INRA – CBGP,
Campus International de Baillarguet, CS 30016
34988 Montferrier-sur-Lez cedex, France
E-mail: nathalie.charbonnel@inra.fr

Montpellier, February 4th 2019

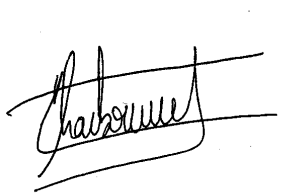
Dear Editor,

Please find enclosed a revision of our manuscript entitled 'Differential immune gene expression associated with contemporary range expansion of two invasive rodents in Senegal', by Charbonnel *et al*., which we would like to submit for recommendations in *PCI Ecology*.

We have addressed all issues raised by the reviewers and detailed responses are provided.

This manuscript is not under consideration for publication in another journal or book. Its submission for publication has been approved by all the relevant authors and institutions, and all individuals entitled to authorship have been so named.

Yours sincerely,
Nathalie Charbonnel (on the behalf of all the co-authors)



Differential immune gene expression associated with contemporary range expansion of two invasive rodents in Senegal

Nathalie Charbonnel1*, Maxime Galan1, Caroline Tatard1, Anne Loiseau1, Christophe Diagne2, Ambroise Dalecky4, Hugues Parrinello5, Stephanie Rialle5, Dany Severac5 and Carine Brouat2

Our responses to reviewers' comments are indicated in colour in the text below. The same colours are used in the revised manuscript to highlight changes made in response to these reviewers' comments.

Reviews by Nadia Aubin-Horth

This ms by Charbonnel and colleagues aims to test predictions associated with the EICA hypothesis (« evolution of increased competitive ability ») in two invasive species. They use transcriptomes from spleen of individuals for populations with different history of invasion (recent versus ancient) and four replicate populations within each invasion history type to find if certain biological functions are differentially activated between recent and ancient populations.

Main comments

I think that this ms is interesting as the RNA-seq data is used to test an evolutionary model. It is also interesting that the dataset seems to support the opposite of the predicted pattern. The ms is also well written and most analyses are useful and appropriate.

Thank you for these positive comments.

However, I have comments on the manuscript that would need to be addressed to streamline it and make it clearer to more generalist readers.

We have tried to address all these comments. Modifications in the manuscript are reported in blue.

1-The EICA hypothesis is central to the ms but its predictions are not spelled out clearly for non specialists.

The authors should explain in greater details what is proposed by the EICA hypothesis. This could be done on page 3 in the introduction. Specifically, it is not clear what increased competitive ability means, since "success" (or fitness) is not measured in the individuals studied.

In the beginning of the introduction, we added some details to better explain what the EICA and EICA-refined hypotheses suggest (P3-4). Briefly, these hypotheses are based on the life history theory and on the idea that defence strategies might influence invasion success. The hypotheses suggest that energy allocation to immunity should be modulated during the course of invasion, considering changes in parasite pressure (e.g., enemy release) and needs for range expansion (dispersal, reproduction).

In our study, we assume that the rodents sampled in recently invaded sites are successful colonizers. The potential changes observed in life history traits between anciently and recently invaded sites may therefore reflect traits favoured at invasion front (*i.e.,* that provided better fitness). But we agree that we do not directly measure individual fitness. We did not include this explanation in the text.

On page 5, the authors state: "We investigated two alternative hypotheses. On one hand, we expected an overall higher immune gene expression of rodent populations in recently invaded sites, as a response to novel parasite pressures encountered » This is clear

« On the other hand, we expected trade-offs between immune pathways in recently invaded sites under the EICA-refined hypothesis. » Trade-offs between immune pathways and …what? With other traits (which ones?) . If it is a trade off with other traits such as life history traits that is expected, shouldn't

these life history traits be measured also? Or were the authors expecting to find that trade offs are visible at the gene expression levels, as in Aubin-Horth et al. 2005 (Proc Roy Soc B) who did this in an unrelated study and another tissue in fish? This is especially important for the EICA-refined hypothesis

Or is it « between immune pathways » as in « different immune strategies? « if yes, explain what those different immune pathways are for non specialists. If you expect that different immune strategies have different costs, make a table describing each one with predictions of what is expected

We have now clarified in the introduction that the EICA-refined hypothesis relies on trade-offs between immune pathways, and we provided details about the different costs of these pathways P4).

