INSTITUTE OF PLANT GENETICS
POLISH ACADEMY OF SCIENCES

Strzeszyńska 34, 60-479 Poznań

Tel.: +48 616550200 or 616550255 E-mail: office@igr.poznan.pl www.igr.poznan.pl
VAT: PL 7811621455 REGON: 000326204 BDO: 000017736

Poznań,15 September 2023

Prof. dr hab. Małgorzata Jędryczka
Department of Pathogen Genetics and Plant Resistance
Institute of Plant Genetics, PAS, Poznań

Review of the PhD Thesis

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

by Piotr Kopeć, MSc.

The work was done at the Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Faculty of Biology of Adam Mickiewicz University in Poznań. The Thesis were supervised by Prof. dr hab. Wojciech Karłowski.

Legal basis

The activities were carried out on the basis of Art. 29 of the Act dated 30 April 2010 on research institutes (Journal of Laws of 2022, item 498), in accordance with the Art. 179 of the Act dated July 3, 2018. Provisions introducing the Act on Higher Education and Science (Journal of Laws of 2018, item 1669) and the Act dated March 14, 2003 on academic degrees and academic titles and on academic degrees and titles in the field of art (Journal of Laws of 2017, item 1789).

Funding

The work was supported by the National Science Centre NCN 2016/22/M/NZ9/00604 and the project POWR 03.02.00-00-1022/16.

Project collaborators and their involvement in the NCN project

The NCN project belonged to HARMONIA grants, it was entitled 'Analiza genetyczna odporności na *Plasmodiophora brassicae* Wor. u rzepaku (*Brassica napus* L.) z wykorzystaniem wysokoprzepustowych technologii sekwencjonowania i mapowania' (Genetic analysis of oilseed rape (*Brassica napus* L.) resistance to *Plasmodiophora brassicae* Wor. using high throughput sequencing technologies and mapping'. The head of the project was Prof. Wojciech Karłowski, the supervisor of the PhD Thesis. Project coordinator was Assoc. Prof. Katarzyna Mikołajczyk, employed by the Institute of Plant Breeding and Acclimatization - National Research Institute (NRI), Branch Poznań (IHAR-PIB). The NCN project lasted from June 2017 till June 2020. The University of Adam Mickiewicz was the leader of the consortium, with two partners: IHAR-PIB and the Institute of Plant Protection - NRI (IOR-PIB). The foreign partner was Dr. Christian Obermeier from Justus-Liebig-Universität Giessen (JLUG), Institut für Pflanzenbau und Pflanzenzüchtung, Germany.

Structure of the Thesis - formal evaluation

The work has a semi-typical layout. It deviates from the standard Ph.D. Thesis (monograph) as it contains already published work and its discussion as well as two chapters of the unpublished work, which relate closely to the preliminary experiment. All parts of the work are divided into clear chapters and subchapters with information in accordance with the title. The work contains the following parts: Acknowledgements, Abstract, Abstract in Polish, List of scientific works included in the dissertation, Funding, Abbreviations followed by the main part of the Thesis composed of the Introduction and Aims of the Thesis. Then the manuscript 'Local Duplication of TIR-NBS-LRR Gene Marks Clubroot Resistance in *Brassica napus* cv. Tosca and Supplementary Figures of the manuscript are presented followed by the additional comments and extensive discussion of the experiment and its results. The next chapter again starts with the Introduction which is a bit confusing, but this introduction relates solely to the 'Functional study of the duplicated TNL gene", that constitutes the second part of this study (unpublished, prepared for publication). This part is supplied with separate Materials and methods followed by the Results and Discussion. The same organisation concerns the next chapter on 'Comparative analysis of the $Crr3^{Tce}$ homologous *loci*'. Again this subchapter has its own Introduction, Methods, does not contain the materials as they have been described already, followed by Results and Discussion. The whole Thesis is finalized by overall Conclusions and the References.

