *8*(1), 17657.

Daun, J. K. (2011). 1 - Origin, Distribution, and Production. In J. K. Daun, N. A. M. Eskin, & D. Hickling (Eds.), *Canola* (pp. 1–27). AOCS Press.

de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H. A. J., & Thomma, B. P. H. J. (2010). Conserved fungal *LysM*

effector *Ecp6* prevents chitin-triggered immunity in plants. *Science*, *329*(5994), 953–955.
Delmas, C. E. L., Fabre, F., Jolivet, J., Mazet, I. D., Richart Cervera, S., Delière, L., & Delmotte, F. (2016). Adaptation of a plant pathogen to partial host resistance: selection for greater aggressiveness in grapevine downy mildew. *Evolutionary Applications*, *9*(5), 709–725.
Delourme, R., Bousset, L., Ermel, M., Duffé, P., Besnard, A. L., Marquer, B., Fudal, I., Linglin, J., Chadœuf, J., & Brun, H. (2014). Quantitative resistance affects the speed of frequency increase but not the diversity of the virulence alleles overcoming a major resistance gene to *Leptosphaeria maculans* in oilseed rape. *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases*, *27*, 490–499.
Deschamps, S., Llaca, V., & May, G. D. (2012). Genotyping-by-Sequencing in Plants. *Biology*, *1*(3), 460–483.
Diederichsen, E., Beckmann, J., Schondelmeier, J., & Dreyer, F. (2006). GENETICS OF CLUBROOT RESISTANCE IN *BRASSICA NAPUS* ´MENDEL´. *Acta Horticulturae*, *706*, 307–312.
Diederichsen, E., Frauen, M., Linders, E. G. A., Hatakeyama, K., & Hirai, M. (2009). Status and Perspectives of Clubroot Resistance Breeding in Crucifer Crops. *Journal of Plant Growth Regulation*, *28*(3), 265–281.
Diederichsen, E., Frauen, M., & Ludwig-Müller, J. (2014). Clubroot disease management challenges from a German perspective. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *36*(sup1), 85–98.
Diederichsen, E., & Sacristan, M. D. (1996). Disease response of resynthesized *Brassica napus* L. lines carrying different combinations of resistance to *Plasmodiophora brassicae* Wor. *Plant Breeding = Zeitschrift Fur Pflanzenzuchtung*, *115*(1), 5–10.
Dixon, G. R. (2009). The Occurrence and Economic Impact of *Plasmodiophora brassicae* and Clubroot Disease. *Journal of Plant Growth Regulation*, *28*(3), 194–202.
Dixon, G. R. (2014). Clubroot (Plasmodiophora brassicaeWoronin) – an agricultural and biological challenge worldwide. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *36*(sup1), 5–18.
Dixon, G. R. (2017). The origins of edible brassicas. *Plantsman* , *16*(3), 180–185.
Djidjou-Demasse, R., Moury, B., & Fabre, F. (2017). Mosaics often outperform pyramids: insights from a model comparing strategies for the deployment of plant resistance genes against viruses in agricultural landscapes. *The New Phytologist*, *216*(1), 239–253.
Dolatabadian, A., Bayer, P. E., Tirnaz, S., Hurgobin, B., Edwards, D., & Batley, J. (2020). Characterization of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. *Plant Biotechnology Journal*, *18*(4), 969–982.
Donald, C., & Porter, I. (2009). Integrated Control of Clubroot. *Journal of Plant Growth Regulation*, *28*(3), 289–303.
EFSA Panel on Genetically Modified Organisms (GMO), Mullins, E., Bresson, J.-L., Dalmay, T., Dewhurst, I. C., Epstein, M. M., Firbank, L. G., Guerche, P., Hejatko, J., Moreno, F. J., Naegeli, H., Nogué, F., Rostoks, N., Sánchez Serrano, J. J., Savoini, G., Veromann, E., Veronesi, F., Fernandez, A., Gennaro, A., … Schoonjans, R. (2022). Criteria for risk assessment of plants produced by targeted mutagenesis, cisgenesis and intragenesis. *EFSA Journal. European Food Safety Authority*, *20*(10), e07618.
Eitas, T. K., & Dangl, J. L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Current Opinion in Plant Biology*, *13*(4), 472–477.
Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238.
Fields, S., & Song, O.-K. (1989). A novel genetic system to detect protein–protein interactions. *Nature*, *340*(6230), 245–246.
Fredua-Agyeman, R., Hwang, S.-F., Zhang, H., Falak, I., Huang, X., & Strelkov, S. E. (2021). Clubroot resistance derived from the European *Brassica napus* cv. “Tosca” is not effective against virulent *Plasmodiophora brassicae* isolates from Alberta, Canada. *Scientific Reports*, *11*(1), 14472.
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/*

