1. It is known that wild tomato species can display distinct phenological development under controlled conditions with artificial light. It is unclear to what extent the phenological development of the S. pimpinellifolium genotypes was similar to cultivated tomato genotypes. Did these plants have similar growth rate, plant height, leaf sizes and shapes, internode distances, flowering times? Phenological differences may influence the outcome of fungal infection experiments in a quantitative manner that is difficult to predict and analyse.   
2. It was published that the susceptibility of tomato leaves to B. cinerea varies quantitatively with plant age and leaf position on the stem. The methods section in the manuscript describes “we selected 5 leaves per plant (expanded leaves from second true leaf or older), and 2 leaflet pairs per leaf”. I presume that the authors meant to say “from second true leaf and younger” as the lowest leaves are the first ones developed and therefore physiologically the oldest. From the description of the methods it is unclear to me to what extent the selection of the leaves has influenced the outcome of disease development.   
3. The inoculation was done with droplets of fungal spores in 50% organic grape juice. From Figure 1 it is obvious that the inoculum droplet was in many cases placed on or just next to the central leaf vein. This is, in my perception, an unfortunate choice as it results in lesions that are irregular in shape, as is evident from the image. The inoculation on the surface of vascular bundles initially provides a more difficult surface penetration, but once the fungus has penetrated the vasculature it will more readily spread phytotoxic metabolites and hydrolytic enzymes to distant host tissues, where they can exert an effect on host cells. At the same time, the vasculature does not impose any barrier to fungal growth and dispersal to neighbouring tissues. The spread of fungal infection thus largely follows the shape of the vasculature which is obvious in several leaflets displayed in Figure 1. Performing inoculations more distant from the central vein would have resulted in more  
circular lesions and would have offered a more consistent and reliable measure of lesion surface. In my judgement, such a simple change of inoculation strategy would have given more reliable quantitative measures of plant susceptibility and fungal virulence.   
4. Another aspect of experimental variation is in the inoculum density used. The methods section describes that the authors inoculated leaflets with a single drop of inoculum containing 40 spores of a B. cinerea isolate. I very much appreciate the relevance of using a low dosage when aiming to quantify plant susceptibility, in order to prevent overkill. A dosage of 40 spores is at the very low end of the spectrum, and it is therefore crucial to ensure an accurate counting and dilution of spores. I can envisage that applying 30 spores instead of 40 would have an impact on the speed of disease development. Also spore quality and viability should be comparable between isolates. It may sound trivial, but considering the use of 97 fungal isolates in a single experiment, it is truly important, yet difficult to standardize. When inoculating this many isolates in a single experiment, the logistics and timing are complex. As soon as spores are diluted in the grape juice medium, they  
germinate within 3-4 hours. It is therefore crucial that they are inoculated onto the leaves before they germinate. With 12 plant genotypes and 97 fungal isolates, this is a challenge. To what extent have the authors ensured that the eventual lesion sizes were not influenced by slight variations in spore concentrations, the state of spore germination and other variables imposed by experimental procedures?  
5. Considering the concerns described above, it is difficult for me to judge whether having six replicates of lesion sizes at a single time point for every individual plant-fungus interaction is sufficient for valid conclusions. The analyses may be statistically correct, however, I cannot judge whether the outcomes truly reflect genetically determined biological differences with a meaningful impact on plant-fungus interactions.   
6. Regarding GWA of the fungal genome for variation in virulence among the fungal isolates, it is quite surprising that the authors mapped the fungal sequence reads to an outdated B. cinerea genome. The methods section reports that read data were mapped to an improved draft genome of strain T4 published in 2012. First of all, strain T4 is a very poorly pathogenic strain. Secondly this genome was far from complete and its annotation was poor. Thirdly this genome was not deposited in NCBI, and only accessible through the Broad Institute Fungal Genome website until that access was discontinued in 2016. It is therefore no longer possible for readers to retrieve the B. cinerea genes with gene IDs “BcT4\_xxxx” (as used in Table S3a). A gapless genome of B.cinerea strain B05.10 (with a community-curated annotation) is available since June 2016 at the EnsemblFungi platform of EBI. The publication describing this gapless genome was online in the spring of 2016. The authors are aware of  
this, since they do use gene IDs from this gapless genome in Table S3a as well. However, it seems that the initial GWA mapping was done on the imperfect genome of strain T4 and the outcome of the analysis was subsequently converted to orthologs in B05.10. The GWA mapping of SNP data on the gapless B05.10 genome might have yielded a more reliable and comprehensive outcome.   
7. In addition, the authors describe in the methods section that the “SNPs were annotated using SNPdat ….. from the T4 gene models for genomic DNA by linking the SNP to genes within a 2kbp window”. It should be noted that intergenic regions in many fungi, including B. cinerea, are often short, at times shorter than 0.5 kbp. Choosing windows of 2 kb surrounding the start and stop of a particular gene may wrongly assign a SNP to the neighbouring gene(s). Examining Table S3a suggests that this appears to have happened: there are several physically clustered genes in the table (BcT4\_6000, 6001, 6002, 6003; BcT4\_8803, 8805). I would encourage the authors to repeat the SNP mapping on the new, gapless, publicly available B. cinerea genome and use a smaller window, or possibly limit the analysis to the coding region only.   
  