As you mentioned, we expect to see these trade offs at the gene expression level: There is no other phenotypic estimation of immune traits.

We next made our expectation clearer, and provided some examples of costly and cost-effective immune strategies.

Unfortunately, considering the complexity of the immune system (and especially for vertebrates), we do not think that predictions about the immune pathways supposed to be costly or cost-effective should be detailed in the manuscript. Moreover, due to the difficulty to estimate immune costs, this information is only available for few immune pathways (see Lee 2006; https://academic.oup.com/icb/article/46/6/1000/714860). We therefore only provided some examples in our manuscript (P4).

This topic comes back on page 12 in the discussion.

"we did not find any evidence of immunity or particular immune pathways being dampened at the expense of other life history traits or of less energetically costly immune defences. «

Which other life history traits did the authors measure? Not clear.

We modified the text to better reflect the hypotheses tested in this manuscript and that have now been clarified in the introduction. Namely, we did not measure any other life history trait than immune related ones. We tested that invasion success is driven by changes in immune strategies that are modulated by trade-offs between energetically and cost-effective immune pathways; or by an overall increase of immune responses that enable to face new parasites encountered during the course of invasion (P3-4).

I understand that the authors found the opposite, with higher immunity expression instead of lower, but still, how could they validate their original prediction with the data they had collected?

We have previously worked on helminths, bacteria and virus communities infecting these rodent populations (and in native species too). These results have been published (Diagne et al. 2016, 2017) and we cite them in the revised version while interpreting the patterns of differential gene expression observed in this study (P12-13).

On page 10, the authors talk about their predictions "Contrary to what was expected, along the mouse invasion route, all immune-related genes detected were over-expressed in recently invaded sites, and among them, inflammation and complement pathways were over-represented." This should be clearer in the introduction

The introduction has been modified and the hypotheses have now been clarified and detailed (P3-4).

We also modified this sentence in the discussion to explain that inflammation is expected to be energetically costly while complement activation is expected to be cost effective, according to Lee (2006). (P12).

Minor comments

INTRODUCTION

p.3 “From an eco-evolutionary perspective, invasion success may rely on pre-adaptation within the original range 5,6 or on the rapid evolution of phenotypic traits that would be advantageous in newly colonized areas 5,7. Some supports for this latter process come from the identification of phenotypic variation along invasion gradients, «

I think that phenotypic gradients could be found even if it is based on standing genetic variation, if the alleles are at low frequency in the original population, such that it is difficult to sample the resulting phenotype and / or the alleles that result in the new phenotype are rarely found in the same individuals, but the smaller effective population size of the invading front could result in the « encounter » of these alleles by chance.

We are sorry that the reviewer might have misunderstood what we indicated. We did not assume that the rapid evolution of phenotypic traits in newly colonized areas was due to new mutations. This phenotypic evolution could result from new / standing genetic variation, or polyploidization or hybridization, or phenotypic plasticity. Moreover, we must recognize that phenotypic gradients might also result from stochastic events (allele surfing), so that the identification of phenotypic variation along invasion gradient was definitely not a relevant argument for selective processes. We therefore removed the sentence about the environmental gradients correlating with phenotypic variations in some invasion cases, as it might be confusing (P3).

RESULTS

Even if this information is given in table 1, it would be best to orient the reader by starting the paragraph with 1-tissue studied (spleen) 2-that there are two invasion history (it is really well explained on page 15, maybe pu tat beginning of results?) 3-that there are 4 sites within each invasion history 4-that POOLS of individuals are used. This is important when we try to understand the analysis strategy presented later with the 4vs4 and 8 vs8

This information has been added at the beginning of the Results section (P6).

The analysis is very complicated with the different ways and different levels of stringency. Could it be possible to only present one?

The two statistical approaches were designed with biostatisticians and bioinformaticians. They are complementary in terms of statistical power and robustness. As recommended by the other referees, we improved the revised version to better explain and argue the way we performed the statistical analyses.

