Dissecting natural variations in quantitative resistances to parasites by genetical genomics

Pjotr Prins, Ritsert Jansen, Jaap Bakker and Geert Smant

Introduction

Plants are threatened by numerous pathogens and have evolved highly sophisticated mechanisms to defend themselves. The development of disease is governed by various biotic and abiotic factors, like temperature, nutrients, pH and soil type. Many research programs are focussed on understanding qualitative resistances based on gene-for-gene interactions, which mostly result in either completely resistant or highly susceptible plants and which are hardly influenced by environmental factors. The goal of this project is to dissect quantitative variations in host plant resistances by studying host-parasite interactions by genetical genomics on genome-wide expression patterns of host and parasite. In this project we will focus on the interaction between *Arabidopsis thaliana* and the root knot nematode *Meloidogyne hapla*.

Approach

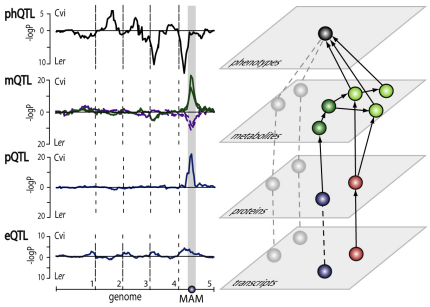

Fig. 1: xQTL analyses (from Jansen et al. 2009)

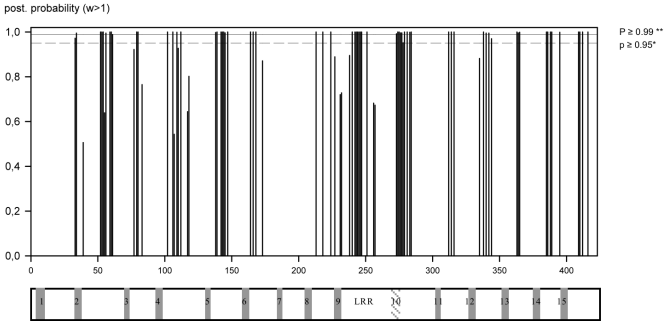

Fig. 2: Positively selected sites in the LRR domains of the Gpa2/Rx resistance gene homologues

Dissecting genome-genome interactions by xQTL

With full genome information available of both the host and the parasite, many types of high-throughput techniques can be integrated. The mapping of quantitative trait loci (QTL) allows using the latest high-throughput technologies for exploring gene regulation and gene networks. Each technology allows mapping of certain measurements, or *traits*, against a genomic location. Figure 1 shows how different technologies confirm each other. Gene expression analysis (eQTL) is compared against metabolite (mQTL), protein (pQTL) and traditional phenotypes (phQTL). The combined techniques have been coined *xQTL* analysis. The goal of this project is to combine eQTL, mQTL and phQTL to identify inter-genomic regulatory networks.

Identification of interacting genes by high-throughput detection of positive selection

We are currently writing and testing software for assessing and visualizing genomic locations that are under positive (diversifying) selection. Current software for the detection of positive selection is too slow to parse complete genomes. The idea is to analyze the complete genomes of Arabidopsis and *Meloidogyne hapla* to identify genes that are positively selected. It is noted that positive selection is often a hallmark of genes that are involved in disease resistance in the host and virulence in the parasite (see Fig. 2). This will be accomplished by optimizing the software program PAML (Phylogenetic Analysis by Maximum Likelihood).

Deep sequencing versus microarrays

Recent developments in sequencing technology (e.g. Solexa, SoLiD, and 454) are revolutionizing sequence analysis. Complete genome sequencing is now possible at relatively low costs. This has led us to investigate the possibilities of using, so-called, deep-sequencing for gene expression analysis - as a replacement for microarrays. With deep sequencing cDNAs are directly sequenced and the number of transcripts is counted. At present we are investigating i) the 'depth' of Solexa sequencing of cDNAs, ii) the stochastic variation for lowly expressed genes, and iii) the noise generated by sequencer misreads (1:1000). We have done a pilot run on *C. elegans* with deep sequencing using 7x coverage. The results are encouraging as we confirmed effects seen in earlier microarray experiments. However, to acquire sufficient sequence information for eQTL analysis, we may need a few times higher coverage as we have to account for combing the transcriptomes from two species in one sample.

Groningen Bioinformatics Centre, University of Groningen, and Laboratory of Nematology, Wageningen University