THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity

Justin Walsh, Luigi Pontieri, Patrizia d'Ettorre and Timothy A. Linksvayer

Article citation details
Proc. R. Soc. B **287**: 20201029.
http://dx.doi.org/10.1098/rspb.2020.1029

Review timeline
Original submission: 4 December 2019
1st revised submission: 5 May 2020
2nd revised submission: 19 May 2020
Final acceptance: 19 May 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2019-2835.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Excellent

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
Yes

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
The manuscript "Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity" by Walsh and colleagues aims to test whether cuticular hyrdrocarbons (CHCs) for the model ant species M. pharaonis are heritable utilizing a systematically produced 8-way intercross with known pedigree, GC/MS analysis of the CHCs, behavioral and life-history assays as well as Bayesian Modeling. As Proc B has established itself in recent year to publish basic, high-quality work on general question of social insect CHCs the manuscript seems highly suitable to be published in the journal.

While I quite like the manuscript (only a few smaller suggestions below), I do not agree with the authors conclusion that they have "increased the understanding of genetic architecture …" in total. There is quite a number of reviews and also research papers (actually most on social insects as well), that dissected the molecular genomic/transcriptomic logic of CHC production (specific elongases, methyl-amino transferases, CYP4G1s, fatty-acyl CoA-reducatases …). As I understand it the Linksvayer lab has genomic as well as transcriptomic resources available for this ant species (and the intercross lines), so I wonder why the authors did not include some RNAseq data (differential expression data of CHC-related genes) do explain some of the patterns they found and also present some more mechanistic insights, which would elevate their work even more and transform it into a truly remarkable piece of work (that being said, its already quite good!)

I quite like the introduction – right length and good outline of the basic problem. I only want to point to some papers by the Menzel lab that I quite enjoyed reading and which – in my opinion – could add some more to parts of the introduction:

- Menzel et al. 2017 - The evolution of a complex trait: cuticular hydrocarbons in ants evolve independent from phylogenetic constraints

- Menzel et al. 2019 - Communication versus waterproofing: the physics of insect cuticular hydrocarbons

For the chemical analysis the authors pooled 15 workers – which is fine – yet in doing so the actual variation in CHC composition within each colony (intra-colony) will be highly underestimated as each GC trace will already represent a "mean" profile. This is discussed later in the manuscript, but as the uses the term "variation" for CHC composition, I feel like the authors need to clarify that earlier.

I have one major concern regarding the analysis, or rather the identification of the cuticular hydrocarbons. As M. pharaonis appears to be transformed into another model ant species (besides the clonal raider ant maybe?), it seems highly appropriate that a paper that deals with the heritability of its CHCs also does a throughout identification. Therefore, I would invite the authors to set a high-quality paper standard and also identify the double bond positions of the M. pharaonis alkenes. As ant material is certainly not a problem and also many alkenes are quite abundant this can be easily done using Iodine-catalyzed DMDS-addition microreaction, clean up and subsequent GC/MS. This method is overall, quite cheap, easy to learn and perform and does not require any major instrumentation aside from a GC/MS

First of all – the authors did a great job and the provided R markdown made it really easy to review (thank you!). What follows is mainly nitpicking, but I'd suggest that the authors make the suggested amendments to improve their already good paper even further.
CLR/Aitchison-transformation: The author state that they "added a small constant value to each peak" as Aitchison suggested in his 1982 paper. This might not influence the over all results much, yet it was shown later that adding a small value is not an appropriate attempted and zero replacement should be performed using the "multiplicative simple replacement" method (Martin-Fernandez et al. 2003, Math Geol), which preserves the covariance structure of the data.
For the Bayesian part: Overall, the statistical analyses look pretty good to me, though the prior selection seems a little iffy since the authors changed the prior based on whether the sampling converged and the value seemed reasonable which is a bit of a circular way of picking priors (I know, people end up doing it a lot). I am wondering whether the authors should have done some prior predictive checks before sampling, as this can help alleviate the need to do the circular change of priors to make sure the model works out. As for calls of significance, I'm a little hesitant with accepting the non-overlap of 0 with the 95% confidence intervals for the genetic correlation stuff as significant as it doesn't consider multiple comparisons, but that's a general statistical concern. To be really top notch, authors should have done a prior predictive check to make sure their priors were reasonable, a parallel coordinate plot would have been nice (though their trace plot is an indicator the sampler is doing well) to check for any pathological behavior in parameter estimates (this helps indicate if any bimodality might be happening), and should have done a posterior predictive check to make sure that the model fits the data well. A corner plot would have been helpful to get a sense of the posterior. The posterior predictive is the most informative of these because it gives you a sense of whether the statistical model can capture the distribution of your data.

The results section seems like the underdeveloped part of this manuscript – it is far too short and for me not understandable only reading the main text + figures. For instance, the phenotypic correlations are basically described as "well, some CHCs and three collective behaviors are somehow – positive, negative, how much? – correlated with each other. Period. As no figure is provided, in the main text, one may as well skip this information, as the reader has simply no idea what the results are, also some of the other result-text is equally vacuous (while the R markdown is really helpful to understand the results). Therefore, the authors have to amend their results section – explain the results, may include fig S3 and S4 in the main manuscript and be clear what is significant.

Review form: Reviewer 2

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Marginal

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Marginal

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
The study by Walsh et al. deals with an important and interesting topic – the heritability of cuticular hydrocarbon profiles, and their association with fitness traits. Thus, a study like this can provide valuable results and basics for future research. The authors report studies from a breeding experiment of the pharaoh ant, which have been bred into the fifth generation. They ask three questions: firstly, which hydrocarbons are heritable, secondly, which ones discriminate best between colonies, and thirdly, which ones are associated to fitness (i.e. productivity) and may be under selection.
The aims of this study read highly promising. However, upon reading results and discussion, the reader is left with lots of correlations (or similar results), some of which are significant and others are not. They are presented (some of them in the supplement only), but not interpreted. Thus, as a reader, I cannot gain much new knowledge from this study apart from being overwhelmed with details. Please think carefully how things can be interpreted, and what new knowledge can be gained from the result that correlation A (or heritability A, or selection gradient A) is significant and B is not. Don't get me wrong, I think this dataset can be very valuable. However, the analyses presented here do not really provide insights since there is very little biological interpretation.
Major issues
- the authors should think carefully about the biological interpretation of the results – e.g. what do we actually learn about the genetic architecture of CHC profiles? What do we learn about selection on CHCs, and why may CHC profile and fitness be correlated?
- the authors should present more details on the results, and apply rigorous multiple test correction (FDR) to each analysis. Fitness measures, behavioural data and chemical data should also be presented separately (not only in combination), such that the reader can see e.g. which

CHCs differ between colonies, whether between-colony variation is larger than within-colony variation, which behavioural traits co-vary etc. This would enhance the credibility of this dataset a lot.
- the authors should consider not only analysing the individual hydrocarbons, but also pool them according to homologous series, and pool them according to substance class (linear alkane, alkene, monomethyl alkane, dimethyl alkane). This might additionally provide evidence on inter-relations between hydrocarbons – CHCs of the same homologous series are likely to be part of the same biosynthetic pathway. The different analyses can be presented alongside each other.
- previous studies on heritability of CHCs are mentioned, but not discussed in detail. It would be interesting to see if the same CHCs were found to be heritable in other studies. You might also discuss studies on QTL mapping of CHC profiles (e.g. Nasonia wasps: Niehuis et al. 2011 Heredity)

Minor comments

Abstract

Line 20 The term "pharaoh ant laboratory mapping population" is a bit odd, please rephrase. What is a 'mapping population'? (see also line 94)

Introduction

Line 52-55 This does not really make sense. Firstly, it is not clear yet whether solitary insects always have simpler profiles than social insects (see e.g. Kather and Martin 2015 J Chem Ecol). Secondly, insects of the same species usually produce the same set of compounds, albeit in different quantities. Hence, inter-individual exchange would not make the profile more complex. Thirdly, inter-individual cuticular hydrocarbon exchange should not make the genetic architecture more complex (it may only make it more difficult to study).

Methods

Line 109-116 The behavioural data should also be presented in the results (independently from the CHC data). Did they co-vary? Did they differ between colonies? A PCA or some other ordination of the five behaviours would be helpful.

Line 138: This is close to 50% of the data being discarded!! Please explain – can you be sure that the results are still valid?

Line 144: replace "normalized" by "standardized"

Line 151: please provide this small constant value. Did you try out different values to see whether this changed the analyses?

Line 156-158 I appreciate the detailed and informative supplementary file! It is nice and easy to follow.

Line 160-186: Can you give the range of possible heritability values? Do they range from 0 to 1? Showing the CHC results would be good (as an ordination) to get an impression of within-replicate and between-replicate variation. Were the behavioural results consistent between colony replicates of the same colony? This would be a valuable additional information.

