Institute of Biochemistry and Biophysics
Polish Academy of Sciences

Pawińskiego 5a; 02-106 Warszawa; Tel.: +48 22 / 592 21 45; Fax: +48 22 / 592 21 90; e-mail: secretariate@ibb.waw.pl; http://www.ibb.waw.pl

The review of the PhD dissertation by Juan Camilo Ochoa Cabezas *"Elucidation of genetic factors determining resistance through genome wide association study, transcriptome profiling and functional genetics in Arabidopsis natural accessions"*

The submitted PhD dissertation was carried out under the supervision of dr. hab. Robert Malinowski, associate professor at the Institute of Plant Genetics of the Polish Academy of Sciences. Dr. William Truman was the auxiliary supervisor of this thesis. The research in the laboratory led by dr. hab. Malinowski focuses on developmental plasticity of plants exposed to various environmental cues. A special emphasis is placed on the plant response to *Plasmodiophora brasicae*, a protist, which is a causal agent of clubroot. This topic is not only of a great economical importance but it has been providing a novel insight into plant physiology. Thus it is of a broad interest to the scientific community, and the previous findings of the group have been published in the very prestigious journals. The aim of the studies conducted by Juan Camilo Ochoa Cabezas was identification of genetic factors that determine resistance of *Arabidopsis thaliana* to *P. brasicae* by analyzing natural variation in this trait.

The dissertation has a classic structure, is written in English and prepared very carefully. It comprises all the standard sections, that is *Abstracts* (both in English and Polish), *Introduction*, *Objective & Hypotheses*, *Materials & Methods*, *Results*, *Discussion*, *Conclusions & References*. The content of the text is densely illustrated with 50 figures and 10 tables. The *Introduction* provides a general background and all pieces of information necessary to understand the experiments. The only shortcoming of this part, in my opinion, is that it makes an impression of being composed of unrelated chapters. Next, the PhD candidate clearly formulated the objectives and hypotheses of his work and described in detail the methods used in the study. This latter section, however, in contrast to the rest of the dissertation, contains relatively many small errors like superfluous spaces, wrong prepositions (for instance, 'grown **in** the conditions' instead of '**under**') etc. More importantly some essential details are lacking, for example in which ratios/concentrations were prepared the precipitation mixtures in the following procedures? "The DNA was purified with 5 M potassium acetate followed by precipitation ***with a mixture of isopropanol and 3 M sodium acetate***", "followed by precipitation ***with a mixture of absolute ethanol and sodium acetate".*** It is also not always defined whether volume per volume or weight per volume percentage was used.

The *Results* section is divided into two separate parts. The objective of the first part was to pinpoint genetic factors that determine resistance to P1B pathotype of *P. brasicae.* To do this the author applied state-of-the art techniques including GWAS, RNAseq and CRISPR/Cas9 combined with the classic physiological and imagining methods. GWAS analysis performed on natural accessions of Arabidopsis revealed two candidate loci putatively conferring resistance to the clubroot. The first region is located between *At1g32020* and *At1g32100* corresponds to the previously mapped *RPB1* locus. It shows high variability and a complex structure. Comparative analysis of this region of 142 Arabidopsis accessions, performed by a PhD student, allowed to propose a hypothesis that *RPB1* gene contributes to the clubroot resistance. *RPB1* encodes a highly conserved transmembrane protein that does not show any homology to known proteins. Surprisingly, the corresponding gene is present also in the susceptible accessions, and while the coding sequence is highly conserved there is a high variability in the upstream sequences and therefore the author ascribed the resistance phenotype to the promoter region. The observation that the promoter region displays substantial variability raises the question which of its elements are potentially involved in the resistance response, and how it could be verified? Transient expression of RPB1 in *Nicotiana tabacum* evoked extensive tissue necrotization. In turn, deletion of a part of the coding sequence of *RPB1* using CRISPR/Cas9 technique compromised the resistance to *P. brasicae*, and abolished upregulation of the defense-related markers, indicating a crucial role of RPB1 in the establishment of resistance. In contrast, the *RPB1* transcript, in the mutant line, still underwent upregulation upon inoculation with *P. brasicae*. Collectively, the major achievement of this work was to identify *RPB1* gene as an essential element of Arabidopsis response to P1B pathotype of *P. brasicae.*

The second region identified in the GWAS analysis was precisely mapped to the sequence encoding TNL-type resistance protein RAC1. Unexpectedly, elimination of *Rac1* did not affect resistance but I would still not exclude that the encoded TNL contributes to the resistance response. For instance it could play a regulatory role. This hit is also interesting in the context that *rpb1* mutant is still able to perceive pathogen presence that could be inferred from the upregulation of *RPB1* transcript in the mutant line. Here again, I would recommend to analyze also gain-of-function phenotypes. I cannot also agree with the statement concerning the identified SNP that since 'the substitution is present between the LRR domain and the C-terminal end [...] it has no obvious impact on the protein function'. Especially, that within this region in Roq1 receptor, a TNL from *Nicotiana benthamiana*, a C-JID domain (C-terminal jelly-roll/Ig-like domain) has been recently identified (Martin et al, 2020 *Science*). This domain attains a horseshoe-shape and is directly involved in binding of the recognized effector XopQ from *Xanthomonas* spp.

Specific comment to this part

It will be much easier for the reader to immerse in the story, if this section would start with a short, 2-3 sentences, introduction.

In the Section 4.6 referring to Fig. 30, the author has written that a transient expression of RPB1 causes a strong hypersensitive response. I do agree that it induces a strong necrotization/tissue collapse but based solely on macroscopic observation one cannot exclude other possibilities including that RPB1 overexpression is toxic. Therefore I would avoid such definite statements and would recommend to perform some assays in order to get more insight into the nature of this phenomenon.