  
  
Reviewer #2 (Remarks to the Author):  
  
Soltis et al. presented a study on the interactions of tomato and Botrytis genomes in terms of the crop domestication and lesion size. Although the overall aim is interesting, the series of analyses performed do not support their main conclusions, e.g. the interaction effect of the two genomes is not properly justified, nor the polygenicity argument due to the biased procedure of GWAS.   
  
  
Major comments:  
  
1. First of all, regarding the multiple linear regression model, the statement that the final model explained 60% of the total variance of lesion size is very misleading. All the factors that include experiment or block are completely residual variation, not relevant to the genetic or domestication factors. Obviously, only a fraction of 20% variance is explained with relevance to the aim of the study. 60% is a meaningless number.  
  
2. Is the small domestication effect caused by some sort of sampling bias as the number of tomato lines is so few? No data is provided to show the kinship across the tomato lines, making it impossible to assess whether the tomato lines were selected properly for the purpose of investigating domestication.  
  
3. Regarding Figure 3, why talk so much about variance in lesion size when it is not statistically significant? The slightly wider range of domesticated plants to me is just chance. I find this part of the results redundant. If the conclusion is just to show no phenotypic bottleneck, no need to have this whole section of text.  
  
4. Does 'visual analysis' of Figure 1c-h really show an interaction between the two genomes? This is not a scientific statement - without further justification, the variation we see in the figure could all be random (non-genetic).  
  
5. Regarding the interaction effect of the two genomes, the test of a factor with 940 degrees of freedom is useless. I don't understand why random effect models were not used to fit and test these factors with high df.   
  
6. Following the last comment, Wilcoxon test with some sort of FDR calculation was used instead to assess the interaction between two genomes. 1) I don't see Table 2! 2) What does it mean by FDR-corrected p-value? Unclear. How was the correction done? 3) I feel it's some simple FDR correction, so was the population structure of the fungus considered? Without the independence of the individuals, the Wilcoxon test and the FDR processes are invalid. I have no clue about the null distribution of the Wilcoxon statistic under this circumstance.  
  
7. Regarding the GWA analysis, to me, there are more severe issues. I'm OK with using the whole-genome marker-specific shrinkage regression method to fit the data. The question is whether this really supports the conclusions thereafter. This is a discovery problem, so as in all GWAS, population structure is an essential confounder that should be considered and corrected, but it is not done here. All the signals revealed by permutation could be confounded with population structure, namely, a structure was mapped instead of causal loci. I'm surprised that a standard mixed model GWAS was not used for this discovery analysis. Without a proper discovery phase or validation study, the subsequent fine-mapping analysis and candidate gene analysis are meaningless and cherry-picking. The polygenicity conclusion is thus not properly justified either.   
  
8. Figure 5a, 'the level of overlap exceeds the expected overlap' - is the difference significant? There must be a way to provide p-value.  
  
9. Regarding the GWA about 'domestication shift', similar major issue applies as above in point 7. Besides, why the genetic effects in domesticated plants are larger than those in wild?  
  
  
Minor but not necessarily small points:  
  
10. Regarding Table 1, it's not a nice table for a paper, too much redundant information, e.g. SS, F-value.  
  
11. Throughout the manuscript, I don't think 'p < XX' follows the requirement of statistical reports in Nature journals - report 'p = ??' instead.  
  
12. The wild v.s. domesticated effect is very small. I simply cannot see the effect from Figure 2 which to me is a useless figure.  
  
