THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

Transcriptional analysis of insect extreme freeze tolerance

Lauren E. Des Marteaux, Petr Hůla and Vladimír Koštál

Article citation details
Proc. R. Soc. B **286**: 20192019.
http://dx.doi.org/10.1098/rspb.2019.2019

Review timeline

Original submission:	30 May 2019
1st revised submission:	30 August 2019
2nd revised submission:	30 September 2019
Final acceptance:	1 October 2019

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

Review History

RSPB-2019-1274.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

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

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

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.
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
Summary/General Thoughts:

In this paper, Des Marteaux et al. characterize transcriptional responses to diapause and cold acclimation in a freeze-tolerant fly, focusing on time points with distinct freeze-tolerance phenotypes. The data show strong evidence that upregulation of protein processing genes coincides with the highest levels of freeze tolerance, which suggests this pathway is involved in protection against freezing. The paper is generally well-written, and the figures are nicely presented. The results are somewhat preliminary and speculative, but given the unique system and potential application of these results, I think the paper is of sufficient general interest for Proc B. My main comments regard minor issues with interpretation and some missed opportunities to synthesize the results with the larger body of freeze tolerance literature.

Major Comments:

1. Overall the experimental design is good, but I was perplexed by the lack of a freeze-intolerant group in the RNA-seq analysis. The LD group was included in the qPCR validation, but inclusion of this group in the RNA-seq experiment would have potentially strengthened the authors' ability to globally identify genes specific to freeze-tolerant larvae. A stated goal of the paper is to enhance "fundamental understanding of its extreme freeze tolerance," and in my opinion inclusion of a non-freeze tolerant group increases your ability to do so, especially if the ultimate goal is to identify proteins that function as cryoprotectants (i.e., lines 86-90). I don't think this issue is a fatal flaw, partly because you still have variation in freeze tolerance among your SD and SDA groups, and SD is an appropriate "control group" for the SDA group. But perhaps this minor flaw should be addressed in the discussion. For example, how does your interpretation of the hsp expression results change with the knowledge that several of the hsps were expressed more highly in LD larvae (Fig. S10)? This issue was partially addressed in lines 328-334. However, the difference in freeze tolerance between LD and other groups is much more dramatic than the differences between SD and SDA.

2. There was a missed opportunity to synthesize the current results with previous studies of freeze tolerance. In particular, Toxopeus and Sinclair published a recent review article that summarizes the current physiological knowledge of freeze tolerance, and it would be good to

critically assess your results in the framework laid out by that review. For example, to what extent do your results agree with the current understanding of freeze tolerance, and what new insights does your research provide? Furthermore, Toxopeus also published a recent transcriptome for a freeze-tolerant cricket, so it would be good to at least superficially compare these datasets. A transcriptome for a freeze-tolerant worm was also recently published, allowing for some cross-phyla comparisons (de Boer et al., 2018, Ecology and Evolution 8, 3774-3786). These comparisons don't need to be extensive, but a paragraph in the Discussion would be warranted, perhaps at the expense of some of the molecular details of protein processing (which are somewhat speculative and mainly based on mammalian literature). I also realize your work has somewhat different goals (i.e., cryopreservation vs. ecologically relevant freeze tolerance), but I still think a synthesis of the freeze-tolerance literature could strengthen the paper.

Minor Comments:

1. Line 13: "exploit" is a strange word choice here, perhaps reword?
2. Line 24-27: This sentence is kind of vague, I suggest ending the Abstract with a stronger concluding sentence
3. Line 38: Saying cryoprotective dehydration results in loss of "most" of body water is somewhat misleading, because that makes it seem like >50% is lost. That may be true in some species, but water is lost until the melting point of hemolymph matches that of surrounding ice, so it may not be "most" of the water
4. Line 40: insert "of" before "insects"
5. Line 56: Ragland and Keep's comprehensive metanalysis of diapause would be a good paper to cite here, and you could also revisit this paper in the Discussion to address the extent to which Chymomyza diapause is similar/different transcriptionally to other diapauses (although I realize diapause per se isn't the primary focus of your paper) Ragland, G. J. and Keep, E. (2017). Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta. Physiological Entomology 42, 246-256.
6. Line 74: I'm not sure a discussion of LEA proteins is warranted, because they seem to be rare in insects. To my knowledge, only one insect (Polypedilum vanderplanki) has LEAs, although they do function in lots of other desiccation/freeze tolerant taxa
7. Line 137-141: These numbers (19% of full transcriptome, 6,800 remaining contigs for gene expression analysis) seem low, especially since the previous transcriptome was also built with larvae. I don't have a lot of experience aligning reads to de novo transcriptomes, but do these results suggest that perhaps you should reassemble your transcriptome with the new reads? What percent of your reads aligned to the previous trnscriptome? It seems like you might be missing a lot.
8. Line 145-148: Perhaps briefly justify your choice of a fold change cutoff. I'm personally against fold change cutoffs, because the biological consequences of "fold change" depend a lot on the particular gene. For example, small fold changes of metabolic or regulatory genes could have big effects on physiology due.
9. Line 164-173: A bit more information (either here or in the Supplemental Information) is needed to evaluate the KEGG analysis. What data went into these KEGG analyses, and what reference set of genes was used to test for enrichment analysis?
10. Line 227: "As may be expected" is subjective, so I suggest removing this phrase
11. Line 254: The statement "unlikely to be biologically relevant" is also subjective
12. Line 269: The "anticipatory" hypothesis is supported by previous diapause transcriptome studies that show photoperiodic induction of heat shock proteins even in the absence of temperature change (e.g., Ragland, G. J., Denlinger, D. L. and Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. Proceedings of the National Academy of Sciences of the United States of America 107, 14909-14914.)
13. Line 287-289: This sentence is overly speculative, because many of these genes are kinases, so they are likely post-translationally regulated.