Pool gyne and worker pupae as an additional measure of fitness. Were the two measures correlated?

Line 191: How long were the sampling dates ("blocks") apart from each other? Did the climate vary between the dates? This is important as climate can influence CHC profiles. Is this what you mean by "environmental variance", or is env. variance an additional thing?

Were FDR corrections applied to the heritability estimates and to the selection gradient estimates? FDR is only reported for the behaviour-CHC correlations, but is definitely necessary for all your analyses.

Supplementary figure 1 – the pedigree is a bit hard to understand – it looks as if there was only one colony sampled in F5. How did you get the 48 colony genotypes?

Line 202 replace "are" by "were"

Line 205-206 please explain "case-bootstrapped"

Line 217 – omit "better"

Line 220 – probably intercolonial (intraspecific), but not interspecific. We don't know which compounds are relevant for interspecific recognition here (other species probably won't care about the colony identity of the pharaoh ant).

Results

Line 230 31 correlations out of how many were significant? Was this after FDR correction?

Line 243-245 please provide more details – the reader should not have to consult the supplement.
Line 248: Is "accuracy of 17.1%" the error rate, or the likelihood to be correct?

Discussion
Line 254 – please also discuss studies on Nasonia (e.g. Büllesbach et al. 2013 Heredity). Which CHCs were heritable in this study, and in the other cited ones?
Line 256 – While the evolutionary forces on CHC profiles may be different in social insects, we don't know whether the genetic architecture of CHC synthesis differs between social and solitary insects.
Line 259 – genetic architecture… but can you pinpoint what we actually learn about the genetic architecture underlying CHC variation from this study?
Correlations between CHCs can be much stronger… please see (and discuss) Martin, S. J., & Drijfhout, F. P. (2009): How reliable is the analysis of complex cuticular hydrocarbon profiles by multivariate statistical methods?. Journal of chemical ecology, 35(3), 375-382.
Line 270-273 This might actually be a statistical effect – if certain hydrocarbons differ between colonies, they will show up in the random forest analysis. In addition, they may be more likely to show a higher heritability estimate – because, if other CHCs do not differ between colonies, the heritability will be estimated as low. Only if they differ, they can have a higher heritability estimate I think. Although you may be right and heritable CHCs are those that differ most between colonies, you should acknowledge that this may also be a statistical effect.
Line 292-294 This should be tested (statistically) and presented in the results.
Line 296-298 Please provide more details about these other studies – which hydrocarbons were correlated in these other studies? Can you compare your results to them? The same applies to study [59] (line 275-278) – even if fewer compounds were found and fewer individuals extracted, results on variation should nevertheless be compared qualitatively.
Line 303-304 Yes, but please provide more interpretation and more details to the link of CHC variation and variation in colony productivity. As it is now, you only present some correlations and leave the interpretation to the reader.
Line 309-310 Were the two fitness measures correlated? Please show this in the results section. I suggest to use the sum of the two measures as a third fitness measure (and present it in the supplement, if necessary).
Line 320-322 50% relative humidity can actually be quite dry (hence stressful) for an ant if the soil is dry. Were the ants kept in open boxes or closed boxes (i.e. with a lid)? If there were lids, the humidity of the climate chamber does not provide much information on the humidity inside the nest box anyway.
Line 359-360 I find this a bit over-interpreted… please state explicitly what we learn about the genetic architecture of the hydrocarbon profile here. That some hydrocarbons correlate to productivity in the lab may be a statistical artefact (the more so as there was no FDR correction), so I would not conclude that this is evidence for natural selection.
Fig. 3b please also show the data points of the regression lines in the plots.

Decision letter (RSPB-2019-2835.R0)

08-Jan-2020

Dear Mr Walsh:

I am writing to inform you that your manuscript RSPB-2019-2835 entitled "Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission,

provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Please note that this decision may (or may not) have taken into account confidential comments.

In your revision process, please take a second look at how open your science is; our policy is that all data involved with the study should be made openly accessible-- see:
https://royalsociety.org/journals/ethics-policies/data-sharing-mining/
Insufficient sharing of data can delay or even cause rejection of a paper.

Sincerely,
Professor John Hutchinson
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
This paper has now been evaluated by two reviewers, who both recommended substantial revisions; the reviewers provided detailed comments that I will not reiterate here, but these comments would need to be addressed head-on in a revision.

From my own reading, I really liked the scope and approach taken in this paper; I would like to see more studies take a quantitative genetics approach to studying the genetic architecture of and selection on ant traits. So I think it is well suited to Proceedings B, as it really helps to push research in ant evolution in a new direction, in my opinion.

I thought the introduction could be broadened somewhat. Instead of starting right in with cuticular hydrocarbons, the authors might consider a general paragraph about ants and social evolution, and then explain why cuticular hydrocarbons are an important trait mediating ant interactions (as well as functioning to prevent water loss). I think it might make for a more compelling read.

I also think just a tad more information (one short paragraph?) about the biology of pharaoh ants is in order--it would help readers understand whether selection measured in a lab population is likely to bear any resemblance to selection in the wild. For example, the fact that this ant species

reproduces by budding needs to be explained before the reader gets to the discussion. Regarding the estimates of selection, I think somewhere it should be acknowledged that the number of sexuals produced by a colony may not be a good measure of fitness; on this subject, I always think of Deborah Gordon's 2013 Journal of Animal Ecology paper ("Colony life history and lifetime reproductive success of red harvester ant colonies"), because she found that there was no relationship between the number of gynes a colony produced and the number of offspring colonies the colony made at her study site. (Sigh.)

Both reviewers brought up the issue of correcting for multiple tests, and I will extend their comments a little further by noting that instead of measuring genetic correlations by running one model for each pairwise combination of hydrocarbons, the authors should probably have used MCMCglmm to estimate a genetic variance-covariance (G) matrix for all traits simultaneously. In a similar vein, instead of estimating a selection gradient on each hydrocarbon trait individually, one possible approach is to do a principal components analysis on all the hydrocarbon variables, and then estimate selection on each principal component; there are some pitfalls with this approach (see Chong et al. 2018 in Evolution Letters, "A note on measuring natural selection on principal component scores"), but I think it is worth considering for this highly multivariate trait dataset.

Because addressing the reviewer comments and doing some new analyses of the data would involve substantial revisions to the paper, I am recommending it be rejected in its current form, but with the opportunity to resubmit.

Megan Frederickson

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
The manuscript “Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity” by Walsh and colleagues aims to test whether cuticular hyrdrocarbons (CHCs) for the model ant species M. pharaonis are heritable utilizing a systematically produced 8-way intercross with known pedigree, GC/MS analysis of the CHCs, behavioral and life-history assays as well as Bayesian Modeling. As Proc B has established itself in recent year to publish basic, high-quality work on general question of social insect CHCs the manuscript seems highly suitable to be published in the journal.

While I quite like the manuscript (only a few smaller suggestions below), I do not agree with the authors conclusion that they have “increased the understanding of genetic architecture …” in total. There is quite a number of reviews and also research papers (actually most on social insects as well), that dissected the molecular genomic/transcriptomic logic of CHC production (specific elongases, methyl-amino transferases, CYP4G1s, fatty-acyl CoA-reducatases …). As I understand it the Linksvayer lab has genomic as well as transcriptomic resources available for this ant species (and the intercross lines), so I wonder why the authors did not include some RNAseq data (differential expression data of CHC-related genes) do explain some of the patterns they found and also present some more mechanistic insights, which would elevate their work even more and transform it into a truly remarkable piece of work (that being said, its already quite good!)

I quite like the introduction – right length and good outline of the basic problem. I only want to point to some papers by the Menzel lab that I quite enjoyed reading and which – in my opinion – could add some more to parts of the introduction:

- Menzel et al. 2017 - The evolution of a complex trait: cuticular hydrocarbons in ants evolve independent from phylogenetic constraints

- Menzel et al. 2019 - Communication versus waterproofing: the physics of insect cuticular hydrocarbons

For the chemical analysis the authors pooled 15 workers – which is fine – yet in doing so the actual variation in CHC composition within each colony (intra-colony) will be highly underestimated as each GC trace will already represent a "mean" profile. This is discussed later in the manuscript, but as the uses the term "variation" for CHC composition, I feel like the authors need to clarify that earlier.

I have one major concern regarding the analysis, or rather the identification of the cuticular hydrocarbons. As M. pharaonis appears to be transformed into another model ant species (besides the clonal raider ant maybe?), it seems highly appropriate that a paper that deals with the heritability of its CHCs also does a throughout identification. Therefore, I would invite the authors to set a high-quality paper standard and also identify the double bond positions of the M. pharaonis alkenes. As ant material is certainly not a problem and also many alkenes are quite abundant this can be easily done using Iodine-catalyzed DMDS-addition microreaction, clean up and subsequent GC/MS. This method is overall, quite cheap, easy to learn and perform and does not require any major instrumentation aside from a GC/MS
First of all – the authors did a great job and the provided R markdown made it really easy to review (thank you!). What follows is mainly nitpicking, but I'd suggest that the authors make the suggested amendments to improve their already good paper even further.