Netherlands Journal of Plant Breeding, *211*(2), 201–213.
Fredua-Agyeman, R., Yu, Z., Hwang, S.-F., & Strelkov, S. E. (2020). Genome-Wide Mapping of Loci Associated With Resistance to Clubroot in *Brassica napus* ssp. *napobrassica* (Rutabaga) Accessions From Nordic Countries. *Frontiers in Plant Science*, *11*, 742.
Freh, M., Gao, J., Petersen, M., & Panstruga, R. (2022). Plant autoimmunity-fresh insights into an old phenomenon. *Plant Physiology*, *188*(3), 1419–1434.
Friedt, W., & Snowdon, R. (2010). Oilseed Rape. In J. Vollmann & I. Rajcan (Eds.), *Oil Crops* (pp. 91–126). Springer New York.
Fuchs, H. (1996). Identification of a gene in *Arabidopsis thaliana* Controlling resistance to clubroot (*Plasmodiophora brassicae*) and characterization of the resistance response. *Molecular Plant-Microbe Interactions: MPMI*, *9*(2), 091.
Gabur, I., Chawla, H. S., Lopisso, D. T., von Tiedemann, A., Snowdon, R. J., & Obermeier, C. (2020). Gene presence-absence variation associates with quantitative *Verticillium longisporum* disease resistance in *Brassica napus*. *Scientific Reports*, *10*(1), 4131.
Gabur, I., Chawla, H. S., Snowdon, R. J., & Parkin, I. A. P. (2019). Connecting genome structural variation with complex traits in crop plants. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *132*(3), 733–750.
Gan, C., Yan, C., Pang, W., Cui, L., Fu, P., Yu, X., Qiu, Z., Zhu, M., Piao, Z., & Deng, X. (2022). Identification of Novel Locus *RsCr6* Related to Clubroot Resistance in Radish (*Raphanus sativus* L.). *Frontiers in Plant Science*, *13*, 866211.
Garber, R. C. (1979). *The Ultrastructure of Meiosis in Plasmodiophora Brassicae (Plasmodiophorales)*. National Research Council of Canada.
García-Nafría, J., Watson, J. F., & Greger, I. H. (2016). IVA cloning: A single-tube universal cloning system exploiting bacterial In Vivo Assembly. *Scientific Reports*, *6*, 27459.
Gloggnitzer, J., Akimcheva, S., Srinivasan, A., Kusenda, B., Riehs, N., Stampfl, H., Bautor, J., Dekrout, B., Jonak, C., Jiménez-Gómez, J. M., Parker, J. E., & Riha, K. (2014). Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense. *Cell Host & Microbe*, *16*(3), 376–390.
Golicz, A. A., Bayer, P. E., Barker, G. C., Edger, P. P., Kim, H., Martinez, P. A., Chan, C. K. K., Severn-Ellis, A., McCombie, W. R., Parkin, I. A. P., Paterson, A. H., Pires, J. C., Sharpe, A. G., Tang, H., Teakle, G. R., Town, C. D., Batley, J., & Edwards, D. (2016). The pangenome of an agronomically important crop plant *Brassica oleracea*. *Nature Communications*, *7*, 13390.
Golicz, A. A., Bayer, P. E., & Edwards, D. (2015). Skim-based genotyping by sequencing. *Methods in Molecular Biology* , *1245*, 257–270.
Gravot, A., Grillet, L., Wagner, G., Jubault, M., Lariagon, C., Baron, C., Deleu, C., Delourme, R., Bouchereau, A., & Manzanares-Dauleux, M. J. (2011). Genetic and physiological analysis of the relationship between partial resistance to clubroot and tolerance to trehalose in *Arabidopsis thaliana*. *The New Phytologist*, *191*(4), 1083–1094.
Han, G.-Z. (2019). Origin and evolution of the plant immune system. *The New Phytologist*, *222*(1), 70–83.
Hasan, M. J., & 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 / National Research Council Canada = Genome / Conseil National de Recherches Canada*, *59*(10), 805–815.
Hatakeyama, K. (2015). Characterization and Utilization of the Clubroot Resistant Genes in Chinese Cabbage (*Brassica rapa* L.). 한국균학회소식: 학술대회논문집, 33–33.
Hatakeyama, K., Niwa, T., Kato, T., Ohara, T., Kakizaki, T., & Matsumoto, S. (2017). The tandem repeated organization of NB-LRR genes in the clubroot-resistant *CRb* locus in *Brassica rapa* L. *Molecular Genetics and Genomics: MGG*, *292*(2), 397–405.
Hatakeyama, K., Suwabe, K., Tomita, R. N., 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.
Hatakeyama, K., Yuzawa, S., Tonosaki, K., Takahata, Y., & Matsumoto, S. (2022). Allelic variation of a clubroot resistance gene (*Crr1a*) in Japanese cultivars of Chinese cabbage (*Brassica rapa* L.). *Breeding Science*, *72*(2), 115–123.