In total, the work contains pages 103, including 65 pages of the novel text, 18 pages of the publication in co-authorship (including 3 pages of references and 2 pages of Supplementary figures), 10 pages of authors' statements, 10 page of references. The volume contains Author's contribution statement and the statements of the co-authors presented in separate documents. Each sub-chapter contains own numbering of figures and tables, own introduction and discussion with additional comments, discussion of study limitations and future directions of work.

The individual input of PhD student

The PhD Dissertation contains the publication with ten co-authors from five scientific institutions, including four above mentioned (AMU, IHAR, IOR, JLUG as well as the co-author from Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany). According the statement of the PhD student Piotr Kopeć he was involved in the phenotyping and genotyping of the mapping population, he processed the RNA-seq samples and generated part of the Nanopore sequencing data. He collected processed analysed and integrated the phenotyping , genotyping and sequencing data. These analyses included genetic map construction, QTL mapping, resistance gene characterization and sequence polymorphism evaluation, gene expression study and genetic marker design. The PhD student prepared all figures and co-wrote the article. The percent of the input has not been expressed, but based on the statement a great part of the work can be attributed to the Candidate (Piotr Kopeć).

Summarising, in formal terms, the monograph has the structure which deviates from the usual doctoral monograph but it is appropriate for thesis of this kind. The use of already published publication makes the form different from the standard, but it is fully acceptable. The work is prepared carefully, it has all the necessary parts, presented in the right order. Illustrations and tables are properly prepared, described and catalogued. Mistakes are extremely rare. I dully accept the form of the Thesis.

Scope and topic of the PhD Thesis

The work concerns winter oilseed rape (*Brassica napus* L.), which is an important cultivated plant from the economic point of view. Rapeseed occupies the area of over 1 million hectares in Poland and is the main domestic source of vegetable oil. This important crop plant is exceptionally valuable in temperate and cool climate conditions and is one of the basis of agricultural economy in Europe, China, Canada and Australia. Interestingly, due to the diversification of crops, it is also introduced into unusual climatic conditions, e.g. in Brazil, Egypt, Tunisia and some African countries. Brassicaceae plants are an important component of the food security in many countries, including the availability of vegetable oils for food. They contain valuable vitamins, minerals and many health-promoting substances useful in phytotherapy. Moreover rapeseed is used for numerous technical purposes, primarily it is a valuable source of renewable fuels. The European Commission directive imposes on all Member States the obligation to increase the consumption of renewable energy sources in their national economy, so the availability of oilseed rape products is a great bonus for the country. Recent problems with food security, including the inability to collect and distribute oilseed rape from Ukraine have shown the importance of this crop.

The potential of plants belonging to the Brassica genus for the economy of many countries is the result of progress in research and breeding. The significant share of oilseed rape in agricultural fields constitutes a great counterweight to cereal crops and interrupts or weakens the development of pathogens of cultivated grasses. However, too intensive use of oilseed rape in crop rotation leads to problems with plant health and often implies the use of pesticides, which cause problems to human health and is harmful to the environment.

The PhD Thesis concerns clubroot – the devastating disease of oilseed rape which causes serious economic problems due to significant yield losses and long-term deterioration of an agricultural soil. The low soil quality caused by the high infestation with the pathogen *Plasmodiophora brassicae* results in long-term inability of agricultural soil to be used in rotations containing any of the Brassica crops, including mustard and brassica vegetables, such as cabbage, cauliflower or broccoli. Brassica crops serve as phytosanitary plants for cereals, so the inability to use them in a cropping system leads to agrotechnical complications. The problems are both at the scale of single farms located in areas with soil infested by the pathogen, but they are also observed at the national and international scale. Recent studies (2016-2020) done in central Europe revealed significant differences in disease incidences at infected fields between countries and sites per country. Approximately 51.4% of fields had a low occurrence of the disease (DI ≤ 30%), 32.3% had moderate clubroot incidence (30% < DI < 60%), and 16.3% showed a high incidence of clubroot disease (DI >= 60%), which means the pathogen is detectable in most of the fields. The severe disease symptoms are observed mostly when the number of spores exceeded $1x10^3$ per plant.

