DanR:

My main concern was with the clarity of how you set up the study and compare it to other works in the intro. This study is clearly about the polygenic basis of the variation in outcomes of interactions between host and pathogen genotype. However the words “polygenic”, “genetic”, and “interaction” all have both statistical and mechanistic meanings which are really quite different and I don’t think you’re being consistent in how you use them. Same with “large effect” and “SNP vs QTL”. I think that being more careful to distinguish the mechanistic meaning (how genes, etc function in pathways, how plants detect and respond to the pathogen) from the statistical meaning (correlation between variation in genotype and variation in phenotype) will help set up the results better.

In the results, it seems like one of the main results is the lack of overlap between the hotspots with respect to B. cinerea genes and Arabidopsis genes. This is surprising (interesting) because you’d expect if something affects the host-pathogen interaction (mechanistic sense), then it would have broad-scale affects on the transcriptomes of both species. From a statistical perspective, I do worry about false negatives here, though. You don’t have a super high-power experiment (really pretty low power for GWAS), and therefore only expect to detect a smallish percent of the true associations. Therefore, if you find a hotspot in one species, you don’t have a very high probability of also finding it in the other species, even if it is actually important in the 2nd species. For the hotspots that you find in one species, you can search specifically for effects of these loci in the genes of the other loci (ignoring the rest of the genome). This effectively lowers the significance threshold for these loci in the second species. This is reasonable statistically for the question of whether the candidate locus also has an effect on the 2nd species.

Another main result is that you can find examples with these pathways of structural variation in cis that is important and missed by the GWAS. For the botcinic acid network in particular, it could be simply that this deleted pathway is at <20% frequency (from Fig 3). So the reason that you don’t pick it up is not because it’s structural, it’s because it is below your MAF threshold. If you reduced your MAF threshold to 10%, you might find a genotyped SNP on this haplotype outside of the deletion that does have a strong cis association. I think that structural variation is very interesting, and its neat that you find some with such clear impacts. But I just would be hesitant to throw out GWAS because of this here. GWAS can find loci that are caused by structural variants as long as there are markers that are in LD with the structural variation (which is really the same as how GWAS finds any other causal variation).

DanK:

Sounds like in the sections on structural variation that we should emphasize that they don’t rule out GWA just that we need bigger samples and better sequencing to help boost the power. It sounds like we didn’t go far enough to reflect this as Dan read it as a GWA bad section.

Maybe check on the wording, we might want to diminish the polygenic aspect in some sections to reflect the power issues.

DanR – we were struggling as the human literature calls a SNP linked to a transcript as a eQTL. It should be an eSNP but the existing literature seems to have settled on what both of us consider to be incorrect. Any thoughts?

DanR:

I’m fine with eQTL. I don’t like calling the SNPs in your GWAS “candidates for causal polymorphisms”, unless you can prove that every genetic variant is actually in your GWAS.

“polygenic” is fine from your results. It’s the conclusion that the hotspots do not overlap among species that I’d be cautious about.

DanK:

Sorry, I must have missed a few causal polymorphisms.

We’ll temper that conclusion even more about overlap. We have some a priori networks that link to these hotspots which I can use for a priori type test as you were suggesting. That might help us to see how the variance is propogating between the species. That will probably have to be an add on after Nicole is done and I can get it done.