Heidrich, K., Tsuda, K., Blanvillain-Baufumé, S., Wirthmueller, L., Bautor, J., & Parker, J. E. (2013). *Arabidopsis* TNL-WRKY domain receptor *RRS1* contributes to temperature-conditioned *RPS4* auto-immunity. *Frontiers in Plant Science*, *4*, 403.

Hurgobin, B., Golicz, A. A., Bayer, P. E., Chan, C.-K. K., Tirnaz, S., Dolatabadian, A., Schiessl, S. V., Samans, B., Montenegro, J. D., Parkin, I. A. P., Pires, J. C., Chalhoub, B., King, G. J., Snowdon, R., Batley, J., & Edwards, D. (2018). Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant Biotechnology Journal*, *16*(7), 1265–1274.

Hwang, S.-F., Howard, R. J., Strelkov, S. E., Gossen, B. D., & Peng, G. (2014). Management of clubroot (*Plasmodiophora brassicae*) on canola (*Brassica napus*) in western Canada. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *36*(sup1), 49–65.

Hwang, S.-F., Strelkov, S. E., Feng, J., Gossen, B. D., & Howard, R. J. (2012). *Plasmodiophora brassicae*: a review of an emerging pathogen of the Canadian canola (*Brassica napus*) crop. *Molecular Plant Pathology*, *13*(2), 105–113.

Jajor, E., Mrówczyński, M., Bartkowiak-Broda, I., Broniarz, J., Danielewicz, J., Dobrzycka, A., Dworzańska, D., Fiedler, Ż., Gorzała, G., Horoszkiewicz-Janka, J., Kierzek, R., Korbas, M., (rolnictwo), E. M., (agronomia), K. M., Matysiak, K., (rolnictwo), A. M., Muśnicki, C., Obst, A., Perek, A., … Zamojska, J. (2017). *Metodyka integrowanej ochrony i produkcji rzepaku ozimego oraz jarego dla doradców: opracowanie zbiorowe*. Instytut Ochrony Roślin - Państwowy Instytut Badawczy.

Javed, M. A., Schwelm, A., Zamani-Noor, N., Salih, R., Silvestre Vañó, M., Wu, J., González García, M., Heick, T. M., Luo, C., Prakash, P., & Pérez-López, E. (2023). The clubroot pathogen *Plasmodiophora brassicae*: A profile update. *Molecular Plant Pathology*, *24*(2), 89–106.

Jørgensen, I. H. (1992). Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. *Euphytica/ Netherlands Journal of Plant Breeding*, *63*(1), 141–152.

Kageyama, K., & Asano, T. (2009). Life Cycle of *Plasmodiophora brassicae*. *Journal of Plant Growth Regulation*, *28*(3), 203–211.

Kapitonov, V. V., & Jurka, J. (2007). Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends in Genetics: TIG*, *23*(10), 521–529.

Katche, E., Quezada-Martinez, D., Katche, E. I., Vasquez-Teuber, P., & Mason, A. S. (2019). Interspecific hybridization for *Brassica crop* improvement. *Crop Breeding, Genetics and Genomics*. https://doi.org/10.20900/cbgg20190007

Kopec, P. M., Mikolajczyk, K., Jajor, E., Perek, A., Nowakowska, J., Obermeier, C., Chawla, H. S., Korbas, M., Bartkowiak-Broda, I., & Karlowski, W. M. (2021). Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in *Brassica napus* cv. Tosca. *Frontiers in Plant Science*, *12*, 639631.