Due to the long life of resting spores, the assessment of the pathogen abundance in agricultural fields can serve as a guideline for disease control at the country-wide level or the regional scale. Poland is one of the countries which has a country-wide map, showing the spread of the pathogen in agricultural soils. Based on the calculation of gene copies specific to *P. brassicae* obtained by qPCR it was possible to formulate the recommendations for Polish farmers in respect to the cultivation guidelines. The studies showed a very high risk of yield losses in defined regions of north, south-west and central Poland and an urgent need to undertake intensive preventative measures.

High use of pesticides in agriculture is not accepted by many 'environmentally conscious' societies. Plant protection against clubroot is the example of skillful use of knowledge about the genetic resistance of plants. Numerous resistance genes and QTLs have been identified in A, B and C genomes of Brassica plants and related plant families. The identification and skillful use of these sources of resistance gives an advantage to the farmers and consumers of Brassica crops as well as the whole society, due to environment protection. Considering all these issues, the PhD Thesis of Piotr Kopeć M.Sc. is well suited to the socially accepted plant breeding strategies, which are based on genetic resistance.

The Thesis aimed at the elucidation and characterisation of the genetic basis of clubroot resistance available in cv. 'Tosca'. This clubroot-resistance source is highly desirable for breeders but the background of this resistance was unknown. The study performed by Piotr Kopeć was expanded to broader understanding of the host-pathogen interaction and the development of molecular markers for marker-assisted selection of new clubroot resistant cultivars, without the need of burdensome resistance testing.

The research topics discussed in the PhD Thesis of Piotr Kopeć fall within the scope of important basic research, the aim of which is to identify, locate and describe resistance genes of oilseed rape to clubroot disease, which is crucial for the cultivation of this crop worldwide. The work has an important theoretical as well as applicatory aspect – the development of markers for disease resistance, to be used in plant breeding of Brassica crops. The research presented in the Thesis creates new knowledge which is a prerequisite for obtaining a doctoral title in biological sciences.

PhD Thesis evaluation

Abstract

In my opinion the summary lacks more detailed information about the main results obtained in this interesting work. The background of the study occupies half of the text, they are followed by materials and methods mixed with the results, but the results are described very generally. The summary should be more specific and informative.

Abbreviations

Informative, very useful, some abbreviations were used in this study but they were not listed eg. CDS, mS, mNG, mFP, crT and other abbreviations used in Chapter 4 (explained in page 61).

General introduction

Informative and clearly written but lacks the details on resistance genes and the novel, international system of their classification. The Introduction contains the general information on oilseed rape, clubroot disease including the pathogen in general, mechanisms of disease resistance and arms race between the pathogen and the plant. The information about the pathotypes and different systems of pathotype classification is superficial and it only lists different systems, without any explanation. Taking into consideration the pathotype-specific resistance present in Brassica crops, the reader expects more details in this chapter. Pinpointing of the novel resistance *crT* found in $Crr3^{TSC}$ locus is one of the main aims of the Thesis. This is why the reader expects more information on resistance genes found to date, as compared to the general (half page) information on qualitative and quantitative resistance found in A and C genomes. The Author concentrates rather on the sources of resistance genes and not on the genes themselves and their location on chromosomes. I would expect a list of known resistance genes and information on gene clusters.

The breeding of cultivars resistant to clubroot and the introduction of clubroot resistance (CR) genes started when the first resistance sources were found in European fodder turnip cultivars belonging to *B. rapa* and were introgressed into *B. napus*. Many genes have been cloned and mapped, especially on A genome, primarily on chromosome A03, where at least 10 loci other than *Crr3* have been described in papers dated between 2006 and 2018. At least seven other genes were found on A01, A02, A05, A06, A08. Some research teams studying CR resistance sources termed differently the same *loci*/genes. The example of this misuse has been shown by the Author in page 23. I quite like the subchapter on genomic technologies bridging the gap between basic research and breeding. Interesting and clear. All in all, the Introduction is well written. A broader and deeper insight in both the pathogen diversity and resistance genes would leave me feel more satisfied.