Briefly, in the ‘4vs4 approach’, we keep only one out of the two replicates analysed by locality. This allows to suppress the factor ‘locality’ of the analysis: we compare the four recently invaded versus the four anciently invaded regions, and here the four samples per condition are independent from each other and considered as replicate. As there is no reason to keep the first or the second locality replicate, all the 256 combinations of replicate selection has been made.

In the ‘8vs8 approach’, the eight samples from each condition are kept, and the locality factor is added to the design in the statistical analysis settings in order to consider that replicates are paired by locality.

As a high variability is observed between the samples, combining the results of these two approaches seemed to be a prudent choice, enabling the results to be as robust as possible. Thus, the genes identified with the '8vs8 approach' and declared as DE in more than 85% of the 256 combinations of the '4vs4 approach' were eventually declared as DE. The threshold of 85% were set according to the visualisation of a barplot representing the number of '4vs4' comparisons in which a gene (highlighted in the '8vs8 approach') is found to be differentially expressed."

The text has been modified to clarify the design and better justify the combination of the two approaches (P7-8 & P18-19).

p.9 the authors use the DE genes in mouse to study protein-protein interactions. They focus on the genes related to immunity (using GO terms) and find that they interact within a cluster. Is that a trivial finding / unsurprising? Are there examples of proteins related to the same biological function that have no interaction? What does it tell us more than what we already knew?

There is a misunderstanding. We analysed protein-protein interactions both on the whole set of DE genes (364 for the house mouse dataset) and on the immune related DE genes (73 genes for the house mouse). Such network analysis is classically performed to identify important clusters and key proteins in

the whole network. As such, we did not point out in our manuscript that all proteins interact. We highlighted main clusters of interactions. This result enabled to better visualize and understand the metabolic and immune pathways that are involved in the differences observed between anciently and recently invaded sites.

Because the network was built on databases considering direct (physical) and indirect (functional associations) interactions, as well as genetic interactions and shared pathways, it is unsurprising that proteins related to the same biological function are found to interact. What was not trivial (especially with regard to the network built on immune related DE genes) concerned the structure of the network. We could have found a random network, with no given cluster isolated, or we could have found a hierarchical cluster (what we observed for the two datasets considered) with proteins organised in groups of small highly connected/functional networks (e.g., complement proteins; fibrinogen and serine ; aloprotein and haptoglobin). Therefore, the network observed indicates that changes occurring between recently and anciently invaded sites involve at least three groups of proteins that are interacting with each other's in functional complexes and pathways. Changes do not rely on a single cluster of immune related proteins.

TABLE

Table 1 Add a column on the left with « invasion time » instead of using superscript for each population name

This has been done.

FIGURES

FIGURE 1 can be supplementary

We have chosen to keep this figure in the manuscript because we think that it well illustrates that immune genes are over-expressed in house mouse recently invaded sites compared to anciently invaded ones, and that this patterns is not found when comparing black rat sites.

Figure 2 should be removed, it is not of enough quality for a public document

We agree and Figure 2 is now included as Fig S3

Figure 3 I really liked how this figure showed us that even tough there are trends at the average level between invasion histories, there are specific populations that have their own expression profiles for these genes. That is very interesting and will probably warrant more attention in the future.

We thank you for this positive comment on this figure (now Fig 2).

DISCUSSION

p.10 « our results suggested that variations of immune phenotypes were a less important strategy for *R. rattus* invasion success «

I don't see where in the ms is invasion success quantified? Could it be that rats are actually not as successful than if they had modulated their immune system? Or that they would be even less successful if they did, since most of the immune activity is up-regulation anyway (which suggest that the invasive mouse is fighting new pathogens)?

We are sorry that our sentence was not clear.

We were not suggesting that the black rat was less or more successful than the house mouse. We considered that invasion was successful for both species. Indeed, our longitudinal surveys confirm i) that native species were present before the black rat or house mouse introduction, and ii) that the native species are becoming rare/absent from human settlings following the settlement of the black rat or house mouse.

We neither did speculate about the fate of black rat invasion considering different levels of modulation of immune responses.