CLR/Aitchison-transformation: The author state that they "added a small constant value to each peak" as Aitchison suggested in his 1982 paper. This might not influence the over all results much, yet it was shown later that adding a small value is not an appropriate attempted and zero replacement should be performed using the "multiplicative simple replacement" method (Martin-Fernandez et al. 2003, Math Geol), which preserves the covariance structure of the data.

For the Bayesian part: Overall, the statistical analyses look pretty good to me, though the prior selection seems a little iffy since the authors changed the prior based on whether the sampling converged and the value seemed reasonable which is a bit of a circular way of picking priors (I know, people end up doing it a lot). I am wondering whether the authors should have done some prior predictive checks before sampling, as this can help alleviate the need to do the circular change of priors to make sure the model works out. As for calls of significance, I'm a little hesitant with accepting the non-overlap of 0 with the 95% confidence intervals for the genetic correlation stuff as significant as it doesn't consider multiple comparisons, but that's a general statistical concern. To be really top notch, authors should have done a prior predictive check to make sure their priors were reasonable, a parallel coordinate plot would have been nice (though their trace plot is an indicator the sampler is doing well) to check for any pathological behavior in parameter estimates (this helps indicate if any bimodality might be happening), and should have done a posterior predictive check to make sure that the model fits the data well. A corner plot would have been helpful to get a sense of the posterior. The posterior predictive is the most informative of these because it gives you a sense of whether the statistical model can capture the distribution of your data.

The results section seems like the underdeveloped part of this manuscript – it is far too short and for me not understandable only reading the main text + figures. For instance, the phenotypic correlations are basically described as "well, some CHCs and three collective behaviors are somehow – positive, negative, how much? – correlated with each other. Period. As no figure is provided, in the main text, one may as well skip this information, as the reader has simply no idea what the results are, also some of the other result-text is equally vacuous (while the R markdown is really helpful to understand the results). Therefore, the authors have to amend their results section – explain the results, may include fig S3 and S4 in the main manuscript and be clear what is significant.

Referee: 2

Comments to the Author(s)
The study by Walsh et al. deals with an important and interesting topic – the heritability of cuticular hydrocarbon profiles, and their association with fitness traits. Thus, a study like this can provide valuable results and basics for future research. The authors report studies from a breeding experiment of the pharaoh ant, which have been bred into the fifth generation. They ask three questions: firstly, which hydrocarbons are heritable, secondly, which ones discriminate best between colonies, and thirdly, which ones are associated to fitness (i.e. productivity) and may be under selection.
The aims of this study read highly promising. However, upon reading results and discussion, the reader is left with lots of correlations (or similar results), some of which are significant and others are not. They are presented (some of them in the supplement only), but not interpreted. Thus, as a reader, I cannot gain much new knowledge from this study apart from being overwhelmed with details. Please think carefully how things can be interpreted, and what new knowledge can be gained from the result that correlation A (or heritability A, or selection gradient A) is significant and B is not. Don't get me wrong, I think this dataset can be very valuable. However, the analyses presented here do not really provide insights since there is very little biological interpretation.
Major issues
- the authors should think carefully about the biological interpretation of the results – e.g. what do we actually learn about the genetic architecture of CHC profiles? What do we learn about selection on CHCs, and why may CHC profile and fitness be correlated?
- the authors should present more details on the results, and apply rigorous multiple test correction (FDR) to each analysis. Fitness measures, behavioural data and chemical data should also be presented separately (not only in combination), such that the reader can see e.g. which CHCs differ between colonies, whether between-colony variation is larger than within-colony variation, which behavioural traits co-vary etc. This would enhance the credibility of this dataset a lot.
- the authors should consider not only analysing the individual hydrocarbons, but also pool them according to homologous series, and pool them according to substance class (linear alkane, alkene, monomethyl alkane, dimethyl alkane). This might additionally provide evidence on inter-relations between hydrocarbons – CHCs of the same homologous series are likely to be part of the same biosynthetic pathway. The different analyses can be presented alongside each other.
- previous studies on heritability of CHCs are mentioned, but not discussed in detail. It would be interesting to see if the same CHCs were found to be heritable in other studies. You might also discuss studies on QTL mapping of CHC profiles (e.g. Nasonia wasps: Niehuis et al. 2011 Heredity)
Minor comments
Abstract
Line 20 The term "pharaoh ant laboratory mapping population" is a bit odd, please rephrase. What is a 'mapping population'? (see also line 94)
Introduction
Line 52-55 This does not really make sense. Firstly, it is not clear yet whether solitary insects always have simpler profiles than social insects (see e.g. Kather and Martin 2015 J Chem Ecol). Secondly, insects of the same species usually produce the same set of compounds, albeit in different quantities. Hence, inter-individual exchange would not make the profile more complex. Thirdly, inter-individual cuticular hydrocarbon exchange should not make the genetic architecture more complex (it may only make it more difficult to study).
Methods
Line 109-116 The behavioural data should also be presented in the results (independently from the CHC data). Did they co-vary? Did they differ between colonies? A PCA or some other ordination of the five behaviours would be helpful.
Line 138: This is close to 50% of the data being discarded!! Please explain – can you be sure that the results are still valid?

Line 144: replace “normalized” by “standardized”
Line 151: please provide this small constant value. Did you try out different values to see whether this changed the analyses?
Line 156-158 I appreciate the detailed and informative supplementary file! It is nice and easy to follow.
Line 160-186: Can you give the range of possible heritability values? Do they range from 0 to 1? Showing the CHC results would be good (as an ordination) to get an impression of within-replicate and between-replicate variation. Were the behavioural results consistent between colony replicates of the same colony? This would be a valuable additional information.
Pool gyne and worker pupae as an additional measure of fitness. Were the two measures correlated?
Line 191: How long were the sampling dates (“blocks”) apart from each other? Did the climate vary between the dates? This is important as climate can influence CHC profiles. Is this what you mean by “environmental variance”, or is env. variance an additional thing?
Were FDR corrections applied to the heritability estimates and to the selection gradient estimates? FDR is only reported for the behaviour-CHC correlations, but is definitely necessary for all your analyses.
Supplementary figure 1 – the pedigree is a bit hard to understand – it looks as if there was only one colony sampled in F5. How did you get the 48 colony genotypes?
Line 202 replace “are” by “were”
Line 205-206 please explain “case-bootstrapped”
Line 217 – omit “better”
Line 220 – probably intercolonial (intraspecific), but not interspecific. We don’t know which compounds are relevant for interspecific recognition here (other species probably won’t care about the colony identity of the pharaoh ant).
Results
Line 230 31 correlations out of how many were significant? Was this after FDR correction?
Line 243-245 please provide more details – the reader should not have to consult the supplement.
Line 248: Is “accuracy of 17.1%” the error rate, or the likelihood to be correct?

Discussion
Line 254 – please also discuss studies on Nasonia (e.g. Büllesbach et al. 2013 Heredity). Which CHCs were heritable in this study, and in the other cited ones?
Line 256 – While the evolutionary forces on CHC profiles may be different in social insects, we don’t know whether the genetic architecture of CHC synthesis differs between social and solitary insects.
Line 259 – genetic architecture… but can you pinpoint what we actually learn about the genetic architecture underlying CHC variation from this study?
Correlations between CHCs can be much stronger… please see (and discuss) Martin, S. J., & Drijfhout, F. P. (2009): How reliable is the analysis of complex cuticular hydrocarbon profiles by multivariate statistical methods?. Journal of chemical ecology, 35(3), 375-382.
Line 270-273 This might actually be a statistical effect – if certain hydrocarbons differ between colonies, they will show up in the random forest analysis. In addition, they may be more likely to show a higher heritability estimate – because, if other CHCs do not differ between colonies, the heritability will be estimated as low. Only if they differ, they can have a higher heritability estimate I think. Although you may be right and heritable CHCs are those that differ most between colonies, you should acknowledge that this may also be a statistical effect.
Line 292-294 This should be tested (statistically) and presented in the results.
Line 296-298 Please provide more details about these other studies – which hydrocarbons were correlated in these other studies? Can you compare your results to them? The same applies to study [59] (line 275-278) – even if fewer compounds were found and fewer individuals extracted, results on variation should nevertheless be compared qualitatively.
Line 303-304 Yes, but please provide more interpretation and more details to the link of CHC variation and variation in colony productivity. As it is now, you only present some correlations and leave the interpretation to the reader.