Korbas, M., Jajor, E., & Budka, A. (2009). Clubroot (*Plasmodiophora Brassicae*) - A Threat for Oilseed Rape. *Journal of Plant Protection Research*, *49*(4). https://doi.org/10.2478/v10045-009-0071-8

Kris A. Wetterstrand, M. S. (2019, March 13). *DNA sequencing costs: Data*. Genome.gov; NHGRI. https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data

Krzymański, J. (1968). Variation in thioglucosides in rapeseed meal (*Brassica napus*). *Meeting of the Associate Commitees of National*.

Kuang, H., Woo, S.-S., Meyers, B. C., Nevo, E., & Michelmore, R. W. (2004). Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *The Plant Cell*, *16*(11), 2870–2894.

Lai, Y., & Eulgem, T. (2018). Transcript-level expression control of plant NLR genes. *Molecular Plant Pathology*, *19*(5), 1267–1281.

Lapin, D., Johanndrees, O., Wu, Z., Li, X., & Parker, J. E. (2022). Molecular innovations in plant TIR-based immunity signaling. *The Plant Cell*, *34*(5), 1479–1496.

Lemarié, S., Robert-Seilaniantz, A., Lariagon, C., Lemoine, J., Marnet, N., Levrel, A., Jubault, M., Manzanares-Dauleux, M. J., & Gravot, A. (2015). *Camalexin* contributes to the partial resistance of *Arabidopsis thaliana* to the biotrophic soilborne protist *Plasmodiophora brassicae*. *Frontiers in Plant Science*, *6*, 539.

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*. *Frontiers in Plant Science*, *7*, 1483.
Liu, L., Qin, L., Cheng, X., Zhang, Y., Xu, L., Liu, F., Tong, C., Huang, J., Liu, S., & Wei, Y. (2020). Comparing the Infection Biology of *Plasmodiophora brassicae* in Clubroot Susceptible and Resistant Hosts and Non-hosts. *Frontiers in Microbiology*, *11*, 507036.
Liu, L., Qin, L., Zhou, Z., Hendriks, W. G. H. M., Liu, S., & Wei, Y. (2020). Refining the Life Cycle of *Plasmodiophora brassicae*. *Phytopathology*, *110*(10), 1704–1712.
Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A. P., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, X., Wang, J., Lu, K., Fang, Z., Bancroft, I., Yang, T.-J., Hu, Q., Wang, X., Yue, Z., … Paterson, A. H. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, *5*, 3930.
Lüders, W. (2017). *Analyses of Virulence of European Isolates of Clubroot (Plasmodiophora Brassicae Wor.) and Mapping of Resistance Genes in Rapeseed (Brassica Napus L.)*. Universitätsbibliothek.
Ludwig-Müller, J., Bennett, R. N., Kiddle, G., Ihmig, S., Ruppel, M., & Hilgenberg, W. (1999). The host range of *Plasmodiophora brassicae* and its relationship to endogenous glucosinolate content. *The New Phytologist*, *141*(3), 443–458.
Ludwig-Müller, J., Prinsen, E., Rolfe, S. A., & Scholes, J. D. (2009). Metabolism and Plant Hormone Action During Clubroot Disease. *Journal of Plant Growth Regulation*, *28*(3), 229–244.
Luke, G. A., Roulston, C., Tilsner, J., & Ryan, M. D. (2015). Growing uses of 2A in plant biotechnology. In *Biotechnology*. InTech.
Lu, K., Wei, L., Li, X., Wang, Y., Wu, J., Liu, M., Zhang, C., Chen, Z., Xiao, Z., Jian, H., Cheng, F., Zhang, K., Du, H., Cheng, X., Qu, C., Qian, W., Liu, L., Wang, R., Zou, Q., … Li, J. (2019). Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. *Nature Communications*, *10*(1), 1154.
Lu, Y., & Tsuda, K. (2021). Intimate Association of PRR- and NLR-Mediated Signaling in Plant Immunity. *Molecular Plant-Microbe Interactions: MPMI*, *34*(1), 3–14.
Lv, H., Fang, Z., Yang, L., Zhang, Y., & Wang, Y. (2020). An update on the arsenal: mining resistance genes for disease management of *Brassica* crops in the genomic era. *Horticulture Research*, *7*, 34.
Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*(W1), W276–W279.
Malinowski, R., Smith, J. A., Fleming, A. J., Scholes, J. D., & Rolfe, S. A. (2012). Gall formation in clubroot-infected *Arabidopsis* results from an increase in existing meristematic activities of the host but is not essential for the completion of the pathogen life cycle. *The Plant Journal: For Cell and Molecular Biology*, *71*(2), 226–238.
Malinowski, R., Truman, W., & Blicharz, S. (2019). Genius Architect or Clever Thief-How *Plasmodiophora brassicae* Reprograms Host Development to Establish a Pathogen-Oriented Physiological Sink. *Molecular Plant-Microbe Interactions: MPMI*, *32*(10), 1259–1266.
Mason, A. S., Higgins, E. E., Snowdon, R. J., Batley, J., Stein, A., Werner, C., & Parkin, I. A. P. (2017). A user guide to the *Brassica* 60K Illumina Infinium™ SNP genotyping array. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *130*(4), 621–633.
Matsumoto, E., Ueno, H., Aruga, D., Sakamoto, K., & Hayashida, N. (2012). Accumulation of three clubroot resistance genes through marker-assisted selection in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Journal of the Japanese Society for Horticultural Science*, *81*(2), 184–190.
Miao, L., Xi, D., Li, X., Zhang, C., Gao, L., Li, Y., Zhu, Y., & Zhu, H. (2023). Genetic Diversity Analysis of Non-Heading Chinese Cabbage of Resistance to Clubroot Disease Based on SNP Molecular Markers. *Agronomy*, *13*(1), 245.
Michelmore, R. W., Christopoulou, M., & Caldwell, K. S. (2013). Impacts of resistance gene genetics, function, and evolution on a durable future. *Annual Review of Phytopathology*, *51*, 291–319.
Michelmore, R. W., & Meyers, B. C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, *8*(11), 1113–1130.
Montarry, J., Cartier, E., Jacquemond, M., Palloix, A., & Moury, B. (2012). Virus adaptation to quantitative plant resistance: erosion or breakdown? *Journal of Evolutionary Biology*, *25*(11),