14. I'm not sure how to articulate this, but I think you may be on to something by paying attention to the species-specific, unannotated genes. Pretty much every species that diapauses or undergoes cold acclimation upregulates heat shock proteins, but none of them other than Chymomyza survive cryopreservation. So that means Chymomyza either has "better" heat shock proteins, or perhaps they are not really an important piece of the puzzle. I have no doubt they're important for deep freezing, but I think the Discussion might be better served by focusing on the new findings. In other words, most of the Discussion focuses on heat shock protein/proteasome/ER, which we already know is important for cold hardiness. So I would perhaps shift the emphasis off heat shock proteins a little, even though the results are compelling, because heat shock proteins seem to be involved in every insect stress response and likely don't explain why Chymomyza can survive in liquid N2 (unless you have reason to believe they have super heat shock proteins).

Review form: Reviewer 2 (Craig Marshall)

Recommendation
Accept as is

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?
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?
No

Comments to the Author
This is a very nice paper indeed that was a pleasure to assess and there is very little that needs comment.

I particularly like the use of four biological replicate populations for the validation of expression by qPCR. It would also be useful for a brief description of how the four reference genes were chosen for qPCR. This is referenced, but the details are buried there and there is little discussion as to why these genes were chosen over others. They are all ribosomal protein genes and I wonder if their expression is entirely independent.

The assignment of only 5% of the contigs in the full transcriptome to KEGG pathways seems low and has the potential to skew the assessment of gene differential expression. It would be helpful to include a little discussion as to what might underlie this low rate of assignment given that the proportion of transcripts matched to homologues was much higher. Does this reflect an absence of insect data in KEGG or an abundance of novel differentially-expressed sequences in C. costata?

In particular, I wonder what was seen in pathways likely to lead to proline synthesis (and degradation) given proline's role as a compatible osmolyte and in cryoprotection in C. crostata. One limitation in using differentially-expressed sequences in these sorts of studies is that sequences constitutively expressed in large amounts might be missed: perhaps sequences associated with proline metabolism might fall into this category and it might be worth looking at the abundance of these sequences even if they are not differentially-expressed.

The finding of a large number of differentially-expressed but unrecognised sequences is intriguing and I would like to know more about these. Although some basic blast-based searches have been used to try and find homologues, I wonder if searches for open reading frames were considered along with a search of these against protein and motif databases. It is reasonable that this work might form the basis of another paper.

I wonder also if there is room for discussion of ways in which interesting transcripts might be tested for their role in freezing tolerance. C. costata is not a model organism and techniques such as RNAi are not readily available, but it would be interesting to know how key candidates for freezing tolerance could be investigated further even if such work might not comprise part of this paper.

Review form: Reviewer 3

Recommendation
Accept with minor revision (please list in comments)

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?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
Yes

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?
No

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
Please see attached file. Thank you for the interesting paper! (See Appendix A)

Decision letter (RSPB-2019-1274.R0)

12-Jul-2019

Dear Dr Des Marteaux:

I am writing to inform you that your manuscript RSPB-2019-1274 entitled "Transcriptional analysis of insect extreme freeze tolerance" 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 important 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.

