

This is a very interesting manuscript that sets out to test the evolution of aggressiveness traits (latent period and lesion size) in *Z. tritici* at two temporal scales – within an growing season i.e., over a number of asexual cycles of infection versus across several years with several different epidemics, punctuated by sexual recombination. The authors (claim to) show evidence for significant evolution within seasons but less or none across seasons, and invoke a “trade-off” in the evolution of aggressiveness within versus among epidemics. It is a very interesting story, but I am rather perplexed by a number of things. I agree that the authors find a significant decrease in phenotypic variation in latent period between the beginning and the end of the 2009-2010 epidemic cycle. Mean latent period also decreased, but not significantly so, between the beginning and the end of the 2009-2010 epidemic cycle. So there was loss of variation but no significant change in mean phenotype. For lesion size there was evidence neither for loss of variation nor for shift in phenotype. So, if I understand correctly, there was not really any significant evolution within the epidemic cycle. However, the authors argue, on Page 11 of the discussion, that “A difference in aggressiveness was observed between the initial and final conidial subpopulations (Ci-2009, Cf-2010)...”. Perhaps I really missed something, but I do not see on what they can base this claim.

Given that I have doubts about the basis of the claim of the intra-epidemic patterns, I have difficulty with the arguments that set out to explain the difference between the intra- and inter-epidemic patterns, which also showed no significant changes in aggressiveness phenotypes.

In addition to this very general objection (which pains me, because this is such a good idea, but I just do not see that the data uphold the claims, but really, perhaps I have not understood something properly) I have a number of more precise suggestions for clarifying and improving the presentation.

I list several things that I would like to see clarified or corrected. Some are details of wording that could be improved to enhance readability and understanding. Some are more fundamental. I provide them in the order they appear in the manuscript.

There is some use of terminology that has me puzzled and I think this must be clarified or corrected.

1) Intro, par 2. It appears that the authors are equating “pathogen fitness” with “pathogenicity”, and defining “fitness” as “a pathogen’s ability to infect a plant host and grow”. I agree that this may be a good enough definition of “pathogenicity” (though I expect “pathogenicity” to include some pathogenic effects, i.e., damage to the host) but I do not AT ALL see that this is a good definition of “pathogen fitness”. Surely infection and growth is not sufficient unless there is also reproduction or production of propagules. Pathogens may, in some cases, infect, grow and cause damage to the host without being able to successfully reproduce. They may have high pathogenicity but zero fitness.

2) Intro, par 3-5. The authors appear to confuse “selection” with “evolution”. These are not the same thing. Indeed, selection CAN lead to evolution, i.e., a change in allele frequencies over time. On the other hand, there can be selection without evolution if the selected phenotypic variation is not heritable. Perhaps non-heritable variation for virulence is unimaginable for a plant pathologist, so the authors feel no need to distinguish between selection and evolution. However, for the more general reader please be more rigorous in your wording. Par 3: host resistance decreases because of the **evolution** of pathogen populations, with an **increase in the frequency** of virulent or aggressive strains. Par 4: I think you mean “... In several cases, **evolution** of greater aggressiveness was independent of host ...” Par 5: “gene flow” is not a “selective force”. It IS an “evolutionary force”. Here again, please distinguish clearly between selection, which is one of several evolutionary forces, and other evolutionary forces that are non-selective (e.g., mutation, drift, gene flow...)

There are also some aspects of structure of the writing that puzzle me.

Paragraph 5 of the Intro presents a lot of different ideas. It starts by presenting reasons why there may be no change or no detectable change in aggressiveness over the course of an annual epidemic (1) the change is too small, 2) selection is too small to cause evolution, 3) selection is less powerful than migration (or I suppose, drift or mutation ... why not?), 4) there may be competition within hosts (but the more aggressive may not always be the more competitive). Then it switches to talking about adaptation to host types, via serial passage or specialisation. This paragraph needs to be reorganised and structured to present a coherent message.

On the other hand, the next 2 paragraphs (6 and 7) both present the idea that the intra- versus inter-season or – epidemic processes may counteract each other. Perhaps this could also be better organised and structured.

