Dear Prof Kliebenstein,   
  
Your manuscript entitled "Crop domestication and pathogen virulence: Interactions of tomato and Botrytis genetic diversity" has now been seen by 3 referees, whose comments are appended below. In the light of their advice I regret to inform you that we cannot publish your manuscript in Nature Communications.  
  
You will see that, while the reviewers find your work of interest, they raise substantive concerns on the correctness of the input data (phenotyping of the disease development, Reviewer #1) and the data analysis (the interference of population structure on, Reviewer #2), which cast doubt on the advance your findings represent over earlier work and the strength of the novel conclusions that can be drawn at this stage. Unfortunately, these reservations are sufficiently important to preclude publication of this study in Nature Communications.   
  
<Response> We have conducted the mixed model with population structure analysis on the new SNP calls as requested and this showed exactly the same architecture as found with the Genome Prediction methodology. Further, we have also clarified several sections of the materials and methods to rectify the concerns from Reviewer 1 and commented on how development and disease are actually highly integrated systems that are nearly impossible to separate. Hence the vast developing literature on growth to defense cross-talk. We hope that these corrections have helped to rectify the concerns.

We are unsure as to where the comment about novelty with regards to our previous work arises as it is not in any of the reviewer comments. We would like to note that in our previous work, all we showed was that there was quantitative variation in *B. cinerea*. A common misperception is that quantitative variation means polygenic but that is not true. It is possible to be polygenic and qualitative as well as monogenic and quantitative (Falconer and Mackay, 1996; Mackay, 2001; Mackay, 2009; Mackay, 2014). As such, we would respectfully like to argue that the previous work showing quantitative variation does not prove polygenic architecture. The only other studies that we know of with similar analysis of quantitative variation are in Zymoseptoria, which is a specialist monocot pathogen and these important studies typically show a few of large effect loci within largely domesticated hosts. We would like to respectfully note that there is not a single similar study looking at a vast collection of necrotrophic pathogen germplasm across domesticated and non-domesticated individuals of a plant species. Nor is there a study that we are aware of which shows such a potential polygenic architecture in a natural pathogen. As such, we feel that the observations about domestication are novel and new insights for the field as stated by multiple reviewers. We did not see any reviewer saying that the work was not novel.

Reviewers' comments:  
  
Reviewer #1 (Remarks to the Author):  
  
This manuscript describes quantitative aspects of disease development in the interaction between isolates of the fungus Botrytis cinerea and genotypes of cultivated and wild tomato. The authors inoculated 97 fungal isolates on leaves of 6 cultivated tomato (S. lycopersicum) genotypes and 6 wild tomato (S. pimpinellifolium) genotypes and quantified lesion sizes at 72 hours post inoculation. The dataset was analysed for the effects of domestication, plant genotype and pathogen genotype on disease development. The analyses indicate that wild tomato (S. pimpinellifolium) is overall slightly more resistant to B. cinerea than the cultivated tomato. Genome-wide association analysis suggested that virulence in the fungus is controlled by multiple genes, and yielded a list of genes where the allelic variation was linked to virulence.   
  
This manuscript contains an impressive amount of data, which have required a rigorous logistic design. The authors have made great effort to standardize the experimental protocol and randomize the experimental design to meet criteria for proper statistical analyses. The design, however, may have had repercussions for undesirable biological variation that I cannot judge based on the provided information. My considerations and concerns are listed below.

<Response> We hope that the responses below have clarified our protocol and alleviated your concerns.

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.

<Response> We have included information in the materials and methods that flowering time was well after our sampling time indicating that the plants had not made a transition to flowering. We agree with the reviewer that there is developmental variation between the plants even within a species as is seen between the domesticated tomato germplasm. This is true of every quantitative genetic system in plants and it is not possible to truly remove these differences. From and ecology and evolution view, it is not clear that these are actually separate systems. For example, it is well known in plants that the defense hormones have influence on development and the development hormones have influence on defense. Further, growth and development play a key role in the main defense mechanism of evasion within natural ecosystems. As such, from our viewpoint, we are not focused on solely identifying MAMP/PAMP pathways but instead we are querying anything and everything that influences the interaction of host and pathogen. In the natural system, any component, MAMP/PAMP or development that minimizes disease sensitivity will be selected on and potentially beneficial to a breeder. Additionally, we would like to note that if our results were predominantly driven by phenological variation between the *S. pimpinellifolium* and domesticated germplasm, then we would not see the overlap between the genotypes of these two species as seen in Figure 2.