Line 309-310 Were the two fitness measures correlated? Please show this in the results section. I suggest to use the sum of the two measures as a third fitness measure (and present it in the supplement, if necessary).
Line 320-322 50% relative humidity can actually be quite dry (hence stressful) for an ant if the soil is dry. Were the ants kept in open boxes or closed boxes (i.e. with a lid)? If there were lids, the humidity of the climate chamber does not provide much information on the humidity inside the nest box anyway.
Line 359-360 I find this a bit over-interpreted… please state explicitly what we learn about the genetic architecture of the hydrocarbon profile here. That some hydrocarbons correlate to productivity in the lab may be a statistical artefact (the more so as there was no FDR correction), so I would not conclude that this is evidence for natural selection.
Fig. 3b please also show the data points of the regression lines in the plots.

Author's Response to Decision Letter for (RSPB-2019-2835.R0)

See Appendix A.

RSPB-2020-1029.R0

Review form: Reviewer 1

Recommendation
Accept as is

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Excellent

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
Yes

Comments to the Author
After reading the authors response and their revised manuscript, I am pleased to see that they addressed or try to addressed the majority of all concerns raised.

Given the recent Covid-19 situation, I can live with the fact that the CHC analysis does not include double bonds positions, as this is truly rather a cosmetic change. Yet, I hope that this will be part of future work of the authors.

Regarding the statistics the authors also sufficiently explain their prior selection and answered my comment about the CLR transformation.

Thanks for the revision and great job.

Decision letter (RSPB-2020-1029.R0)

19-May-2020

Dear Mr Walsh

I am pleased to inform you that your Review manuscript RSPB-2020-1029 entitled "Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity" has been accepted for publication in Proceedings B. Congratulations!!

The referee(s) do not recommend any further changes. Therefore, please proof-read your manuscript carefully and upload your final files for publication. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let me know immediately.

To upload your manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, upload a new version through your Author Centre.

Before uploading your revised files please make sure that you have:

1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".

2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. Please note that PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file from the main text and the file name should contain the author's name and journal name, e.g authorname_procb_ESM_figures.pdf
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI. Please see: https://royalsociety.org/journals/authors/author-guidelines/

4) Data-Sharing and data citation
It is a condition of publication that data supporting your paper are made available. Data should be made available either in the electronic supplementary material or through an appropriate repository. Details of how to access data should be included in your paper. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link
http://datadryad.org/submit?journalID=RSPB&manu=RSPB-2020-1029 which will take you to your unique entry in the Dryad repository.

If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.

5) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit https://royalsociety.org/journals/authors/author-guidelines/.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your final version. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Dr John Hutchinson, Editor
mailto:proceedingsb@royalsociety.org

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
After reading the authors response and their revised manuscript, I am pleased to see that they addressed or try to addressed the majority of all concerns raised.

Given the recent Covid-19 situation, I can live with the fact that the CHC analysis does not include double bonds positions, as this is truly rather a cosmetic change. Yet, I hope that this will be part of future work of the authors.

Regarding the statistics the authors also sufficiently explain their prior selection and answered my comment about the CLR transformation.

Thanks for the revision and great job.

Sincerely,
Proceedings B
mailto: proceedingsb@royalsociety.org

Decision letter (RSPB-2020-1029.R1)

19-May-2020

Dear Mr Walsh

I am pleased to inform you that your manuscript entitled "Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Paper charges
An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit https://royalsociety.org/journals/ethics-policies/media-embargo for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

Associate Editor
Comments to Author:
This paper has now been evaluated by two reviewers, who both recommended substantial revisions; the reviewers provided detailed comments that I will not reiterate here, but these comments would need to be addressed head-on in a revision.

From my own reading, I really liked the scope and approach taken in this paper; I would like to see more studies take a quantitative genetics approach to studying the genetic architecture of and selection on ant traits. So I think it is well suited to Proceedings B, as it really helps to push research in ant evolution in a new direction, in my opinion.

> We thank the Editor for the kind words and we certainly agree that the type of research we present is valuable and very interesting.

I thought the introduction could be broadened somewhat. Instead of starting right in with cuticular hydrocarbons, the authors might consider a general paragraph about ants and social evolution, and then explain why cuticular hydrocarbons are an important trait mediating ant interactions (as well as functioning to prevent water loss). I think it might make for a more compelling read.

> We agree with the Editor that a broadened introduction would make the manuscript more compelling. We reworked the beginning of the Introduction (lines 33 - 44) to discuss the role of altruism and kin selection in the evolution of sociality and how cuticular hydrocarbons function in social insect nestmate recognition.

I also think just a tad more information (one short paragraph?) about the biology of pharaoh ants is in order--it would help readers understand whether selection measured in a lab population is likely to bear any resemblance to selection in the wild. For example, the fact that this ant species reproduces by budding needs to be explained before the reader gets to the discussion.

> We agree that more information about the biology of pharaoh ants would be useful. We added a few sentences to the beginning of the Methods section (lines 94 - 96) describing that pharaoh ants tend to live in association with humans and reproduce by budding. Furthermore, we added a sentence (lines 192 - 196) to the selection estimates section of the Methods to remind the reader that pharaoh ants reproduce by budding and therefore both reproductives and workers are necessary to produce new colonies. Finally, we added a sentence to the discussion (lines 389 - 392) arguing that because pharaoh ants live in association with humans, the laboratory conditions of our study may be more similar to the natural conditions of pharaoh ants than other species.

Regarding the estimates of selection, I think somewhere it should be acknowledged that the number of sexuals produced by a colony may not be a good measure of fitness; on this subject, I always think of Deborah Gordon's 2013 Journal of Animal Ecology paper ("Colony life history

and lifetime reproductive success of red harvester ant colonies"), because she found that there was no relationship between the number of gynes a colony produced and the number of offspring colonies the colony made at her study site. (Sigh.)

> We agree that the Ingram et al. 2013 paper is very interesting and we cite it. We also acknowledge that our measure, and indeed any measure of fitness, is imperfect. However, we note that their measure of fitness (the number of daughter colonies that have survived at least one season) also has issues because it potentially conflates the fitness of the founding gyne and her incipient colony (e.g., gyne mating success and survival, early colony founding success and survival) with parental fitness (i.e. fitness of the founding colony). These issues aren't unique to social insects and the implications of assigning fitness to parents or offspring have been discussed in detail by Wolf and Wade 2001 (https://onlinelibrary.wiley.com/doi/full/10.1046/j.1420-9101.2001.00277.x). We suspect that one reason gyne production does not correlate with the number of daughter colonies in harvester ants is that gyne survival rate and founding success is very low, so that incipient colony founding success may depend more on the traits and genotype of the gyne (and possibly her mates) -- which implies selection acting on gyne traits and not on the traits of the parental colony. (Of course, in reality, gyne traits and colony-level traits are likely linked, for example through size-number trade-offs at the colony level.) In contrast, in a budding species such as *M. pharaonis*, we expect that the number of reproductives produced, and the number of workers produced, are better measures of fitness. We added a brief discussion of these issues and a further explanation of why we think that our productivity measures are good measures of colony fitness (lines 338 - 350).

Both reviewers brought up the issue of correcting for multiple tests, and I will extend their comments a little further by noting that instead of measuring genetic correlations by running one model for each pairwise combination of hydrocarbons, the authors should probably have used MCMCglmm to estimate a genetic variance-covariance (G) matrix for all traits simultaneously. In a similar vein, instead of estimating a selection gradient on each hydrocarbon trait individually, one possible approach is to do a principal components analysis on all the hydrocarbon variables, and then estimate selection on each principal component; there are some pitfalls with this approach (see Chong et al. 2018 in Evolution Letters, "A note on measuring natural selection on principal component scores"), but I think it is worth considering for this highly multivariate trait dataset.

> We agree that the multiple comparisons issue is an obvious issue when considering so many individual peaks, and we also agree that there are unfortunate issues with all possible solutions (e.g., using PCA). We have carefully weighed the pros and cons of the various approaches and made several additions in the revised manuscript. Below, we first explain how we deal with the issue of accounting for multiple comparisons and statistical significance, and next we explain how we added multivariate analyses to our previously reported univariate (and bivariate) analyses.

For the selection estimates, it is in particular of biological interest whether a given trait is actually associated with fitness or not (i.e. assessing the significance of fitness-trait associations is important). In accordance with reviewer suggestions, we added an FDR correction to the p values from our univariate selection estimates. Initially we had 14 (when defining fitness as the production of reproductives) and 9 (fitness as the production of workers) significant linear estimates. After an FDR correction, we have 10 and 6 significant estimates so our conclusion, that selection is shaping the hydrocarbon profile of our ants, is still valid after correcting for multiple comparisons.