Intro, last paragraph. This is mostly fine, but “remote dates” is not very clear. Perhaps try “at the temporal extremes” or simply “early and late over the time course of the experiment”.

Also, “sympatric and allopatric” hosts is surprising here. It is not explained, so it is rather incomprehensible. Perhaps describe more directly what you did. On the other hand, it is not so clear to me that Apache can be considered “allopatric” since it was grown during the entire experimental period over rather large areas within the broad geographical area of the study. You even state, in the Methods, on page 4, column 2, top of the page, that you consider that Apache was a likely source of the local pathogen subpopulations...

Figure 3. The circled numbers are not explained in the legend. We learn about them on pages 7 and 8... They need some sort of explanation here.

Methods: The spores for the Ai-2009 subpopulations were 6 years old when tested and compared with the (fresh) Af-2015 spores. Is this a problem? Should it be explicitly mentioned? What problems could this pose?

Methods, Page 6 2<sup>nd</sup> paragraph: What does “Three stems per plant were kept.” mean? I have never worked with wheat. I understand that it tillers. Does this mean that you trimmed or clipped the plants so they had exactly three tillers? At 7 weeks of age? This seems strange to me but perhaps that is what you did. Please explain it more completely.

Methods, *Data analysis*: This is not completely clear. Do you mean that the aggressiveness traits of each of the two pairs of subpopulations were assessed in a separate greenhouse trial, i.e., one trial for each pair of subpopulations? Or do you mean that for each of the two pairs of subpopulations there were two, independent greenhouse trials, i.e., two independent trials for each pair of subpopulations?

Results Figure 6: The figure legend is not clear or complete. Please state the subpopulation from which the two “representative” isolates were drawn. Please state, for the test of I47 (which I recognise must be an ascospore isolate, whether the test was on cv. Soisson or Apache.

And why do you present “representative” isolates instead of a central tendency and confidence intervals? Is this because you want to stress the variation observed for the different replicates?

In the text, halfway down the first column of page 8 you state that you present “A representative example ... for two Cf-2012 and Af-2015 isolates”. You mean ““A representative example ... for two isolates, one from Cf-2012 and one from Af-2015...”

Page 8 column 2 end of first paragraph: I do not really like the formulation “maladapted to their allopatric host”. I have a bit of difficulty with the term “maladapted” in general. I know what “adaptation” is. It is the product (“an adaptation”) or the process (“adaptation”) of natural selection, which we generally consider leads to an increase in performance or fitness. I am less clear about what “maladapted” means, particularly in the context of allopatry. Does “maladapted” refer to a response to natural selection that did not lead to an increase in performance or fitness, or even one that leads to a decrease in performance or fitness? This can happen of course, and there are examples, but is that what you mean? On the other hand, if Apache is really “allopatric” (though I rather doubt that, since it is planted in the cultivation area) then there is no (or little) natural selection occurring on Apache, so there is no adaptive process if there is no selection. So perhaps you mean “maladapted” as a total absence of adaptation, because there is no natural selection? But then you should rather state that the ascospore subpopulations were not adapted to the allopatric host (which is what you would expect, if it were allopatric (unless you assume that adaptation to one host (Soissons) also generates adaptation to the other host (Apache), i.e., a general scheme of adaptation rather than a specific one. What I am trying to say is that I think it is inappropriate to talk about “adaptation” (mal- or otherwise) in the absence of selection. You can say they are not adapted ... of course, because there has been no selection... There will be more on this below when we get to the Discussion.

Page 9, first column: “data not shown), but there was NO significant ...”

Page 9, second column: The “The aggressiveness of isolates depended on the cultivar on which it was assessed” implies a significant interaction. Perhaps mention this and refer to the interaction.