2242–2252.
Murat, F., Louis, A., Maumus, F., Armero, A., Cooke, R., Quesneville, H., Roest Crollius, H., & Salse, J. (2015). Understanding Brassicaceae evolution through ancestral genome reconstruction. *Genome Biology*, *16*, 262.
Narusaka, M., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2010). The floral inoculating protocol: a simplified *Arabidopsis thaliana* transformation method modified from floral dipping. *Plant Biotechnology* , *27*(4), 349–351.
Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., & Narusaka, Y. (2009). *RRS1* and *RPS4* provide a dual Resistance-gene system against fungal and bacterial pathogens. *The Plant Journal: For Cell and Molecular Biology*, *60*(2), 218–226.
Neik, T. X., Barbetti, M. J., & Batley, J. (2017). Current Status and Challenges in Identifying Disease Resistance Genes in *Brassica napus*. *Frontiers in Plant Science*, *8*, 1788.
Nei, M., & Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, *39*, 121–152.
Neuhauser, S., Kirchmair, M., Bulman, S., & Bass, D. (2014). Cross-kingdom host shifts of phytomyxid parasites. *BMC Evolutionary Biology*, *14*(1), 33.
Ngou, B. P. M., Ahn, H.-K., Ding, P., & Jones, J. D. G. (2021). Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature*, *592*(7852), 110–115.
Ngou, B. P. M., Ding, P., & Jones, J. D. G. (2022). Thirty years of resistance: Zig-zag through the plant immune system. *The Plant Cell*, *34*(5), 1447–1478.
Nishimura, M. T., Anderson, R. G., Cherkis, K. A., Law, T. F., Liu, Q. L., Machius, M., Nimchuk, Z. L., Yang, L., Chung, E.-H., El Kasmi, F., Hyunh, M., Osborne Nishimura, E., Sondek, J. E., & Dangl, J. L. (2017). TIR-only protein *RBA1* recognizes a pathogen effector to regulate cell death in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(10), E2053–E2062.
Ochoa, J. C., Mukhopadhyay, S., Bieluszewski, T., Jedryczka, M., Malinowski, R., & Truman, W. (2022). Natural variation in *Arabidopsis* responses to *Plasmodiophora brassicae* reveals an essential role for *RPB1*. In *bioRxiv* (p. 2022.10.06.511079). https://doi.org/10.1101/2022.10.06.511079
Olszak, M., Truman, W., Stefanowicz, K., Sliwinska, E., Ito, M., Walerowski, P., Rolfe, S., & Malinowski, R. (2019). Transcriptional profiling identifies critical steps of cell cycle reprogramming necessary for *Plasmodiophora brassicae*-driven gall formation in *Arabidopsis*. *The Plant Journal: For Cell and Molecular Biology*, *97*(4), 715–729.
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*). *Frontiers in Plant Science*, *9*, 653.
Pang, W., Liang, Y., Zhan, Z., Li, X., & Piao, Z. (2020). Development of a Sinitic Clubroot Differential Set for the Pathotype Classification of *Plasmodiophora brassicae*. *Frontiers in Plant Science*, *11*, 568771.
Piao, Z., Ramchiary, N., & Lim, Y. P. (2009). Genetics of Clubroot Resistance in *Brassica* Species. *Journal of Plant Growth Regulation*, *28*(3), 252–264.
Piętka, T., Krzymański, J., Krótka, K., Bartkowiak-Broda, I., & Others. (2014). Double low white mustard (*Sinapis alba* L. syn. *Brassica hirta*) is a source of protein and oil. *Oilseed Crops*, *35*, 21–35.
Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., Durel, C.-E., & Delourme, R. (2017). Quantitative Resistance to Plant Pathogens in Pyramiding Strategies for Durable Crop Protection. *Frontiers in Plant Science*, *8*, 1838.
Poland, J., & Rutkoski, J. (2016). Advances and Challenges in Genomic Selection for Disease Resistance. *Annual Review of Phytopathology*, *54*(1), 79–98.
Quan, J., & Tian, J. (2009). Circular polymerase extension cloning of complex gene libraries and pathways. *PloS One*, *4*(7), e6441.
rapa Genome Sequencing Project Consortium, B., Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., & Others. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, *43*(10), 1035–1039.
Rashid, A., Ahmed, H. U., Xiao, Q., Hwang, S. F., & Strelkov, S. E. (2013). Effects of root exudates