With regard to the black rat, we discussed the point that we did not observe any clear pattern of changes in immune related gene expression between anciently and recently invaded sites. This result suggested that variations of immune phenotypes (as reflected by splenic gene expression) were not likely to have driven the black rat invasion success (considering that invasion success relies on the rapid evolution of phenotypic traits in newly colonized areas and not on prior adaptation in the anciently invaded range).

We modified the sentence to make it clearer (P13-14).

p.11 « Although mouse and rat populations experienced reduced genetic diversity due to founder events, they may have developed adaptive responses to novel selection pressures through high levels of plasticity «

Are the authors suggesting that there is genetic accommodation? Please explain in more detail and propose what it means for invasive species in general. Also, this study is a correlation study, and we do not have relationship between the gene expression phenotype and fitness, such that wording should be modulated accordingly.

We agree that our study is only correlative. Therefore we can describe patterns and propose hypotheses for scenarios resulting in these patterns, but we can not speculate too much about the mechanisms involved. We therefore discussed the three potential scenario that could explain changes in gene expression during rodent invasion, namely stochasticity versus adaptation (plasticity or selection).

With regard to plasticity, we did not discuss the possibility of genetic accommodation (evolutionary shifts in gene expression plasticity), genetic assimilation (loss of plasticity and fixation of favoured traits) as we think that it would be too speculative considering our data.

We modified the paragraph to better highlight the scenario that we want to discuss (stochasticity versus adaptation through plasticity or selection). (P11).

p.12 « The up-regulation of all immune genes found to be differentially expressed in sites recently invaded by the house mouse strongly supported the assumption of an increased overall infection risk in recently invaded sites. « How do we know this? Please add relevant references

In the original manuscript, the lines that followed this sentence provided evidence of such potential increase of infection risk, through a brief description of our work and results on the pathogenic bacteria found in these rodent populations (Diagne et al. 2017).

We have slightly modified these lines to make it clearer (P12).

Reviews by anonymous reviewer, 2018-11-22 10:32

In the manuscript entitled: « Differential immune gene expression associated with contemporary range expansion of two invasive rodents in Senegal » the authors assess the gene expression patterns between anciently and recently established populations of two rodent species in Senegal. They hypothesized that invasion success may relies on immune phenotypic traits that would be advantageous in recently invaded sites. The authors indeed showed that the species M*us musculus domesticus* showed an over-expression of immune related genes (notably the complement activation pathway), in recently invaded sites compared to anciently invaded sites and likely related to novel parasite pressures encountered in recently invaded sites. Regarding the species *Rattus rattus* the authors suggest that some stochastic events may be associated with colonization history since no particular pattern of differential gene expression related to immunity were found.

First, I would like to congratulate the authors for this very interesting work. I found the objectives clear and concise notably with a well written introduction. The analyses are various with a good use of the replicates and seem robust with a deep investigation regarding the biological processes involved. I nevertheless think that the manuscript could be improved notably regarding the structure (see comments below).

We thank you for these positive comments. We have tried to address the following comments. Modifications in the manuscript are reported in green.

Major comments:

1) Even if I found the objectives very clear, I think that the manuscript could still **explicit some information earlier** in order to gain in clarity. Notably, the whole study focuses on one tissue (the **spleen**) to analyze gene expression patterns and this information is given a bit late (in the discussion section). However, since the spleen is an immune related tissue I think that this could be explicitly mentioned in the introduction and related to the mains hypothesis investigated, since studying other tissue would test other hypothesis (such as using brain to study behavioural related genes as mentioned page 14).

We agree with this comment and we have now included this information in the introduction (P5).

2) Two approaches are used to test for differentially expressed genes (4x4 and 8x8), however, the subsequent functional analysis mainly focus on the genes identified with the 8x8 approach. I think that the 4x4 approach could be better justified regarding the main questions addressed or better linked to the others results. For instance, are some of the differentially expressed genes of the 4x4 approach related to immune processes?

We detailed and argued below, in response to the previous reviewers, the design of the two statistical analyses (4vs4 and 8vs8 approaches) performed. Because there is a high variability between samples, we explained that combining the results of these two approaches seemed to be a prudent choice, enabling the results to be as robust as possible. Thus, the genes identified with the '8vs8 approach' and declared as DE in more than 85% of the 256 combinations of the '4vs4 approach' were declared as DE. The threshold of 85% were set according to the visualisation of a barplot representing the number of '4vs4' comparisons in which a gene (highlighted in the '8vs8 approach') is found to be differentially expressed.