Aims of the thesis

The chapter starts with the information what project was realized and who was the project leader and main participant. I find it strange as this has nothing to do with the aim of PhD Thesis. Then the overall aim is given followed by more detailed explanation. We learn about the aims but there is no scientific hypothesis given. The research steps are logically presented, but it looks like the research plan was formed on spot, as a logical consequence of the previous results rather than like the research planned from the very beginning.

Chapter: "Local duplication of TIR-NBS-LRR Gene marks clubroot resistance in *Brassica napus* cv. 'Tosca'"

This chapter consists of the paper published in 2021 in Frontiers of Plant Sciences by Kopeć PM, Mikołajczyk K., Jajor E., Perek A., Nowakowska J., Obermeyer C., Chawla HS, Korbas M, Bartkowiak-Broda I and Karłowski WM (IF 6,627, MEiN 100 pts). The clubroot resistance locus derived from cv. 'Tosca' was characterized. The population of double haploids was derived from cv. 'Tosca' and the disease severity was evaluated, genotyped with Brassica 60K array and selected using SSR/SCAR marker. The construction of the genetic map helped to narrow down the resistance locus to 0.4 cM fragment on A03 chromosome corresponding to the region previously described as *Crr3*. The close study of this region using Oxford Nanopore long read genome resequencing and RNA-seq revealed the duplication of resistance gene of the TIR-NBS-LRR (TNL) type. Clubroot-resistant and susceptible inoculated and control DH lines showed differences at the transcription level. The Authors concluded that the TNL gene was a promising candidate for the resistance factor. The publication is very interesting and gives and insight in the valuable source of clubroot resistance. The only (but substantial) drawback of this study is the lack of purification and a poor characterization of the pathotype used in this study. The pathogen chosen for the phenotyping of clubroot resistance was evaluated solely using the basic and crude system designed by Somé et al. In contrast to many well-known and popular pathotyping systems in *P. brassicae*, the Somé system uses only three genotypes: *Brassica napus* var. *napus* (fodder rape) Dc 101 Nevin, *B. napus* var. *rapifera* Dc 130 Wilhelmsburger and *B. napus* var. *oleifera* cultivar Brutor. Moreover, the spores for inoculation were collected from one or more galls collected in the field without any further purification. This lead to a very basic knowledge on the population of *P. brassicae* used in the experiment. The Authors call it 'P3-dominant environmental sample', what does not characterise the pathotype to the current standards. The pathotype used has not been challenged against cv. Mendel to check whether it belongs to the group P3 or P3+. Detailed characterization of the pathotype would bring the better insight into this study. From one hand the Authors used DH lines (plants were homozygous) and from the other hand the pathogen was a mixture of "who knows what".

There are many ways of pathogen purification. The ideal way is single sporing, which is difficult in *P. brassicae*. In contrast, obtaining single club isolates is relatively easy and could be used in this study. The pathogen is a living organism reacting to photoperiod and the temperature. So do the plants. The control of these parameters in the glasshouse is possible only to some extent. It is strange that the Authors did not test all DH plants in one run but the inoculation was done from April to August in batches of 68 plants. This resulted in uncontrolled variability and the necessity of result adjustments to reduce the batch effect. In phytopathological experiments such things should be avoided. I guess this effect did not change the final result of this study. I would like to get the opinion of the Candidate on the effect of the variability of plant disease phenotyping on bioinformatic analyses of experiment results: To which extent this variability can be misleading when data from many sources are integrated? Otherwise, the experiments were well designed. The Supplementary Figures are hard to read from the Thesis as they are copied from the article but the publication is available online and can be studied without any problems.