and pH on *Plasmodiophora brassicae* resting spore germination and infection of canola (*Brassica napus* L.) root hairs. *Crop Protection* , *48*, 16–23.
Rennie, D. C., Holtz, M. D., Turkington, T. K., Leboldus, J. M., Hwang, S.-F., Howard, R. J., & Strelkov, S. E. (2015). Movement of *Plasmodiophora brassicae* resting spores in windblown dust. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *37*(2), 188–196.
Řičařová, V., Kaczmarek, J., Strelkov, S. E., Kazda, J., Lueders, W., Rysanek, P., Manolii, V., & Jedryczka, M. (2016). Pathotypes of *Plasmodiophora brassicae* causing damage to oilseed rape in the Czech Republic and Poland. *European Journal of Plant Pathology / European Foundation for Plant Pathology*, *145*(3), 559–572.
Rimbaud, L., Papaïx, J., Barrett, L. G., Burdon, J. J., & Thrall, P. H. (2018). Mosaics, mixtures, rotations or pyramiding: What is the optimal strategy to deploy major gene resistance? *Evolutionary Applications*, *11*(10), 1791–1810.
Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24–26.
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*). *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *117*(5), 759–767.
Samans, B., Snowdon, R., & Mason, A. S. (2018). Homoeologous Exchanges and Gene Losses Generate Diversity and Differentiate the *B. napus* Genome from that of Its Ancestors. In S. Liu, R. Snowdon, & B. Chalhoub (Eds.), *The Brassica napus Genome* (pp. 131–148). Springer International Publishing.
Schiessl, S., Huettel, B., Kuehn, D., Reinhardt, R., & Snowdon, R. J. (2017). Targeted deep sequencing of flowering regulators in *Brassica napus* reveals extensive copy number variation. *Scientific Data*, *4*, 170013.
Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., Davidson, M. W., & Wang, J. (2013). A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nature Methods*, *10*(5), 407–409.
Snowdon, R., Lühs, W., & Friedt, W. (2007). Oilseed Rape. In C. Kole (Ed.), *Oilseeds* (pp. 55–114). Springer Berlin Heidelberg.
Somé, A., Manzanares, M. J., 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 Pathology*, *45*(3), 432–439.
Song, J.-M., Guan, Z., Hu, J., Guo, C., Yang, Z., Wang, S., Liu, D., Wang, B., Lu, S., Zhou, R., Xie, W.-Z., Cheng, Y., Zhang, Y., Liu, K., Yang, Q.-Y., Chen, L.-L., & Guo, L. (2020). Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of *Brassica napus*. *Nature Plants*, *6*(1), 34–45.
Song, X., Wei, Y., Xiao, D., Gong, K., Sun, P., Ren, Y., Yuan, J., Wu, T., Yang, Q., Li, X., Nie, F., Li, N., Feng, S., Pei, Q., Yu, T., Zhang, C., Liu, T., Wang, X., & Yang, J. (2021). *Brassica carinata* genome characterization clarifies U's triangle model of evolution and polyploidy in *Brassica*. *Plant Physiology*, *186*(1), 388–406.
Stefansson, B. R., & Hougen, F. W. (1961). Note on the isolation of rape plants with seed oil free from erucic acid. *Canadian Journal of Plant Science* https://cdnsciencepub.com/doi/abs/10.4141/cjps61-028
Stefansson, B. R., & Kondra, Z. P. (1975). Tower summer rape. *Canadian Journal of Plant Science. Revue Canadienne de Phytotechnie*, *55*(1), 343–344.
Stein, A., Coriton, O., Rousseau-Gueutin, M., Samans, B., Schiessl, S. V., Obermeier, C., Parkin, I. A. P., Chèvre, A.-M., & Snowdon, R. J. (2017). Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*. *Plant Biotechnology Journal*, *15*(11), 1478–1489.
Strelkov, S. E., Hwang, S.-F., Manolii, V. P., Cao, T., Fredua-Agyeman, R., Harding, M. W., Peng, G., Gossen, B. D., Mcdonald, M. R., & Feindel, D. (2018). Virulence and pathotype classification of *Plasmodiophora brassicae* populations collected from clubroot resistant canola (*Brassica napus*) in Canada. *Canadian Journal of Plant Pathology. Revue Canadienne de Phytopathologie*, *40*(2),

