Title

"Elucidation of genetic factors determining resistance or susceptibility to clubroot disease through genome wide association study, transcriptome profiling and functional genetics in Arabidopsis natural accessions".

Abstract

Clubroot disease, caused by the obligate biotroph *P. brassicae* has become one of the most important limitations to *Brassica* crop production; the use of resistant cultivars is the most advantageous way to manage this problem. As the prevalence of pathogen races that can overcome the resistance of currently available cultivars is expanding, the elucidation and characterization of durable mechanism for immunity is a priority.

In our study a collection of 142 Arabidopsis accessions was used to identify genetic factors responsible for the resistance or susceptibility to a Polish *P. brassicae* pathotype predominant countrywide. Through a genome wide association study (GWAS) two loci associated with resistance with a high degree of significance were identified, one of these regions contained *RPB1* (*Resistance to Plasmodiophora brassicae 1*) and *RPB1* homologs in the resistant accessions Uod-1 and Est-1; the second locus was in the coding sequence of *RAC1 (Resistance to Albugo candida 1)*, that codes for a TIR-NB-LRR protein. Through the generation of loss of function mutants in resistant accessions created with the CRISPR/Cas9 (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-Associated protein 9) technique, a fundamental role for *RPB1* in resistance to clubroot disease was demonstrated, as the *rpb1* lines presented symptoms and pathogen DNA quantification comparable to the highly susceptible accession Col-0. *RPB1* codes for a small, putative transmembrane protein (148 aa, 16.0 kDa) with no known function or homology to characterized genes. The coding sequence for *RPB1* is present in several of the *P. brassicae* susceptible accessions and exhibits a high degree of conservation among the accessions that contain it, both resistant and susceptible. However, there is substantial variation in the sequence of the upstream promoter region, and this may explain the resistance phenotype. *RPB1* is strongly upregulated in the resistant accessions following inoculation with *P. brassicae,* and it is required for the induction of defense genes such as *PR5* (*Pathogenesis Related 5*) and *CYP71A13 (Cytochrome P450 Monooxygenase 71A13)*, moreover transient expression in *N. tabacum* leaves triggers a hypersensitive reaction, which supports the hypothesis that RPB1 is a positive regulator of defense responses. The

rpb1 deletion lines were still able to respond to *P. brassicae* infection by activating transcription of the non-functional, truncated *RPB1* gene but failed to induce other defense responses, indicating that RPB1 may not be involved in the pathogen recognition but is, however, crucial for immunity.

In the screen of Arabidopsis accessions, one line, Pro-0, was prominent because it displays comparatively less severe clubroot disease symptoms but has abnormally high pathogen DNA levels. We followed up on this finding and characterized the differences observed between the susceptible Arabidopsis accessions Col-0 and Pro-0 through histological observation and transcriptomic analysis. In the hypocotyls of healthy plants, it was observed that Pro-0 contains a higher proportion of xylem relative to the hypocotyl area, and in the sections of inoculated plants the xylem anatomy appeared less disrupted compared with Col-0. With comparative transcriptomic analysis of the response to *P. brassicae* infection, in Pro-0 we observed signs of delayed progression in the pathogen driven reprogramming of host developmental processes compared with Col-0. In the dynamics of clubroot gall development, there is a proliferative phase where host cell division is stimulated, followed by an expansive phase where hypertrophy of colonized cells occurs. Based on transcriptional changes in genes regulating cell cycle progress, cell growth and vascular patterning we observed that, at the time point profiled (19 dpi), Pro-0 galls retained the transcriptome signature of the proliferative state, while Col-0 had already entered the expansive phase. Based on these results, we hypothesize that variations in host growth and development, particularly vascular development may have a strong influence on clubroot disease progression. Further understanding of this phenomenon will require detailed dissection of the role of key host genes involved in regulation of host developmental responses and how they are targeted by the pathogen; characterizing these events at the molecular level could provide strategies for developing crops that are less severely affected by *P. brassicae*.

In summary, by exploiting the natural diversity in Arabidopsis we were able to identify factors essential for resistance to clubroot disease in Arabidopsis and propose possible mechanisms explaining differences in disease progression in susceptible accessions.