We have included the following comments in the first part of the results wherein we describe the experimental setup to ensure that our goal for the experiment and our reasoning are clear to prevent any misunderstanding as may have occurred with the reviewer. “It should be noted that we are not focusing on MAMP or PAMP specific host/pathogen interactions with this study but that we are instead allowing the identification of any mechanism that may influence the host/pathogen interaction including metabolism, development or any other unknown component. As long as there is genetic variation affecting the trait and the trait influences the interaction of host and pathogen it will be a component of the experiment. This fits with the recently developing view that growth, development and resistance in plants are highly integrated processes that may not be as distinct as once believed (Campos et al., 2016; Ballaré and Pierik, 2017; Karasov et al., 2017; Züst and Agrawal, 2017).”.

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.

<Response>We have corrected this typo in the text to say “second true leaf and younger”. We would also like to note that in our original models as described in the materials and methods, we explicitly tested for effects of leaf position and leaflet number but that neither of these terms significantly influenced our results. Thus, these were dropped from the final model to streamline the model and to minimize any potential for over-parametrization.

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.

<Response> As the reviewer noted, the spore application spot was kept consistent near the primary vasculature given the complexity of the structure of the tomato leaf and the differences in primary and secondary vascular patterning and spacing. This was considered the best solution to ensure similar infection locations. We should note that we are not aware of evidence that the vasculature is truly defenseless or that circular lesions are “truer” measurements of the interaction. In fact, we have published results that show exactly the opposite. Specifically that there is genetic variation in the host and the pathogen that interact to control the shape of the lesion independently of the size of the lesion (Fordyce et al., 2018). Further, when conducting microscopic analysis, we realized that the growth on the vasculature is actually linked to genetic variation in the pathogen and host with regards to preferential tip growth in the pathogen (Corwin et al., 2016). To illustrate this, in Figure 1, the droplets are all placed in the same spot as best as possible and represent different isolates on different genotypes and it is quite clearly visible that only some isolates track along the vasculature. As such, we would like to respectfully disagree with the reviewer and instead suggest that lesion area and shape are both genetic aspects of the interaction that can be studied separately and we focused solely on lesion area in this manuscript.

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?

<Response> We would like to thank the reviewer for identifying a key point that we know realize that we had foolishly not discussed in our papers over the past 13 years on this protocol. We had a specific method to ensure even application exactly as the reviewer was concerned about. We have now included the following information in the materials and methods “The diluted spore suspensions were homogenized by agitation continuously during the entire process of applying the spores to all samples. This maintains the spores in the suspension and ensures even application across samples,.” The pipettors are calibrated routinely during the year to ensure that there RSE is never >5% which means that our inoculum will have a SD of +/- 2 spores. We went back over the past 13 years of papers where we utilized this protocol plus a sampling of the hundreds of papers where others have utilized the same protocol and realized that we had foolishly never included this information. We would like to thank the reviewer for identifying this essential oversight and hope that this information helps to clarify this concern. We would like to note that given the randomized aspect of the design that if the entire experiment was driven by pipetting error that there would be no significance for genotype or isolate as they were randomized. Additionally, in response to the concern of variation in germination timing, we note that *Botrytis cinerea* growth is significantly slowed under refrigeration, so cold incubation is effective to prevent germination prior to inoculation on the leaf surface. We have added this to our materials and methods “Spores in grape juice were maintained in 4°C refrigeration or over ice from the time of collection, to inhibit germination prior to inoculation.”

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.

<Response> We hope that the above responses have helped clarify our protocols and the long history of these protocols in our lab and a large number of other labs. We should also note that the entire experiment is randomized in a block design. As such, if the data was entirely technical error, then there should be no significant signature of any genetic component of the system. The technical error would be randomized across randomized interactions leading to solely noise. Any block specific technical error would be captured solely by the block term in the model. This is a standard approach for all quantitative genetics in all plants and animals across thousands of papers and as such, we are unsure what further evidence can be provided.