Table 3: Did you test the CxP interactions? Indeed, I am not really clear on what the CxI(P) interaction means. A significant I(P) term means that isolates within one or both of the subpopulations are significantly heterogeneous for the phenotype measured/tested. A significant interaction with cultivar implies that the degree of heterogeneity differs between the cultivars, I think, but does not necessarily imply that the means across isolates differ between the cultivars (which is what is suggested by the text of the results I copied above (“The aggressiveness of isolates depended on the cultivar on which it was assessed”). For example, you would find a

significant CxI(P) interaction if there were strong heterogeneity on Apache but little heterogeneity on Soissons, even if the means were similar on the two cultivars... I think...

Table 3. Please, in the note in the legend, call this "lesion size" instead of "sporulation area" (if it is the same thing) for consistency.

The second paragraph of the discussion is rather difficult. I here once again have trouble with some of the ideas expressed. Top of Page 11 you suggest that "pathogens "become maladapted" to allopatric hosts following repeated cycles on a sympatric host." There is nothing really wrong with this, and there are plenty of examples (all the serial passage experiments that work to attenuate diseases), but there are several implicit, unstated assumptions behind this idea. Indeed, IF repeated cycles of selection on host A leads to an evolutionary response and IF this evolutionary response leads to a change in the frequency of alleles that WOULD BE relevant to performance on host B (either because there is a trade-off and alleles relevant to performance on host A have negative pleiotropic effects on host B or simply because, in absence of selection on host B alleles that WOULD BE relevant to performance on host B are lost by drift), then yes, repeated cycles on host A will lead to loss of performance on host B. And I guess I would accept this being called "become maladapted" because indeed, it is natural selection on host A that leads to a change in genotype frequencies and the phenotype of performance on host B. So after objecting initially to this statement, I have to admit that it's all right, given the implicit, unstated assumptions.

However, I find the paragraph after the list of citations ("...When variation in fitness across different spatiotemporal...") quite confusing:

First, "local adaptation" can only occur under certain conditions of selection differential and migration, i.e., only when migration does not counterbalance selection. Hence there are some spatiotemporal scales at which selection is not adequate to generate local adaptation.

Second, a host shift may alter the current selection environment, but it cannot "alter the environment in which immigrant strains underwent preadaptation". You appear to be suggesting some retroactive effect, whereby the initial selective conditions are modified by what happens subsequently... This cannot be what you mean.

The next sentence: I suggest: "Immigrant strains may perform poorly on their new hosts." It is clearer and simpler.

The next sentence (Several studies have shown how...) is really unclear. I think that, again, "selection" is used where "evolution" is meant. Perhaps you mean: "Several studies have shown that virulence and aggressiveness increase over time on predominant host genotypes..."

And the last "trade-off" sentence is not really clear to me either. What causes the "selection for higher levels of aggressiveness on susceptible versus moderately resistant cultivars"?

Next paragraph: What is the "evidence for local host (mal)adaptation? Please explain this.

Page 11 column 2: First paragraph: I just have a bit of trouble with the statement "... the effects of selection are much smaller than those of diversity and the evolutionary forces responsible for its (whose?) maintenance, so the probability of detecting the same clone several times is very low." Perhaps just delete this?

I am also perplexed by your use of the term "trade-off". For me, a trade-off implies either some sort of pleiotropy, with alleles that improve one aspect decreasing another, or a common resource base that can be spent on one function or another. Here what is suggested is that the selective advantage within the season limits access to mating partners, because the successful isolates find themselves alone without a compatible partner in the upper parts of the plants, such that ascospores are only produced in the debris bearing the earlier infections, where selection had not already sorted out the genotypes and reduced the genetic variation. Here, indeed, successful isolates end up in the upper regions of the host plants in non-mating situations, more or less as victims of their own success, but they do not lose mating opportunities in the older infections where they are represented, but so are the other, less intra-epidemically successful isolates. What I am trying to say is that the more aggressive isolates do not lose, they simply fail to win, i.e., the less successful ones have similar mating chances, and hence inter-epidemic transmission chances. For me this is not a trade-off. I find the Allee effect argument convincing and compelling, but I would not call it a trade-off, i.e., I do not think that all "self-limiting" processes imply an underlying trade-off.