The second part of the *Results* focuses on elucidation of the discrepancy between relatively high level of pathogen DNA and low disease index displayed by Pro-0 accession in response to *P. brassicae*. To address this question the author compared hypocotyl structure of Pro-0 with the susceptible accession Col-0 and performed transcriptional profiling of both lines under control conditions and following *P. brassicae* inoculation. He found anatomical differences in vessels and correlated this phenomenon with differentially expressed XPV, a transcription factor (TF) from the NAC family.

Fascinatingly, comparative transcriptomic analysis of two *Brassica napus* genotypes, that differed in resistance to *P. brassicae* revealed also a NAC domain TF (NAC-containing protein 83, NAC083), that binds to the VASCULAR-RELATED NAC-DOMAIN7 negatively regulating xylem formation (Galindo-Gonzalez et al, 2020 *Frontiers in Plant Science*). This paper has not been discussed in the dissertation. It would be interesting to compare also other differentially expressed transcripts identified in Arabidopsis vs. those of rape seed plants. I would like that during public presentation of the PhD thesis these issues will be discussed.

In my opinion, the second part of the results, although very interesting, is preliminary and requires further elaboration. My main concern is that ecotypes Col-0 and Pro-0 apparently differ in their growth rates, and that is why I would recommend to perform infection experiments on plants of the same developmental stage (routinely plants with 6 or 8 fully developed leaves are used in pathoassays). I can expect that the changes observed in the transcription profiles reflect also differences in the development, for example I assume that genes associated with flowering are upregulated in the inoculated Pro-0 plants that were bolting or flowering during samples collection (Fig. 38). Due to these growth differences I have also some reservation about the conclusions inferred from this part of the studies. Namely, I would not exclude that the xylem volume changes during the plant development. Similarly, it is difficult to discern without doubt, whether observed changes in expression levels of host genes related to cell cycle or cell growth, are indeed indicative of the dynamics of clubroot gall development rather than accession specific growth rates. I have also a problem to understand why in uninfected hypocotyls (Fig. 39) the diameters of cross sections derived from Pro-0 accession are significantly bigger than those of Col-0, which actually

corresponds to plant phenotype (Fig. 38), whereas upon inoculation (Fig. 41) the diameters of Col-0 sections are bigger, and this is not at all reflected in the plant size.

The experiments could be corroborated by analysis of Arabidopsis lines manipulated for expression of XVP. Do they display abnormalities in xylem formation compared to the control plants? Additionally, it is always a good strategy to analyze loss-of-function and gain-of-function phenotypes to confirm involvement of a candidate gene in the process studied. The same I would recommend for RPB1 studies.

Finally, one could consider an alternative scenario. The wild-type Pro-0 accession possesses the autoimmune allele of *ACD6* (accelerated cell death 6), that strongly enhances resistance to unrelated pathogens (Todesco et al, 2010, *Nature*). However, its activity in Pro-0 is suppressed by an allelic variant of *SNC1* and thus this accession does not spontaneously form necroses (Zhu et a, 2019, *PLOS Genetics*). Since resistance to clubroot relies upon hypersensitive response one can hypothesize that the phenotype observed in response to *P. brasicae,* that is a low disease index and abundant pathogen DNA, is related rather to a negative regulation of necroses' development - the onset and/or spread- that would exert variant of *SNC1* in this genetic background. I wonder what an experimental approach would the PhD candidate suggest to test this hypothesis?

Especially valuable aspect of this thesis is the way of data interpretation. The PhD candidate

is aware of some limitations of his studies and suggests further experiments that could complete the story.

Minor comments:

The genus name of *Plasmodiophora brasicae* is not written in full in the summary, where it appears for the first time. Moreover, I would recommend to write the full name at its first use in each section and figure, as it is practiced in the manuscripts.

The figures should be self-explanatory whereas some captions are really cryptic, lacking overview of the experiment, containing unknown abbreviations etc. (for example see Fig. 17).

Fig. 40 it is not clear from the legend, whether control plants or inoculated with the pathogen were assayed.

'data' although in standard English used as a mass noun, in scientific language is still treated as a plural (lexico.com)

p.55 it reads Ludwig-**m**üller should read Ludwig-**M**üller

The abbreviation of Public Library of Science has changed from PL**o**S to PL**O**S

In the section *Summary and discussion* some sentences contain imprecise descriptors for instance 'healthier rosettes' or 'less swollen hypocotyls'.

There are also some grammar errors in the text, below a few examples

p. 25 it reads 'zoosporangium undergo' should read undergo**es**

p. 29 it reads 'each species possess' should read each species possess**es**

p. 74 it reads '13 ORF' should read 13 ORF**s**

p. 79 it reads 'are completely susceptibility' should read are completely susceptible

p. 90 it reads 'there were an abundance' should read there was an abundance

In general the thesis is written in a proper and correct English, the text is carefully prepared and the occurring errors do not affect the reception.

In my opinion the results presented in the dissertation are original and they contribute to our better understanding of the mechanisms underlying the defense response of Arabidopsis to *P. brasicae*. Moreover, these findings open new exciting avenues and form a solid base for further studies. Thus I am fully convinced, that the presented dissertation fulfils all the statutory requirements for doctoral dissertations, therefore I recommend to the Institute of Plant Genetics PAS Council that Mr Juan Camilo Ochoa Cabezas is admitted to further stages of a doctorate defense.

Magdalena Krzymowska, Ph. D.

The Head of Plant & Microbial Lab.

at the Institute of Biochemistry & Biophysics
Polish Academy of Sciences

Warsaw, January 20, 2022