> However, this issue is not as clear for the heritability and genetic correlation estimates. Our goal is not to identify “significant” heritability and genetic correlation estimates (we don’t calculate or report p values) but rather to identify patterns that increase our understanding of the genetic architecture underlying the hydrocarbon profile. In accordance with this goal, we never say anything about significance in regards to our heritability estimates. Instead, we simply discuss the estimates and CIs. On the other hand, we initially had discussed significance (defined as when the CIs did not overlap with zero) in terms of our genetic correlation estimates. Although this approach is common in the field of quantitative genetics, it is also common to not discuss significance in terms of genetic correlations and instead only discuss the estimates (often describing them as “high” or “low”). We feel that the latter approach is more in line with our goal for this manuscript and therefore we have removed all discussion of significance of our genetic correlation estimates, including the asterisks previously designing significance in Figure 2.

> We agree with the Editor that a variance-covariance (G) matrix would be ideal and we did previously try to do calculate it. However, a full G matrix including all 32 peaks was not feasible using our MCMCglmm approach (or any other approach we are aware of) as the models fail to converge when so many traits are included (see Wilson et al. 2009 “An ecologist’s guide to the animal model” for a brief consideration of this issue). Additionally, even if it were feasible for a 32-trait model to work without any problems, such a model would likely take months to run as models including only a few traits typically take a few days. A common approach to getting around the issue of too many traits is to estimate all the possible bivariate combinations and take the average of the bivariate estimates (e.g. Lihoreau et al. 2016 “Kin discrimination increases with odor distance in the German cockroach”). We estimated the heritability of all the bivariate combinations (we already had these estimates from our bivariate genetic correlation estimates) between our 32 peaks and found that the average of the bivariate estimates were nearly indistinguishable from the univariate estimates. We included a plot showing the univariate and bivariate estimates in the supplemental material (supplementary figure 8) and mentioned this in the Methods (lines 183 - 185) and Results sections (lines 243 - 244).

> We also agree with the Editor that a principal component analysis would be useful to better explore our dataset, especially to allow us to use formal multivariate models (given the caveats for PCA). Therefore, in the revised manuscript we added a principal components analysis on 29

of our 32 peaks (removing peaks 23, 26, and 31). We removed the three peaks that included zeros (0.001 after we added this small number to all peaks) because PCA is very sensitive to small values and these three peaks dominated the loadings for the first few PCs when we included them and the samples that included zeros were large outliers.

> We estimated the heritability and strength of selection of the first 8 PCs, which explained about 90% of the variance in our dataset, using multivariate models. We did not estimate genetic correlations between PCs because, by definition, PCs are orthogonal to each other and therefore not correlated. Our heritability estimates on the 8 PCs range from 0.004 to 0.20 (mean = 0.12) and those with higher heritabilities tend to have large contributions from the peaks with high heritabilities from our univariate analyses (e.g. PC 6 has the highest heritability and the strongest loading for PC 6 is peak 12, which has one of the highest heritabilities from our univariate estimates). Our selection analyses suggest that there is positive linear selection acting on PC1. We included the results of the heritability and selection analyses on the PCs in the text (lines 245 - 246; 264 - 266) and included figures (PCA bi plots supplementary figures 4-7; PC heritability supplementary figure 9) and tables (PCA variance supplementary table 2, PCA loadings supplementary table 3, PC selection estimates supplementary table 7). .

> Overall, we feel that these changes and additional analyses strengthen our claim that the cuticular hydrocarbon profile is heritable and shaped by selection in our lab population of *M. pharaonis.* We thank the Editor and reviewers for pointing out these issues.

Because addressing the reviewer comments and doing some new analyses of the data would involve substantial revisions to the paper, I am recommending it be rejected in its current form, but with the opportunity to resubmit.

Megan Frederickson

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
The manuscript “Ant cuticular hydrocarbons are heritable and associated with variation in colony productivity” by Walsh and colleagues aims to test whether cuticular hyrdrocarbons (CHCs) for the model ant species M. pharaonis are heritable utilizing a systematically produced 8-way intercross with known pedigree, GC/MS analysis of the CHCs, behavioral and life-history assays as well as Bayesian Modeling. As Proc B has established itself in recent year to publish basic, high-quality work on general question of social insect CHCs the manuscript seems highly suitable to be published in the journal.

> We thank the reviewer for the positive comments.

While I quite like the manuscript (only a few smaller suggestions below), I do not agree with the authors conclusion that they have “increased the understanding of genetic architecture …” in total. There is quite a number of reviews and also research papers (actually most on social insects as well), that dissected the molecular genomic/transcriptomic logic of CHC production (specific elongases, methyl-amino transferases, CYP4G1s, fatty-acyl CoA-reducatases …).

> Thank you, we are glad that you liked our manuscript. We agree with the reviewer that there has been a lot of great genetic/molecular/transcriptomic work characterizing the molecular pathways involved in the production of cuticular hydrocarbons. However, we do feel that our paper fills a large gap in our understanding of the genetic contribution and fitness consequences of *variation* (as opposed to the molecular mechanisms underlying *expression*) in social insect hydrocarbons. Specifically, our paper is one of the first to demonstrate that social insect cuticular hydrocarbons are heritable, genetically correlated, and shaped by selection (and the first to use a large, multi-generational pedigree to do so).

We generally think it is important to note that researchers use two distinct but complementary definitions of “genetic basis” and “genetic architecture”: 1. the molecular mechanisms underlying *trait expression*, and 2. the contribution of DNA sequence variation to *trait variation* within a population. Our manuscript contributes to elucidating the genetic architecture underlying variation in CHC profile and the fitness consequences, in the sense of the second definition, which complements research focused on the first definition.

As I understand it the Linksvayer lab has genomic as well as transcriptomic resources available for this ant species (and the intercross lines), so I wonder why the authors did not include some RNAseq data (differential expression data of CHC-related genes) do explain some of the patterns they found and also present some more mechanistic insights, which would elevate their work even more and transform it into a truly remarkable piece of work (that being said, its already quite good!)

We do have population genomic and a great deal of transcriptomic data (e.g., for foragers and nurses, queens, brood stages, etc.). Such transcriptomic data could potentially elucidate the molecular mechanisms underlying the development and expression of CHC profiles, but the most straightforward approach would be to compare transcriptomic profiles in tissues involved in CHC production (e.g., oenocytes) or perhaps across developmental stages with differences in CHC profile. It is less clear how our transcriptomic datasets based on whole bodies or body segments can be used to identify genes involved in CHC production. In any case, such datasets would try to get at the molecular mechanisms underlying CHC production (definition 1 of genetic basis/architecture), while our current manuscript is focused on the genetic contribution to variation in CHC profile (definition 2). In the future, we hope to combine these two to identify genetic variants underlying variation in expression (i.e. expression quantitative trait loci) for

genes associated with the production of CHCs, but this is well beyond the scope of the current manuscript.

I quite like the introduction – right length and good outline of the basic problem. I only want to point to some papers by the Menzel lab that I quite enjoyed reading and which – in my opinion – could add some more to parts of the introduction:

- Menzel et al. 2017 - The evolution of a complex trait: cuticular hydrocarbons in ants evolve independent from phylogenetic constraints

- Menzel et al. 2019 - Communication versus waterproofing: the physics of insect cuticular hydrocarbons

>Thank you very much for these suggestions. We agree that these two papers are great studies on the evolution and physical properties of cuticular hydrocarbons and we have added multiple citations of these papers to our manuscript.

For the chemical analysis the authors pooled 15 workers – which is fine – yet in doing so the actual variation in CHC composition within each colony (intra-colony) will be highly underestimated as each GC trace will already represent a “mean” profile. This is discussed later in the manuscript, but as the uses the term “variation” for CHC composition, I feel like the authors need to clarify that earlier.

> We agree that this should be explained earlier in the manuscript and therefore added text (lines 127 - 128) to the Methods section.

I have one major concern regarding the analysis, or rather the identification of the cuticular hydrocarbons. As M. pharaonis appears to be transformed into another model ant species (besides the clonal raider ant maybe?), it seems highly appropriate that a paper that deals with the heritability of its CHCs also does a throughout identification. Therefore, I would invite the authors to set a high-quality paper standard and also identify the double bond positions of the M. pharaonis alkenes. As ant material is certainly not a problem and also many alkenes are quite abundant this can be easily done using Iodine-catalyzed DMDS-addition microreaction, clean up and subsequent GC/MS. This method is overall, quite cheap, easy to learn and perform and does not require any major instrumentation aside from a GC/MS

> We agree with the reviewer that identifying the position of the double bonds of the *M. pharaonis* alkenes is a very worthwhile endeavor. However, this would clearly require collecting and analyzing more data, which in particular given the current COVID-19 situation, is not possible.

First of all – the authors did a great job and the provided R markdown made it really easy to review (thank you!). What follows is mainly nitpicking, but I'd suggest that the authors make the suggested amendments to improve their already good paper even further.

> We thank the reviewer for the positive comments and are thankful the R markdown was appreciated!