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.

<Response> In conjunction with the GEMMA analysis, we have utilized the B05.10 genome to re-call all the SNPs. This showed that the general architecture results stayed largely the same and the architecture of the traits is similar. It should be noted that the general virulence of the reference genome is not inherently linked to the validity of the SNP calls. Additionally, the total number of SNPs called with regards to T4 or B05.10 was not significantly different. Interestingly, the B05.10 reference genome identified more SNPs with a MAF less than 0.20 suggesting that it is somewhat of a minor genome within the species. We would like to note that when the B05.10 gapless genome was published was the first time we were alerted to this genome and this manuscript was in a full draft stage and that led us to be hesitant to conduct the entire process again. We had cross-annotated the gene codes to allow translation. To increase potential reliability, we focused any candidate causal discussion on genes that were found using both SNP sets and both algorithms suggesting that they are not susceptible to variation between these. We hope that this is sufficient.

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.

<Response> We would prefer to keep this window gap as a large number of publications have shown that GWA may find a SNP linked to the causal variation rather than the causal polymorphism itself. As such, we don’t consider the identified polymorphisms as the specific casual variants but instead indicators of a potential causal variation in the region. Additionally, given the role of operonic structure in virulence mechanisms like toxin production in virulence, we would prefer to stick to a more inclusive definition at this point. We are wary that going to an ORF only definition would create an artificial false negative error rate without inherently boosting the true positive rate. Further work to develop cross-indexing like wide-ranging transcriptomic or proteomic studies within this system are required to help filter this dataset to increase the causal assessment (Chan et al., 2011). We feel that those are well outside the realm of this manuscript.  
  
  
  
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.

<Response> We hope that the new analysis with GEMMA and population structure showing the same genetic architecture as with BigRR helps to address the revewers concern.  
  
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.  
  
<Response> We have deleted the 60% reference.

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.  
  
<Response> We have now included a new Supplemental Figure S1 showing the sampling of the accessions in regards to 400+ tomato accessions to show that they do in fact sample the broad collection of germplasm and are not a biased focal selection as suggested.

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.  
  
<Response> We are unsure which section of text is considered redundant. The wider range of domestication may be by chance but we felt it was an observation to note for potential research that may follow-up with a broader collection of germplasm.

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).  
  
<Response> We respectfully disagree with the reviewer. We state in the lines that follow this statement that there are two possibilities. One that the visual analysis is misleading and there is no interaction or two that the statistical model as run is not capable of properly testing this interaction. Then we proceed to utilized standard approaches in the ecology and evolutionary field to test these large interactions spaces for such an interaction using the Wilcoxon rank test.

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.   
  
<Response> We have run the analysis with both fixed and random effect models and they do not change the underlying variance structure and did not alter the conclusions. We are unsure if the *B. cinerea* collection is a prefect random sample of the species and thus prefer sticking with the fixed effect model.

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.  
  
<Response> Table 2 was previously included in the supplementary files as Table S2. We have corrected this and placed Table 2 into the body of the manuscript. P-value corrections of the Wilcoxon test were by the R p.adjust method “fdr” from the stats package (Benjamini and Hochberg, 1995). We have clarified this in the manuscript “The lower left corner of the chart includes BH FDR-corrected p-values, the upper right corner includes the test statistic (W) (Benjamini and Hochberg, 1995).” Each Wilcoxon test was conducted on the distribution of mean lesion sizes across all isolates on the selected host pair. We have clarified this in the table 2 caption “Wilcoxon signed-rank test comparing mean *B. cinerea* lesion area on tomato accessions. This tests for a change in the rank order of the 97 isolates between each pair of tomato accessions.”. We would like to note that when using Structure, there is a signal of population structure but upon analyzing the entirety of the data, the Structure signal only links to about 5% of the total pairwise SNP variation indicating that population structure is not likely a driving factor in the Wilcoxon rank test. We agree that if there was extensive population structure, the isolates could move in tandem as sub-populations. Given this low fraction of variance linked to population’s structure, it is likely not driving the results. We would like to respectfully submit that in every biological system including randomized F2s, there is some inter-relatedness between genotypes and as such it is not clear if any Wilcoxon rank test in the literature would be considered appropriate if absolute independence was universally required.