Finding the local duplication of the TNL gene in *Crr3* locus is a solid jump in clubroot resistance studies. By now the paper was cited 11 times. The Authors contribution information states that the work was conceived by KM, CO and WK, 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 and the generation of the Oxford Nanopore sequencing data was done by HSC. The responsibility of the Author of the PHD Thesis was collecting, processing and analyzing the phenotyping, genotyping and sequencing data and performing the association analyses as well as writing the manuscript with his supervisor (WK). The bioinformatic study helped to pinpoint the crucial part of the Brassica genome responsible for CR resistance and gave rise to deeper studies of the region, locus and gene, which is shown in two next chapters of the PhD Thesis. The marker of 577 bp was characteristic for DH lines resistant to clubroot. Is this marker available for the farmers and checked for its robustness against different pathotypes of *P. brassicae*, present in Poland, Europe or worldwide? This aspect of work has a high applicatory value, provided the marker is robust. The susceptibility of cv. Tosca to most common pathotypes of *P. brassicae* found in Canada further support the necessity to convey works with well-defined pathotypes.

Chapter "Functional study of the duplicated TNL gene"

This chapter of the Thesis is the continuation of the previous investigation to elucidate the function of the duplicated gene. Interestingly the chapter has its own short introduction, which is more of the explanation of the rationale behind this study than a description of the role of gene duplication. The introduction explains the Author circumvented the need of gene cloning by synthetic CDS. The methodology was described in detail, including the construct of the cassettes and their cloning into vectors, all primers and markers used (kanamycin and glufosinate resistance) as well as the way of selection of glufosinate resistant transformants (using Basta-based solution, BASF), expression validation and phenotype assessment. Why the disease symptom evaluation was changed from 0-4 to 0-3 scale? The experiment was successful and gave 95%-true positive transformants with transgenes present. Figure 2A should be bigger to show a close up of stunted and transformed plants, I had a problem to see the stunted plants. Figure 2B is not self-explanatory and one cannot read it without the text (crT1 or mNG, mS etc. explained neither under the Figure nor in the list of abbreviations, but only in Materials and methods, page 61).

The results of this study support the duplication of TNL gene, pinpointed at *Crr3* locus at the preliminary experiment. Moreover, the Author proves the activity of both gene copies in contrast to previous studies of Hatakeyama, who found that only one gene of *CRa* and *CRb* pair confers resistance. In the Discussion the Author himself asks several questions about the role of gene duplication, the possibilities of responding to different effectors like in other dual TNL systems found in plants as well as the probable redundancy of one gene out of two paralogues. The discussion is quite unusual, reminds "thinking aloud" but it is very juicy and interesting. It shows a deep knowledge of the Candidate about the subject. It strangely contrasts with the short, slightly boring and very general introduction. The interesting part of the Thesis is also the discussion of study limitations. Again it refers to the way of plant disease phenotyping, which was used as a standard procedure (7 weeks post inoculation) whereas the glufosinate-selected plants most probably required the evaluation at a shorter time, eg. 5-6 weeks post inoculation). These questions could be solved by the work which involves rhizotrones and root phenotyping rather than phenotyping at the pre-defined time-point. As indicated by the Author, future studies should examine other pathotypes of *P. brassicae* and use each gene combination (two single and double knockouts in cv. 'Tosca' and two single and double substitutions in a susceptible cultivar). The study was a great continuation of the published work. Congratulations!

Chapter "Comparative analysis of the *Crr3*^{Tsc} homologous *loci*"