CLR/Aitchison-transformation: The author state that they "added a small constant value to each peak" as Aitchison suggested in his 1982 paper. This might not influence the over all results much, yet it was shown later that adding a small value is not an appropriate attempted and zero replacement should be performed using the "multiplicative simple replacement" method (Martin-Fernandez et al. 2003, Math Geol), which preserves the covariance structure of the data.

> We agree that the multiplicative simple replacement method is a good approach and to verify our results, we used this approach to deal with the zeros in our data in addition to the approach we had used previously (adding a small value to each sample). We followed the method outlined in Martin-Fernandez et al. 2003, Math Geol and used an R package written by the authors called "zCompositions" (https://rdrr.io/cran/zCompositions/man/zCompositions-package.html). After transforming our data using the multiplicative simple replacement method, we re-calculated the heritability and genetic correlations of a subset of our data. The new data set was extremely similar to our old dataset. We only re-calculated heritability and genetic correlations on a subset because to redo the entire set of estimates would take 2-3 months of computing time. We re-did the estimates for the peaks that contained zeros (peaks 23, 26, and 31) as these peak values were the most likely to be changed by using the new method. The new results are very similar from our previous results and do not change the conclusions of our paper. We added text to the manuscript (lines 251 - 256) describing that we used the multiplicative simple replacement method on a subset of our estimates to verify that our estimates using the CLR/Aitchinson transformation were accurate and included two tables in the supplemental material showing the old and new heritability and genetic correlation estimates.

For the Bayesian part: Overall, the statistical analyses look pretty good to me, though the prior selection seems a little iffy since the authors changed the prior based on whether the sampling converged and the value seemed reasonable which is a bit of a circular way of picking priors (I know, people end up doing it a lot). I am wondering whether the authors should have done some prior predictive checks before sampling, as this can help alleviate the need to do the circular change of priors to make sure the model works out. As for calls of significance, I'm a little hesitant with accepting the non-overlap of 0 with the 95% confidence intervals for the genetic correlation stuff as significant as it doesn't consider multiple comparisons, but that's a general statistical concern. To be really top notch, authors should have done a prior predictive check to make sure their priors were reasonable, a parallel coordinate plot would have been nice (though their trace plot is an indicator the sampler is doing well) to check for any

pathological behavior in parameter estimates (this helps indicate if any bimodality might be happening), and should have done a posterior predictive check to make sure that the model fits the data well. A corner plot would have been helpful to get a sense of the posterior. The posterior predictive is the most informative of these because it gives you a sense of whether the statistical model can capture the distribution of your data.

> We understand the reviewers concerns about prior specification and the possibility of circularity in how priors are selected. However, we feel that the priors we used are traditional and highly recommended (e.g. Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models). We did not simply test different priors until we found priors that seemed to perform well. Rather, we started with priors recommended elsewhere for data similar to ours and found that, based on our multiple diagnostic tests, performed well with our data.

> In regards to considering genetic correlations significant, as we described in our response to the Editor, we removed all mention of significance when referring to the genetic correlation estimates.

The results section seems like the underdeveloped part of this manuscript – it is far too short and for me not understandable only reading the main text + figures. For instance, the phenotypic correlations are basically described as “well, some CHCs and three collective behaviors are somehow – positive, negative, how much? – correlated with each other. Period. As no figure is provided, in the main text, one may as well skip this information, as the reader has simply no idea what the results are, also some of the other result-text is equally vacuous (while the R markdown is really helpful to understand the results). Therefore, the authors have to amend their results section – explain the results, may include fig S3 and S4 in the main manuscript and be clear what is significant.

> Although we unfortunately do not have space to include more figures in the main text, we added additional information to the Results section to better describe our findings (lines 272 - 273).

Referee: 2

Comments to the Author(s)
The study by Walsh et al. deals with an important and interesting topic – the heritability of cuticular hydrocarbon profiles, and their association with fitness traits. Thus, a study like this can provide valuable results and basics for future research.

> We thank the reviewer for the encouraging words.

The authors report studies from a breeding experiment of the pharaoh ant, which have been bred into the fifth generation.

> Just to clarify, the colonies in our heterogeneous stock mapping population were systematically intercrossed for nine generations. The supplementary figure illustrating the pedigree for a single colony for 5 generations is included to help the reader understand the breeding scheme. We have added text to the caption to further clarify this issue. Furthermore, we added an additional supplementary figure that shows the entirety of our pedigree.

They ask three questions: firstly, which hydrocarbons are heritable, secondly, which ones discriminate best between colonies, and thirdly, which ones are associated to fitness (i.e. productivity) and may be under selection.
The aims of this study read highly promising. However, upon reading results and discussion, the reader is left with lots of correlations (or similar results), some of which are significant and others are not. They are presented (some of them in the supplement only), but not interpreted. Thus, as a reader, I cannot gain much new knowledge from this study apart from being overwhelmed with details. Please think carefully how things can be interpreted, and what new knowledge can be gained from the result that correlation A (or heritability A, or selection gradient A) is significant and B is not. Don't get me wrong, I think this dataset can be very valuable. However, the analyses presented here do not really provide insights since there is very little biological interpretation.
Major issues
- the authors should think carefully about the biological interpretation of the results – e.g. what do we actually learn about the genetic architecture of CHC profiles? What do we learn about selection on CHCs, and why may CHC profile and fitness be correlated?

> Thank you for your comments. We note that our study is a quantitative genetic study estimating heritabilities, genetic correlations, and selection, and like all such quantitative genetic studies (and like most transcriptomic, genomic, etc. studies) is descriptive. We of course can try to provide explanations for the heritability, genetic correlation, and selection patterns we observe, but we also want to be very careful and not overinterpret the results. Overall, we feel that we carefully and sufficiently interpret our results and answer all of these questions (discussed below) posed by the reviewer. Furthermore, we are hesitant to add much more text because we are already up against the page limit of the journal.

1. What do we actually learn about the genetic architecture of CHC profiles?
> The first four paragraphs of our Discussion section (lines 284 - 330) highlights what our study contributes to our understanding of the genetic architecture of cuticular hydrocarbon profiles. In these paragraphs, we discuss that our study is only of the first studies to estimate the heritability of and genetic correlations between social insect hydrocarbons. We continue to discuss that we expect hydrocarbons with high heritabilities to be among the best at discriminating between colonies and point out that our results fit this prediction. Next, we discuss the implications of our findings for nestmate recognition and how our results fit the prediction that alkenes and branched hydrocarbons are expected to play a larger role in nestmate recognition than linear

alkanes. Finally, we discuss the implications and possible explanations for significant genetic correlations.

2.What do we learn about selection on CHCs?
> In lines 331 - 371 of our Discussion section, we highlight our selection results. Our study is one of the first to demonstrate that natural selection (albeit in a laboratory setting) shapes the social insect cuticular hydrocarbon profile, a point that was often assumed but rarely tested. Furthermore, we found that selection on the production of workers and reproductives seems to be aligned, which is noteworthy because selection can act on workers and reproductives differently.

3. Why may CHC profile and fitness be correlated?
> We discuss why the hydrocarbon profile and fitness may be correlated in lines 351 - 371, including possibly through desiccation resistance and foraging rates, influences on colony division of labor or task allocation, and/or through nestmate recognition and inter-colonial aggression.

- the authors should present more details on the results, and apply rigorous multiple test correction (FDR) to each analysis. Fitness measures, behavioural data and chemical data should also be presented separately (not only in combination), such that the reader can see e.g. which CHCs differ between colonies, whether between-colony variation is larger than within-colony variation, which behavioural traits co-vary etc. This would enhance the credibility of this dataset a lot.

> We added more detail to our Results section to better describe our findings (lines 243 - 250; 272 - 273; 279 - 281). Additionally, we added an FDR correction to our univariate selection estimates and removed all mention of significance in regards to our genetic correlation estimates. More detail on the behavioral data can be found in Walsh et al. 2019 “Ant collective behavior behavior is heritable and shaped by selection” https://www.biorxiv.org/content/10.1101/567503v1. Unfortunately, we do not have room to provide additional figures to better show the variation in our data in the main text. However, the PCA figure and table in the supplement provide a useful way for readers to get a sense of which hydrocarbons differed between colonies.

- the authors should consider not only analysing the individual hydrocarbons, but also pool them according to homologous series, and pool them according to substance class (linear alkane, alkene, monomethyl alkane, dimethyl alkane). This might additionally provide evidence on inter-relations between hydrocarbons – CHCs of the same homologous series are likely to be part of the same biosynthetic pathway. The different analyses can be presented alongside each other.