  
  
Reviewer #3 (Remarks to the Author):  
  
The manuscript "Crop domestication and pathogen virulence: Interactions of tomato and Botrytis genetic diversity" is a welcome inquiry into quantitative interactions between plant hosts and broad-host range pathogens. Much of our understanding of resistance/virulence interactions in plant pathogen interactions, and especially where there are comparisons between domesticated and wild pathosystems, comes from studies of host-specialized biotrophs and the large-effect loci involved in resistance to them. Biotrophs are of course important pathogens and are tractable for elegant studies in evolutionary ecology of plant-pathogen interactions, but they represent a small subset of the diversity of important plant diseases. This manuscript presents an experimental study of quantitative variation in disease response associated with inoculation of a large number (97) of genotypes of the important generalist fungal pathogen Botrytis cinerea onto a suite of a dozen domesticated and wild  
genotypes of Solanum. Combined with genome-wide association mapping, the authors demonstrate the strong polygenic nature of variation in resistance in the host, and the ready availability of allelic variation in the pathogen to overcome any introgression of wild-host resistance alleles into domesticated crops. This is an exceptionally clean and robust demonstration of what is expected to be a limitation to breeding for resistance to generalist pathogens - novel as far as I know for a generalist pathogen. I appreciate having a quantitative estimate of just how much of an effect domestication has on resistance to generalist pathogens, within the context of genotypic variation. Botrytis cinerea is the "poster child" generalist plant pathogen, and an ideal candidate for this endeavor. I would expect this paper to become a classic citation; I know it will directly inform my own research and will immediately become part of required readings in my classes on plant disease.   
  
I thought the manuscript was extremely well written. I enjoyed reading it, thought the framing that included both evolutionary ecology theory and issues of domestication to be strong, and the description of what was done, and what it means, to be clear.   
  
I appreciate the careful distinction in the study to differentiate roles of variation in the host, variation in the pathogen, and variation in the interactions between them -- and the 3276 inoculations needed to test this in a robust way. My curiosity leads me to want to know what happens beyond these two closely related host species -- are there patterns of disjuncture with greater phylogenetic distance? -- but I recognize that that must be a story for another paper, and appreciate that they point out the possibility of such work in the conclusions.   
  
I though figure 1 was a creative way to illustrate the various result patterns -- by reading through the figure legend in conjunction with the graphs, I got a much clearer appreciation than just from reading the text. I also like figure 3.   
  
Minor wording choice, line 283: "we identified a significant increase in the resistance of wild tomato in comparison to domesticated tomato". The resistance was GREATER in wild, but it did not INCREASE (which implies a change). It may be OK to say "observe the expected decreased resistance in domesticated tomato" because the domesticates come from wild, but they didn't really come from these particular wild genotypes, and so even here the increase/decrease wording is troubling.  
  
Another minor wording thought on the section title "Pathogen Specialization to Host Variation" This section is quite strong and interesting -- I liked the approaches to looking at subsets to address specific questions. But I do not quite think the title is appropriate to the results, because the specialization is not to variation -- but rather there is specialization to genotypes within the host variation (host x pathogen genotypic interactions). I think the idea of "specialization TO host variation" is intriguing, but that is not quite what is being addressed here.   
  
Although I just said I like this approach of subsetting the data to address specific questions, I am just a bit uncomfortable with using 2/97 isolates that have statistically significantly different patterns from others as a strong basis for analysis of effect of domestication. Since so much of the story here is about the great variation, even with FDR correction, it seems likely to find isolates that represent nearly any desired pattern. I do not necessarily doubt the interpretation, I am just leary of the use of a couple unusual strains to make central points. Perhaps a bit more attention to caveats would be helpful.   
  
I do not have a strong background in GWA mapping, and cannot comment on the robustness of the approach. But as an end user who thinks about variation, I found this compelling and useful and generally easy to follow, except for Figure 5. In Figure 5 the figure legend and text do not do an adequate job of explaining the figures -- for instance, what the meaning of the inset graphs is, and how to interpret them (line 406) for "levels of overlap exceed the expected overlap due to random chance (Figure 5a).   
  
Methods: I would appreciate a little more detail on the ancestry of the domesticated genotypes. The wild genotypes come from across the geographic range, and represent a diversity of genotypes. What is known about the origins of the domesticates? Were they all derived from a single previous domestication event, or from six separate domestications, or something in between? Just having a bit more context for these genotypes would help in thinking abut interpretation of results. 