284–298.
Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., Fukuoka, H., Matsumoto, S., & Hirai, M. (2003). Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *107*(6), 997–1002.
Tang, H., Zhang, X., Tong, C., Chalhoub, B., Liu, S., & Lyons, E. (2018). From Alpha-Duplication to Triplication and Sextuplication. In S. Liu, R. Snowdon, & B. Chalhoub (Eds.), *The Brassica napus Genome* (pp. 99–109). Springer International Publishing.
Tian, D., Traw, M. B., Chen, J. Q., Kreitman, M., & Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature*, *423*(6935), 74–77.
Tonu, N. N., Wen, R., Song, T., Guo, X., Murphy, L. A., Gossen, B. D., Yu, F., & Peng, G. (2023). Canola with Stacked Genes Shows Moderate Resistance and Resilience against a Field Population of *Plasmodiophora brassicae (*Clubroot) Pathotype X. *Plants*, *12*(4). https://doi.org/10.3390/plants12040726
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 Molecular Biology*, *80*(6), 621–629.
U, Nagaharu. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Japanese Botany*, *7*, 389–452.
USDA Foreign Agricultural Service. (2023, May 12). *Oilseeds: World markets and trade*. USDA Foreign Agricultural Service. https://www.fas.usda.gov/data/oilseeds-world-markets-and-trade
Vlot, A. C., Sales, J. H., Lenk, M., Bauer, K., Brambilla, A., Sommer, A., Chen, Y., Wenig, M., & Nayem, S. (2021). Systemic propagation of immunity in plants. *The New Phytologist*, *229*(3), 1234–1250.
Vollrath, P., Chawla, H. S., Schiessl, S. V., Gabur, I., Lee, H., Snowdon, R. J., & Obermeier, C. (2021). A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *134*(4), 1217–1231.
Wagner, G., Laperche, A., Lariagon, C., Marnet, N., Renault, D., Guitton, Y., Bouchereau, A., Delourme, R., Manzanares-Dauleux, M. J., & Gravot, A. (2019). Resolution of quantitative resistance to clubroot into QTL-specific metabolic modules. *Journal of Experimental Botany*, *70*(19), 5375–5390.
Walerowski, P., Gündel, A., Yahaya, N., Truman, W., Sobczak, M., Olszak, M., Rolfe, S., Borisjuk, L., & Malinowski, R. (2018). Clubroot Disease Stimulates Early Steps of Phloem Differentiation and Recruits SWEET Sucrose Transporters within Developing Galls. *The Plant Cell*, *30*(12), 3058–3073.
Wallenhammar, A.-C. (1996). Prevalence of *Plasmodiophora brassicae* in a spring oilseed rape growing area in central Sweden and factors influencing soil infestation levels. *Plant Pathology*, *45*(4), 710–719.
Waminal, N. E., Perumal, S., Liu, S., Chalhoub, B., Kim, H. H., & Yang, T.-J. (2018). Quantity, Distribution, and Evolution of Major Repeats in *Brassica napus*. In S. Liu, R. Snowdon, & B. Chalhoub (Eds.), *The Brassica napus Genome* (pp. 111–129). Springer International Publishing.
Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Wang, J., Wang, X., Freeling, M., Pires, J. C., Paterson, A. H., Chalhoub, B., … *Brassica rapa* Genome Sequencing Project Consortium. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, *43*(10), 1035–1039.
Wang, Z., Megha, S., Kebede, B., Kav, N. N. V., & Rahman, H. (2022). Genetic and molecular analysis reveals that two major loci and their interaction confer clubroot resistance in canola introgressed from rutabaga. *The Plant Genome*, *15*(3), e20241.
Warwick, S. I. (2011). Brassicaceae in Agriculture. In R. Schmidt & I. Bancroft (Eds.), *Genetics and Genomics of the Brassicaceae* (pp. 33–65). Springer New York.
Weigel, D., & Glazebrook, J. (2006). Transformation of *Agrobacterium* using the freeze-thaw method. *CSH Protocols*, *2006*(7). https://doi.org/10.1101/pdb.prot4666
Williams, P. H., & Others. (1966). A system for the determination of races of *Plasmodiophora*