A brief comparative analysis of the locus conferring clubroot resistance was conducted. The syntenic genomic blocks in several Brassicaceae genomes were identified and their ancestry was traced back in relation to the Brassica lineage polyploidization events together with the examination of gene presence or absence dynamics and the copy number variation. The genomes of nine Brassicaceae species were analysed, which included 29 genomes, primarily 13 genomes of *B. napus*, 7 genomes of *B. rapa* and 3 genomes of *B. oleracea* as well as *B. nigra, Raphanus sativus, Thlaspi arvense, Arabis alpina, Arabidopsis lyrata* and *A. thaliana* (one genome per species). The binary presence-absence matrix was a foundation of further analyses. The copy number variation was inspected visually in A03 and C03 chromosomes, checking for TLR duplication. A phylogenetic tree was constructed based on annotated protein sequences. In contrast to the previous chapter Figure 2 was explained in more details, still missed the explanation of L1, L3, L5 (lineages). The Author found that *crT* gene was present in L3 only. The presence of the insertion was universal on A03 chromosome and it was a single copy, with the exception of *B. rapa* ssp. *chinensis* (Pak Choi) cultivar 'PC-fu'. The Author found similar structure and sequence identity both for cv. 'Tosca' and for 'cv. 'PC-fu'. In-paralogous sequence within the genomes showed a higher similarity than the orthologous sequence between them. The study shows that the TNL gene is a new element within the locus and its duplication is very uncommon. The Author found only two examples and a high sequence similarity between them, suggesting origination from the same event rather than independent duplications. The insertion seemed to be more linked to A clade, however the Author mentioned that alternatively the C03 insertion could have been secondarily lost, which was also a theoretical possibility. Some parts of the Discussion were speculative, such as the fragment on potential mechanism of insertion attributed to transposon. The proximity of the insertion to sugar transport protein *STP6* may have played an important role in inducing of *crT* expression by the pathogen. The Author reminded the *STP6*-like fragments are in fact transposable elements belonging to Helitron family. In spite of numerous studies the role of gene duplication in clubroot resistance remained unknown.

Conclusions

The Author stated that among the genomes studied thus far, only *B. napus* 'Tosca' is the one with resistance gene duplication, while only one copy was present in a susceptible line. He also found a similar duplication event in *B. rapa* ssp. *chinensis* (Pak Choi) cultivar 'PC-fu'. The candidate resistance gene *crT* showed the classical structure of R genes from TNL family. In cv. 'Tosca' both copies were highly polymorphic in the pattern recognizing LRR domain as compared to the cultivar susceptible to clubroot. Transfer of the *crT* gene to *Arabidopsis* conferred the resistance to *P. brassicae*. High evolutionary dynamics of the $Crr3^{TSC}$ locus was shown, particularly in respect to *crT* gene. The conclusions were truly based on the results of experiments as well as skillful computation. Well done!

Remarks and overall evaluation of the Thesis

The PhD Thesis of Piotr Kopeć explain the reasons of the resistance conferred by cv. 'Tosca' and present the uniqueness of this cultivar in respect to clubroot. The doctoral monograph elaborated by Piotr Kopeć is an important contribution to the knowledge on resistance genes to *P. brassicae*, a devastating oilseed rape pathogen worldwide. Piotr Kopeć elucidated the difficult and complex issue of the cv. 'Tosca' resistance to infection with the protozoan *P. brassicae* from the Rhizaria infrakingdom. The study contains basic research as well as applicatory aspect for resistance breeding. The work used modern molecular methods, numerous bioinformatic tools were implemented. Piotr Kopeć proved his comprehensive knowledge and the ability to use the appropriate research tools for different purposes of his study. The experiments were performed and described in a logical and clear way.

The Doctoral Thesis on 'Genetic determinants of rapeseed (*Brassica napus* L.) resistance against *Plasmodiophora brassicae* Wor. infection' prepared by Piotr Kopeć, MSc. meets the requirements of the Act on academic degrees and titles. With full confidence I recommend Piotr Kopeć for the next steps of the doctoral procedure, leading to the degree of philosophy doctor in natural sciences in the discipline of biology.

Due to the complexity of the study, advancement of the research techniques used, the importance of studied problem, high scientific value of obtained results as well as clarity of writing and illustrating the results, I appeal to the Disciplinary Council and the Faculty Council of *Collegium Biologicum* of Adam Mickiewicz University to award Mr. Piotr Kopeć with a special prize or distinction.

Małgorzata Jędryczka