Keeping only the 4vs4 approach would be far too conservative considering the high variability between replicates. Among the 18 genes found to be DE from this approach, none was directly related to immune processes.

Note that keeping only the results of the 8vs8 approach did only marginally change the results in comparison to what is presented in our manuscript. Briefly, the analyses highlighted 'acute inflammatory response' as one of the main enriched biological process, and the Complement cascade as one of the main enriched pathway for the house mouse. The interaction network was also significant with an important node corresponding to complement cascade proteins. With regard to the black rat, considering only the 8vs8 approach highlighted 38 enriched GO. The Revigo analysis few emphasized biological processes related with metabolism only, and no pathway was detected as significantly enriched.

As we think that these results based on the '8vs8 approach' are less reliable because they could include false positive, we did not include them in the new version of the manuscript.

3) Discussion is a bit redundant (the mention of stochastics events involved in the rat invasion history for instance are discussed at the end of the first paragraph, in the second paragraph and in the fifth paragraph) and could beneficiate from a better structure, notably by addressing the question mentioned in the end of the introduction more directly. For instance in the introduction addressing whether the EICA or the EICA refined hypothesis are supported comes as the last question but is discussed quite early in the discussion (before the functional categories involved that is the second question presented in introduction).

We have rewritten the discussion to avoid redundancies and follow the questions addressed in the introduction.

Minor comments

4) Page 6, the number of transcriptome libraries is given, but I think that giving at this stage the number of studied sites and replicates would also ease the comprehension of the subsequent analysis.

This point was also noticed by N. Aubin-Horth. Information has been added at the beginning of the Results section (P6).

5) Page 18, it is mentioned that genes with less than 20 or 40 reads were discarded. I suppose that it echoes the "10 occurrences" filter mentioned page 7. If so, I suggest to homogenize and explicit this in order to avoid confusion.

You are right, we discarded all genes with less than 10 occurrences (*i.e.,* 20 reads for the 4 vs 4 strategy and 40 reads for the 8 vs 8 strategy, as we cumulate all the 8 or 16 analysed samples). We homogenized the text (10 occurrences) to avoid potential confusion for the reader (P8).

6) Page 6, the authors mentioned that they used a PCA on standardized read counts. I may be wrong, but I think that a **multidimensional scaling analysis (MDS) would fit better over dispersed count data** (see Bankers et al. 2017, fig. 3 for some example on transcriptomic data).

Bankers, L. et al. 2017. Genomic evidence for population-specific responses to co-evolving parasites in a New Zealand freshwater snail. - Mol. Ecol. 26: 3663–3675.

We now have included a MDS analysis based on a log fold change distance, following the edgeR package in R and the plotMDS function (P7 and Fig S1).

7) Fig. S1, the **PCA** is difficult to read, take into consideration to increase front size for instance.

Fig S1 has been changed

8) Fig. S3 and S4, check carefully the supplementary there is **some mismatch between the figure caption for Fig. S3 and S4 and the actual figures displayed**.

We have modified the figure caption in the main text.

Typography comments:

- A coma just be added just after "e.g." and "i.e.".

done

- Page 13, remove the double "x" in exhibit.

Done

Simon Blanchet

P4: Please elaborate a bit further on this example for readers that are unfamiliar with this literature

We have provided details to explain what genome scans are (P4).

P5: provide a better description of the method (e.g. "a whole RNA sequencing (i.e. RNAseq)"

Done (P5)

P6: the way you've done this PCA is unclear for me. Although explained in the Methods, I was expecting vectors to be genes, not sites. My own feeling is that a discriminant analysis (eg NMDS, DAPC) would have been better suited.

As this modification was also recommended by a reviewer, we have now included a MDS analysis, based on log fold change distance, instead of the PCA previously performed (P7 and Fig S1).

P17, description of the statistical approaches. This is an important section of the MS and I think you should provide more details or explain more clearly what has been done exactly.

This section is now more detailed in the revised manuscript (P7-8 & P18-19).