brassicae that infect cabbage and rutabaga. *Phytopathology*, *56*(6), 624–626.
Woronin, M. (1878). *Plasmodiophora brassicae*, Urheber der Kophpflanzenhernie. *Jahrbucher Fur Wissenschaftliche Botanik*, *1*, 548–574.
Yang, J., Liu, D., Wang, X., Ji, C., Cheng, F., Liu, B., Hu, Z., Chen, S., Pental, D., Ju, Y., Yao, P., Li, X., Xie, K., Zhang, J., Wang, J., Liu, F., Ma, W., Shopan, J., Zheng, H., … Zhang, M. (2016). The genome sequence of allopolyploid *Brassica juncea* and analysis of differential homoeolog gene expression influencing selection. *Nature Genetics*, *48*(10), 1225–1232.
Yang, Z., Jiang, Y., Gong, J., Li, Q., Dun, B., Liu, D., Yin, F., Yuan, L., Zhou, X., Wang, H., Wang, J., Zhan, Z., Shah, N., Nwafor, C. C., Zhou, Y., Chen, P., Zhu, L., Li, S., Wang, B., … Zhang, C. (2022). R gene triplication confers European fodder turnip with improved clubroot resistance. *Plant Biotechnology Journal*, *20*(8), 1502–1517.
Yuan, M., Jiang, Z., Bi, G., Nomura, K., Liu, M., Wang, Y., Cai, B., Zhou, J.-M., He, S. Y., & Xin, X.-F. (2021). Pattern-recognition receptors are required for NLR-mediated plant immunity. *Nature*, *592*(7852), 105–109.
Yu, X., Feng, B., He, P., & Shan, L. (2017). From Chaos to Harmony: Responses and Signaling upon Microbial Pattern Recognition. *Annual Review of Phytopathology*, *55*, 109–137.
Zamani-Noor, N., & Rodemann, B. (2018). Reducing the build-up of P*lasmodiophora brassicae* inoculum by early management of oilseed rape volunteers. *Plant Pathology*, *67*(2), 426–432.
Zhang, X.-C., & Gassmann, W. (2003). *RPS4*-mediated disease resistance requires the combined presence of *RPS4* transcripts with full-length and truncated open reading frames. *The Plant Cell*, *15*(10), 2333–2342.
Zhang, X., Dodds, P. N., & Bernoux, M. (2017). What Do We Know About NOD-Like Receptors in Plant Immunity? *Annual Review of Phytopathology*, *55*, 205–229.
Zheng, X., Koopmann, B., Ulber, B., & von Tiedemann, A. (2020). A global survey on diseases and pests in Oilseed rape—current challenges and innovative strategies of control. *Frontiers in Agronomy*, *2*. https://doi.org/10.3389/fagro.2020.590908
Ziegler, D. J., Khan, D., Pulgar-Vidal, N., Parkin, I. A. P., Robinson, S. J., & Belmonte, M. F. (2023). Genomic asymmetry of the *Brassica napus* seed: epigenetic contributions of DNA methylation and small RNAs to subgenome bias. *The Plant Journal: For Cell and Molecular Biology*. https://doi.org/10.1111/tpj.16254
Zuo, J., Niu, Q. W., & Chua, N. H. (2000). Technical advance: An estrogen receptor-based transactivator *XVE* mediates highly inducible gene expression in transgenic plants. *The Plant Journal: For Cell and Molecular Biology*, *24*(2), 265–273.