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.   
  
<Response> We have completely rerun all of the analyses with GEMMA to conduct a mixed model analysis with population structure correction as requested by the reviewer and included a suite of supplemental figures. These show that the exact same general conclusions about genetic architecture arise about specificity of the interaction, polygenicity and all other major conclusions. As expected GEMMA came up with different putative causal loci. However as we had noted in the citations and text of the previous manuscript, controlling for population structure in a mixed model is not a panacea to solve all problems as it introduces a major false negative error rate (Chan et al., 2011; Brachi et al., 2015; Kooke et al., 2016; Corwin and Kliebenstein, 2017). We did work to refine our analysis by querying for putative causal loci that were identified by both algorithms suggesting that they are stronger candidates than ones found by either approach. As noted, these did agree with a general function in virulence. Unfortunately, there is not a vast body of molecular characterization of genome wide virulence activities; there is not a large body of known causal loci to use for a comparison. There is also not a ready-made collection of standing mutations to order for large validation studies. We agree that these would be important for a follow-up study but given the biological system it is not readily achievable in a single manuscript that is not inherently focused on the specific causal loci.

8. Figure 5a, 'the level of overlap exceeds the expected overlap' - is the difference significant? There must be a way to provide p-value.  
  
<Response> Within the figure, we have provided the line that shows the maximal overlap expected by permutation at each overlap level. This shows that this is larger than maximally expected using 1000 permutations of the data.

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?  
  
<Response> Please see the response to point 7 about the re-analysis and the new supplemental Figures developed using the GEMMA analysis that show a similar image to that found with BigRR. The reason that the genetic effects in domesticated plants are larger is simply because virulence is on average large in domesticated genotypes so the phenotypic variation going into the model is larger on the domesticated than on the 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.  
  
<Response> We would prefer to maintain the table in the format that it is in. We have learned over time that different readers have been taught different ways of reading these tables and the minimal tables that are maximally informative for one person are not as useful for another person even though it should be relatively easy to interconvert.

11. Throughout the manuscript, I don't think 'p < XX' follows the requirement of statistical reports in Nature journals - report 'p = ??' instead.  
  
<Response> We provided the specific p values where they were meaningfully calculatable. For example, in Table 1, we report p <2e-16 because that is the lowest value that R will calculate and any more specific resolution on the p value would be largely meaningless. There is not a dramatic difference in a p of e-10 versus e-16 or e-20. We would prefer to keep it this way to not convey an incorrect impression of the precision in the statistics.

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.  
  
<Response> We agree that the wild vs domesticated effect is very small and that is the reason for both Figure 2 and Figure 3. Figure 3 provides a more direct view of the wild vs domesticated effect while Figure 2 is focused on showing the distribution of the individual plant accessions. We feel that the combination of the two figures is necessary to show that the wild vs domesticated difference is not driven by outlier accessions in either sample. Hence Figure 2 is critical to this interpretation. We have included a call out to Figure 3 at the first description of this difference.  
  
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.

<Response> Thanks for the support. We hope that the improved manuscript based on the responses to all the reviewers and editors is even stronger.  
  
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.   
  
<Response> We agree that those are exciting questions and unfortunately require multiple species to conduct. Given that these experiments take several months, we felt that combining numerous species in a single first paper was a bit much.

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.   
  
<Response> We agree that increase implies either time or a specific baseline. We have changed this to significantly greater to remove both connotations: “In our analysis, we identified a significantly greater, 18%, resistance of wild tomato in comparison to domesticated tomato across the population of *B. cinerea* isolates (Figure 2 and 3, Table 1, Figure S1)”.

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.   
  
<Response> We agree with the reviewer that we’d actually have to measure the CV of virulence across a specific host species population and assess that to get at specialization to host variation. We have changed this to “Host Genotype” to better reflect what is being discussed.

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.   
  