> We agree with the reviewer that such analyses would be interesting but we feel that we do not have room in our manuscript for additional analyses. We organized our figures to separate out four different hydrocarbon classes (linear alkanes, alkenes, monomethyl alkanes, and dimethyl alkanes) and order the hydrocarbons within each class by branch length. We feel that this organization allows the reader to get a sense of differences between compound types and branch lengths. In our discussion, we make the claim that linear alkenes and monomethyl alkanes have better discrimination power than linear alkanes so we did add a test of whether hydrocarbon structural classes varied in their ability to discriminate between colony genotypes using a linear model (lines 279 - 281).

- previous studies on heritability of CHCs are mentioned, but not discussed in detail. It would be interesting to see if the same CHCs were found to be heritable in other studies. You might also discuss studies on QTL mapping of CHC profiles (e.g. Nasonia wasps: Niehuis et al. 2011 Heredity)

> We agree that a more in depth review of previously hydrocarbon heritability studies would be worthwhile but unfortunately we do not have room to expand on our discussion of them. We did add the Nasonia wasp citation to the introduction (lines 55 - 56)

Minor comments
Abstract
Line 20 The term “pharaoh ant laboratory mapping population” is a bit odd, please rephrase. What is a ‘mapping population’? (see also line 94)

> A mapping population is a population suitable for association studies/qtl mapping and is a standard phrase used in quantitative genetic studies such as ours. We feel that this is an accurate way to describe our laboratory population and have decided to keep the phrase.

Introduction
Line 52-55 This does not really make sense. Firstly, it is not clear yet whether solitary insects always have simpler profiles than social insects (see e.g. Kather and Martin 2015 J Chem Ecol). Secondly, insects of the same species usually produce the same set of compounds, albeit in different quantities. Hence, inter-individual exchange would not make the profile more complex. Thirdly, inter-individual cuticular hydrocarbon exchange should not make the genetic architecture more complex (it may only make it more difficult to study).

> We did not mean to imply that solitary insects have simpler cuticular hydrocarbon profiles than social insects or that inter-individual exchange between social insects makes their profile more complex. Rather, our point was that the genetic architecture underlying the profile of social insects is more complex than solitary insects simply because individuals can acquire hydrocarbons from social partners. This inter-individual exchange would certainly make the genetic architecture underlying the trait more complex because, like all socially-influenced traits, the genetic architecture would depend on the collective genetic makeup of the entire colony,

rather than the genetic makeup of a single individual. To clarify our point, we added text (lines 58 - 59) to highlight that genes in both a focal individual and the individual's social partners affect the genetic architecture of the hydrocarbon profile.

Methods
Line 109-116 The behavioural data should also be presented in the results (independently from the CHC data). Did they co-vary? Did they differ between colonies? A PCA or some other ordination of the five behaviours would be helpful.

> The behavioral data is presented and analyzed in detail in a separate manuscript, Walsh et al. 2019 https://www.biorxiv.org/content/10.1101/567503v3 . Unfortunately, we do not have room in our manuscript to describe the behavioral data in detail but we added a sentence (lines 120 - 122) here to better describe the behavioral and colony productivity data.

Line 138: This is close to 50% of the data being discarded!! Please explain – can you be sure that the results are still valid?

> Yes, we are sure our results are valid despite having to discard some of our samples -- our estimates of heritability, genetic correlations, and selection are based on the included samples. Sample contamination is common in GC/MS work and we had to ship our samples (extracted hydrocarbons) from the US to France which could have resulted in a higher percentage of unusable samples since the extracted hydrocarbons must remain chilled. We used relative peak abundance (rather than absolute abundance) to control for arbitrary differences in peak heights. Additionally, we discarded samples from colony genotypes with only one good sample which increased the number of samples we did not include.

Line 144: replace "normalized" by "standardized"

> We made the suggested change (line 143)

Line 151: please provide this small constant value. Did you try out different values to see whether this changed the analyses?

> The small value that we added to all of the samples was 0.0001 and added this information to the text (line 146). Because of the amount of time it takes to run these models and because our results using the multiplicative simple replacement method suggested by reviewer 1 were extremely similar to our initial results, we did not try different small values. However, we are confident that any small value would yield extremely similar results.

Line 156-158 I appreciate the detailed and informative supplementary file! It is nice and easy to follow.

> We thank the reviewer for the kind words and are happy to know that our markdown file was appreciated!

Line 160-186: Can you give the range of possible heritability values? Do they range from 0 to 1? Showing the CHC results would be good (as an ordination) to get an impression of within-replicate and between-replicate variation. Were the behavioural results consistent between colony replicates of the same colony? This would be a valuable additional information. Pool gyne and worker pupae as an additional measure of fitness. Were the two measures correlated?

> We added (lines 158 - 159) the possible range of heritability values (0 to 1) and genetic correlations (-1 to 1). We also included plots of our the PCs from our PCA with groupings by colony genotype as supplemental material (supplementary figures 4-) to give the reader a sense of within- and between-replicate variation. Detailed information about behavioral variation is reported in Walsh et al. 2019 “Ant collective behavior is heritable and shaped by selection” and we do not have room in our paper to include it.

> Yes, the two fitness measures were correlated (spearman rank, rho = 0.611, $p < 0.001$) and we added this information to the Results section (lines 259 - 261)

> Finally, we do not think it is worthwhile to pool gyne and worker pupae as a third measure of fitness. We added a further explanation of why we think that our productivity measures are good measures of colony fitness (lines 338 - 350). We also added a few sentences to the beginning of the Methods section (lines 94 - 96) describing that pharaoh ants reproduce by budding and therefore reproductives and workers are required to establish new colonies. Furthermore, we added a sentence (lines 192 - 196) to the selection estimates section of the Methods to remind the reader that pharaoh ants reproduce by budding and therefore both reproductives and workers are necessary to produce new colonies.

Line 191: How long were the sampling dates (“blocks”) apart from each other? Did the climate vary between the dates? This is important as climate can influence CHC profiles. Is this what you mean by “environmental variance”, or is env. variance an additional thing?

> Because it took a lot of time to set up the colonies, standardize colony size, allow the ants time to acclimate, conduct behavioral assays, and collect hydrocarbon samples, we could not sample all our colonies at once. Instead, we sampled across different blocks (usually containing 15 to 18 colony replicates) from May 2016 to November 2016 (we added this information about separate blocks to the Methods section, lines 104 - 105). The ants were always maintained on a 12:12 hour light:dark cycle and at 27 ± 1 °C and 50% relative humidity (lines 106 - 107). The “environmental variance” term in our heritability and genetic correlation models is a standard quantitative genetic term that includes all non-genetic variance. In practice, the environmental

variance term is expected to include random, uncontrolled environmental factors, since our samples were kept in the laboratory in climate control chambers, where temperature, humidity, food, etc. are controlled as tightly as possible.

Were FDR corrections applied to the heritability estimates and to the selection gradient estimates? FDR is only reported for the behaviour-CHC correlations, but is definitely necessary for all your analyses.

> We discussed this issue above in our response to the Editor but just to reiterate we agree that correcting for multiple tests is important and therefore, we added an FDR correction to the p values from our univariate selection estimates. In regards to the heritability and genetic correlation estimates, like other quantitative genetic studies, our goal is not to identify "significance" (we don't calculate or report p values) but rather to identify patterns that increase our understanding of the genetic architecture underlying the hydrocarbon profile. In accordance with this goal, we never say anything about significance in regards to our heritability estimates. Instead, we simply discuss the estimates and CIs. On the other hand, we initially had discussed significance (defined as when the CIs did not overlap with zero) in terms of our genetic correlation estimates. Although this approach is common in the field of quantitative genetics, it is also common to not discuss significance in terms of genetic correlations and instead only discuss the estimates (often describing them as high or low). We feel that the latter approach is more in line with our goal for this manuscript and therefore we have removed all discussion of significance of our genetic correlation estimates, including the asterisks previously designing significance in Figure 2.

Supplementary figure 1 – the pedigree is a bit hard to understand – it looks as if there was only one colony sampled in F5. How did you get the 48 colony genotypes?

> Supplementary figure 1 (now supplementary figure 2) is supposed to show a representative colony from the F5 generation and is primarily meant to show how genetic variation present in the parental lineages can be found in subsequent generations. We added further explanation to the figure legend to better explain that this is just an example and is not meant to show our entire pedigree. To show the entire pedigree, we made and included an additional supplementary figure (the current supplementary figure 1) that shows the entire pedigree from parental generation to the current generation and reference this new figure in the main text (line 99)

Line 202 replace "are" by "were"

> We reworked this sentence (lines 200 - 201) to be in active voice rather than passive voice and corrected the incorrect verb tense.

Line 205-206 please explain "case-bootstrapped"

> This is a nonparametric bootstrapping option offered by the R package “gsg” and is commonly used by researchers using the selection gradient approach outlined by Morrissey and Sakredjda (2013). We moved this sentence to the Results section and added “(nonparametric”) (line 264 - 267) to better describe this approach.