<Response> We have reworked this section to make it clearer that two isolates show explicit domestication sensitivity and that the broader collection shows more evidence of general sensitivity to genetic variation in the host using the Rank tests. In the discussion, we worked to make it clear that the majority of our evidence on a potential domestication interaction are predominantly built on the GWA results that were consistent across methodology. We hope that this provides the proper level of caveating.

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).   
  
<Response> We have worked to improve this figure legend to clarify it and hope that this improves the interpretation.

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.

<Response> We have worked throughout the manuscript to address these questions by pointing out that Tomato is typically a single domestication event followed by further improvement and that we attempted to sample the geographic range of the progenitor and the germplasm of the domesticate. We have also included a new supplemental figure showing the distribution of the accessions for which sequence data is available across 400+ accessions of both species. This shown in Supplemental Figure 1.

**References**

**Ballaré, C.L., and Pierik, R.** (2017). The shade‐avoidance syndrome: multiple signals and ecological consequences. Plant, cell & environment **40,** 2530-2543.

**Benjamini, Y., and Hochberg, Y.** (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological)**,** 289-300.

**Brachi, B., Meyer, C.G., Villoutreix, R., Platt, A., Morton, T.C., Roux, F., and Bergelson, J.** (2015). Coselected genes determine adaptive variation in herbivore resistance throughout the native range of Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **112,** 4032-4037.

**Campos, M.L., Yoshida, Y., Major, I.T., de Oliveira Ferreira, D., Weraduwage, S.M., Froehlich, J.E., Johnson, B.F., Kramer, D.M., Jander, G., and Sharkey, T.D.** (2016). Rewiring of jasmonate and phytochrome B signalling uncouples plant growth-defense tradeoffs. Nature communications **7,** 12570.

**Chan, E.K., Rowe, H.C., Corwin, J.A., Joseph, B., and Kliebenstein, D.J.** (2011). Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. PLoS Biol **9,** e1001125.

**Corwin, J.A., and Kliebenstein, D.J.** (2017). Quantitative Resistance: More Than Just Perception of a Pathogen. Plant Cell **29,** 655-665.

**Corwin, J.A., Subedy, A., Eshbaugh, R., and Kliebenstein, D.J.** (2016). Expansive Phenotypic Landscape of Botrytis cinerea Shows Differential Contribution of Genetic Diversity and Plasticity. Molecular Plant-Microbe Interactions **29,** 287-298.

**Falconer, D.S., and Mackay, T.F.C.** (1996). Introduction to Quantitative Genetics. (Essex: Longman, Harlow).

**Fordyce, R.F., Soltis, N.E., Caseys, C., Gwinner, R., Corwin, J.A., Atwell, S., Copeland, d., Feusier, J., Subedy, A., Eshbaugh, R., and Kliebenstein, D.** (2018). Combining Digital Imaging and Genome Wide Association Mapping to Dissect Uncharacterized Traits in Plant/Pathogen Interactions. bioRxiv.

**Karasov, T.L., Chae, E., Herman, J.J., and Bergelson, J.** (2017). Mechanisms to mitigate the trade-off between growth and defense. The Plant Cell **29,** 666-680.

**Kooke, R., Kruijer, W., Bours, R., Becker, F., Kuhn, A., van de Geest, H., Buntjer, J., Doeswijk, T., Guerra, J., Bouwmeester, H., Vreugdenhil, D., and Keurentjes, J.J.B.** (2016). Genome-Wide Association Mapping and Genomic Prediction Elucidate the Genetic Architecture of Morphological Traits in Arabidopsis. Plant physiology **170,** 2187-2203.

**Mackay, T.F.C.** (2001). The genetic architecture of quantitative traits. Annual Review Of Genetics **35,** 303-339.

**Mackay, T.F.C.** (2009). Q&A: Genetic analysis of quantitative traits. Journal of Biology **8,** 23.

**Mackay, T.F.C.** (2014). Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nature Reviews Genetics **15,** 22-33.

**Züst, T., and Agrawal, A.A.** (2017). Trade-offs between plant growth and defense against insect herbivory: an emerging mechanistic synthesis. Annual review of plant biology **68,** 513-534.