Line 217 – omit “better”

> We revised this from better to best (lines 228 - 229) because that sentence is meant to explain that the point of the random forest analysis is to identify which hydrocarbons are the most variable about colony genotypes. We feel that the new sentence is clear and informative for the reader.

Line 220 – probably intercolonial (intraspecific), but not interspecific. We don’t know which compounds are relevant for interspecific recognition here (other species probably won’t care about the colony identity of the pharaoh ant).

> We agree with the reviewer that the compounds identified by the random forest analysis as being the best at discriminating between *M. pharaonis* colonies are likely more important in intraspecific recognition. We revised this sentence (lines 229 - 232) to say “...thus highlighting compounds that might be involved in nestmate recognition.”

Results
Line 230 31 correlations out of how many were significant? Was this after FDR correction?

> We removed all discussion of significance related to our genetic correlation estimates.

Line 243-245 please provide more details – the reader should not have to consult the supplement.

> Unfortunately, we do not have room to add an additional figure to the main text but we added a sentence (lines 273 - 274) to better describe the correlations between hydrocarbons and behaviors.

Line 248: Is “accuracy of 17.1%” the error rate, or the likelihood to be correct?

> 17.1% was the error rate. We corrected this in the text (lines 277 - 278)

Discussion
Line 254 – please also discuss studies on Nasonia (e.g. Büllesbach et al. 2013 Heredity). Which CHCs were heritable in this study, and in the other cited ones?

> We agree that a more in depth review of previously hydrocarbon heritability studies would be worthwhile but unfortunately we do not have room to expand on our discussion of them. We did add the Nasonia wasp citation to the introduction (lines 55 - 56)

Line 256 – While the evolutionary forces on CHC profiles may be different in social insects, we don't know whether the genetic architecture of CHC synthesis differs between social and solitary insects.

> We agree with the reviewer that there is a lot we do not know about the differences between social and solitary cuticular hydrocarbon profiles. We feel that our sentences pointed out by the reviewer (now on lines 254 - 259) highlight this fact.

Line 259 – genetic architecture… but can you pinpoint what we actually learn about the genetic architecture underlying CHC variation from this study?

> We answered this above when the reviewer first asked this:

> The first four paragraphs of our Discussion section (lines 285 - 331) highlights what our study contributes to our understanding of the genetic architecture of cuticular hydrocarbon profiles. In these paragraphs, we discuss that our study is only of the first studies to estimate the heritability of and genetic correlations between social insect hydrocarbons. We continue to discuss that we expect hydrocarbons with high heritabilities to be among the best at discriminating between colonies and point out that our results fit this prediction. Next, we discuss the implications of our findings for nestmate recognition and how our results fit the prediction that alkenes and branched hydrocarbons are expected to play a larger role in nestmate recognition than linear alkanes. Finally, we discuss the implications and possible explanations for significant genetic correlations.

Correlations between CHCs can be much stronger… please see (and discuss) Martin, S. J., & Drijfhout, F. P. (2009): How reliable is the analysis of complex cuticular hydrocarbon profiles by multivariate statistical methods?. Journal of chemical ecology, 35(3), 375-382.

> On lines 324 - 331, we mention that other studies in fruit flies have found strong genetic correlations between hydrocarbons. We then continue to discuss why we might see strong genetic correlations and what they may mean for the evolution of hydrocarbons. We agree with the reviewer that we could devote more space to comparing our correlation estimates to other studies but given our lack of space (we are right at the page limit), we feel that a brief discussion of the evolutionary implications of genetic correlations is more valuable.

Line 270-273 This might actually be a statistical effect – if certain hydrocarbons differ between colonies, they will show up in the random forest analysis. In addition, they may be more likely to show a higher heritability estimate – because, if other CHCs do not differ between colonies, the heritability will be estimated as low. Only if they differ, they can have a higher heritability

estimate I think. Although you may be right and heritable CHCs are those that differ most between colonies, you should acknowledge that this may also be a statistical effect.

> We do not believe our results are due to statistical effects. The random forest analysis does not take into account the pedigree information and therefore it is possible that hydrocarbons that do well at discriminating between colonies could vary among colonies due to random, uncontrolled environmental factors. We do not expect that to be the case- we expect these hydrocarbons to vary largely because of genetic effects (i.e. have high heritabilities, where CHC profile is predicted by pedigree) and that is what we see in our heritability estimates.

Line 292-294 This should be tested (statistically) and presented in the results.

> We added a statistical test (linear model; results on lines 280 - 281) to test whether the four compound types differed in their discrimination ability and found that alkenes and monomethyl alkanes were significantly better at discriminating than linear alkanes.

Line 296-298 Please provide more details about these other studies – which hydrocarbons were correlated in these other studies? Can you compare your results to them? The same applies to study [59] (line 275-278) – even if fewer compounds were found and fewer individuals extracted, results on variation should nevertheless be compared qualitatively.

> We discussed the issue of further comparing our study to other studies that estimated genetic correlations above:
On lines 324 - 331, we mention that other studies in fruit flies have found strong genetic correlations between hydrocarbons. We then continue to discuss why we might see strong genetic correlations and what they may mean for the evolution of hydrocarbons. We agree with the reviewer that we could devote more space to comparing our correlation estimates to other studies but given our lack of space (we are right at the page limit), we feel that a brief discussion of the evolutionary implications of genetic correlations is more valuable.

> Similarly, we do not feel that we have enough space to further compare our results to the other study on cuticular hydrocarbons in *M. pharaonis*

Line 303-304 Yes, but please provide more interpretation and more details to the link of CHC variation and variation in colony productivity. As it is now, you only present some correlations and leave the interpretation to the reader.

> In the paragraph (lines 352 - 372) following the lines referenced here by the reviewer we interpret our selection results in detail. We agree with the reviewer that a more in depth discussion of why hydrocarbon variation may be linked to colony fitness would be ideal. However, we unfortunately do not have space in our current manuscript. Additionally, we can't really say much more without speculating because we are not really certain what is going on. We feel that our current interpretation is carefully worded to not over-interpret our findings.

Line 309-310 Were the two fitness measures correlated? Please show this in the results section. I suggest to use the sum of the two measures as a third fitness measure (and present it in the supplement, if necessary).

> Yes, the two fitness measures were correlated (spearman rank, rho = 0.611, $p < 0.001$) and we added this information to the Results section (lines 259 - 261). We do not think it is worthwhile to pool gyne and worker pupae as a third measure of fitness. We added a further explanation of why we think that our productivity measures are good measures of colony fitness (lines 339 - 351). We also added a few sentences to the beginning of the Methods section (lines 94 - 96) describing that pharaoh ants reproduce by budding and therefore reproductives and workers are required to establish new colonies. Furthermore, we added a sentence (lines 192 - 196) to the selection estimates section of the Methods to remind the reader that pharaoh ants reproduce by budding and therefore both reproductives and workers are necessary to produce new colonies.

Line 320-322 50% relative humidity can actually be quite dry (hence stressful) for an ant if the soil is dry. Were the ants kept in open boxes or closed boxes (i.e. with a lid)? If there were lids, the humidity of the climate chamber does not provide much information on the humidity inside the nest box anyway.

> We do not believe that our ants were under desiccation stress. Many different species of ants are commonly kept at 50% RH. Our ants do not nest in soil, rather they nest in between two glass slides. Their nest boxes do not have lids, so 50% RH is an accurate measure of the humidity experienced by the ants, at least outside the nest. Inside the nest, the humidity is almost certainly higher. Our ants have thrived in 50% RH for over 10 years and always had access to water. Regardless, we believe that our language in the text leaves open the possibility that our ants experienced at least low levels of water stress.

Line 359-360 I find this a bit over-interpreted… please state explicitly what we learn about the genetic architecture of the hydrocarbon profile here. That some hydrocarbons correlate to productivity in the lab may be a statistical artefact (the more so as there was no FDR correction), so I would not conclude that this is evidence for natural selection.

> We have previously addressed how our study increased our understanding of the genetic architecture of the hydrocarbon profile. We agree that our laboratory-based study may not reflect selection in the field but we feel that we sufficiently address this concern in the previous paragraph (lines 386 - 396). Furthermore, we did add an FDR correction to our selection estimates and still find significant estimates of selection, so we feel confident in concluding that the link between hydrocarbon variation and fitness is not a statistical artifact. We have tried to be as careful as possible and describe possible explanations for the patterns we observed so

that future studies (e.g. with experimental manipulation) can further attempt to tease apart how variation in the traits we measured is causally associated with variation in fitness.

Fig. 3b please also show the data points of the regression lines in the plots.

> We made Figure 3b using the fitness.landscape function of the R package “gsg.” These plots show the population mean fitness across a subset of the data range and therefore plotting the individual points would be misleading (see Morrissey & Sakrejda 2013 “Unification of regression-based methods for the analysis of natural selection” for a more in depth description of the fitness landscape plots).