Sincerely,

Professor Hans Heesterbeek
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
The manuscript has been assessed by three expert reviewers and myself, and we all find it interesting, technically very well done, and of importance to the field of freeze tolerance. There are several important points that have been raised by the reviewers that should be attended to carefully in revision. I will highlight a few here that I consider critical, but please consider carefully all points raised by the reviewers and comment on your decisions. Please incorporate the literature on freeze-tolerance suggested by Reviewer 1, and comment on the potential role of HSPs given that many insects express them but are not extremely freeze tolerant. Please carefully respond to Reviewer 2's query about the low percentage of transcripts mapped to KEGG pathways, and Reviewer 1s comment about the generally low number of transcripts/ quality of transcriptome. As suggested by Reviewer 3, please clarify early in the manuscript that the transcriptional changes occur in response to long-term cold acclimation and not exposure to liquid nitrogen, and provide a rationale for studying the transcriptional regulation. It is not necessary to perform protein quantitation on the key candidate proteins that are overexpressed, but please remind the readers during the discussion that transcript abundance does not necessarily map to protein abundance. You might consider pointing out any key conclusions that rely on that assumption. I would be happy to consider a revised manuscript for publication in Proc B.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
Summary/General Thoughts:

In this paper, Des Marteaux et al. characterize transcriptional responses to diapause and cold acclimation in a freeze-tolerant fly, focusing on time points with distinct freeze-tolerance phenotypes. The data show strong evidence that upregulation of protein processing genes

coincides with the highest levels of freeze tolerance, which suggests this pathway is involved in protection against freezing. The paper is generally well-written, and the figures are nicely presented. The results are somewhat preliminary and speculative, but given the unique system and potential application of these results, I think the paper is of sufficient general interest for Proc B. My main comments regard minor issues with interpretation and some missed opportunities to synthesize the results with the larger body of freeze tolerance literature.

Major Comments:

1. Overall the experimental design is good, but I was perplexed by the lack of a freeze-intolerant group in the RNA-seq analysis. The LD group was included in the qPCR validation, but inclusion of this group in the RNA-seq experiment would have potentially strengthened the authors' ability to globally identify genes specific to freeze-tolerant larvae. A stated goal of the paper is to enhance "fundamental understanding of its extreme freeze tolerance," and in my opinion inclusion of a non-freeze tolerant group increases your ability to do so, especially if the ultimate goal is to identify proteins that function as cryoprotectants (i.e., lines 86-90). I don't think this issue is a fatal flaw, partly because you still have variation in freeze tolerance among your SD and SDA groups, and SD is an appropriate "control group" for the SDA group. But perhaps this minor flaw should be addressed in the discussion. For example, how does your interpretation of the hsp expression results change with the knowledge that several of the hsps were expressed more highly in LD larvae (Fig. S10)? This issue was partially addressed in lines 328-334. However, the difference in freeze tolerance between LD and other groups is much more dramatic than the differences between SD and SDA.

2. There was a missed opportunity to synthesize the current results with previous studies of freeze tolerance. In particular, Toxopeus and Sinclair published a recent review article that summarizes the current physiological knowledge of freeze tolerance, and it would be good to critically assess your results in the framework laid out by that review. For example, to what extent do your results agree with the current understanding of freeze tolerance, and what new insights does your research provide? Furthermore, Toxopeus also published a recent transcriptome for a freeze-tolerant cricket, so it would be good to at least superficially compare these datasets. A transcriptome for a freeze-tolerant worm was also recently published, allowing for some cross-phyla comparisons (de Boer et al., 2018, Ecology and Evolution 8, 3774-3786). These comparisons don't need to be extensive, but a paragraph in the Discussion would be warranted, perhaps at the expense of some of the molecular details of protein processing (which are somewhat speculative and mainly based on mammalian literature). I also realize your work has somewhat different goals (i.e., cryopreservation vs. ecologically relevant freeze tolerance), but I still think a synthesis of the freeze-tolerance literature could strengthen the paper.

Minor Comments:

1. Line 13: "exploit" is a strange word choice here, perhaps reword?
2. Line 24-27: This sentence is kind of vague, I suggest ending the Abstract with a stronger concluding sentence
3. Line 38: Saying cryoprotective dehydration results in loss of "most" of body water is somewhat misleading, because that makes it seem like >50% is lost. That may be true in some species, but water is lost until the melting point of hemolymph matches that of surrounding ice, so it may not be "most" of the water
4. Line 40: insert "of" before "insects"
5. Line 56: Ragland and Keep's comprehensive metanalysis of diapause would be a good paper to cite here, and you could also revisit this paper in the Discussion to address the extent to which Chymomyza diapause is similar/different transcriptionally to other diapauses (although I realize diapause per se isn't the primary focus of your paper) Ragland, G. J. and Keep, E. (2017).

Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta. Physiological Entomology 42, 246-256.
6. Line 74: I'm not sure a discussion of LEA proteins is warranted, because they seem to be rare in insects. To my knowledge, only one insect (Polypedilum vanderplanki) has LEAs, although they do function in lots of other desiccation/freeze tolerant taxa
7. Line 137-141: These numbers (19% of full transcriptome, 6,800 remaining contigs for gene expression analysis) seem low, especially since the previous transcriptome was also built with larvae. I don't have a lot of experience aligning reads to de novo transcriptomes, but do these results suggest that perhaps you should reassemble your transcriptome with the new reads? What percent of your reads aligned to the previous trnscriptome? It seems like you might be missing a lot.
8. Line 145-148: Perhaps briefly justify your choice of a fold change cutoff. I'm personally against fold change cutoffs, because the biological consequences of "fold change" depend a lot on the particular gene. For example, small fold changes of metabolic or regulatory genes could have big effects on physiology due.
9. Line 164-173: A bit more information (either here or in the Supplemental Information) is needed to evaluate the KEGG analysis. What data went into these KEGG analyses, and what reference set of genes was used to test for enrichment analysis?
10. Line 227: "As may be expected" is subjective, so I suggest removing this phrase
11. Line 254: The statement "unlikely to be biologically relevant" is also subjective
12. Line 269: The "anticipatory" hypothesis is supported by previous diapause transcriptome studies that show photoperiodic induction of heat shock proteins even in the absence of temperature change (e.g., Ragland, G. J., Denlinger, D. L. and Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. Proceedings of the National Academy of Sciences of the United States of America 107, 14909-14914.)
13. Line 287-289: This sentence is overly speculative, because many of these genes are kinases, so they are likely post-translationally regulated.
14. I'm not sure how to articulate this, but I think you may be on to something by paying attention to the species-specific, unannotated genes. Pretty much every species that diapauses or undergoes cold acclimation upregulates heat shock proteins, but none of them other than Chymomyza survive cryopreservation. So that means Chymomyza either has "better" heat shock proteins, or perhaps they are not really an important piece of the puzzle. I have no doubt they're important for deep freezing, but I think the Discussion might be better served by focusing on the new findings. In other words, most of the Discussion focuses on heat shock protein/proteasome/ER, which we already know is important for cold hardiness. So I would perhaps shift the emphasis off heat shock proteins a little, even though the results are compelling, because heat shock proteins seem to be involved in every insect stress response and likely don't explain why Chymomyza can survive in liquid N2 (unless you have reason to believe they have super heat shock proteins).

Referee: 2

Comments to the Author(s)
This is a very nice paper indeed that was a pleasure to assess and there is very little that needs comment.

I particularly like the use of four biological replicate populations for the validation of expression by qPCR. It would also be useful for a brief description of how the four reference genes were chosen for qPCR. This is referenced, but the details are buried there and there is little discussion as to why these genes were chosen over others. They are all ribosomal protein genes and I wonder if their expression is entirely independent.

The assignment of only 5% of the contigs in the full transcriptome to KEGG pathways seems low and has the potential to skew the assessment of gene differential expression. It would be helpful to include a little discussion as to what might underlie this low rate of assignment given that the proportion of transcripts matched to homologues was much higher. Does this reflect an absence of insect data in KEGG or an abundance of novel differentially-expressed sequences in C. costata?

In particular, I wonder what was seen in pathways likely to lead to proline synthesis (and degradation) given proline's role as a compatible osmolyte and in cryoprotection in C. crostata. One limitation in using differentially-expressed sequences in these sorts of studies is that sequences constitutively expressed in large amounts might be missed: perhaps sequences associated with proline metabolism might fall into this category and it might be worth looking at the abundance of these sequences even if they are not differentially-expressed.

The finding of a large number of differentially-expressed but unrecognised sequences is intriguing and I would like to know more about these. Although some basic blast-based searches have been used to try and find homologues, I wonder if searches for open reading frames were considered along with a search of these against protein and motif databases. It is reasonable that this work might form the basis of another paper.

I wonder also if there is room for discussion of ways in which interesting transcripts might be tested for their role in freezing tolerance. C. costata is not a model organism and techniques such as RNAi are not readily available, but it would be interesting to know how key candidates for freezing tolerance could be investigated further even if such work might not comprise part of this paper.

Referee: 3

Comments to the Author(s)
Please see attached file. Thank you for the interesting paper!

Author's Response to Decision Letter for (RSPB-2019-1274.R0)

See Appendix B.

RSPB-2019-2019.R0

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

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

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

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.
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
I think the authors have made some important improvements to their manuscript. The rationale for the study is improved, and appropriate caveats for RNA-seq are now included. I am mostly satisfied with the authors' responses to my previous review, although there are a couple of cases I thought they could have done a better job. See below:
1. In my first review, I indicated that I thought the authors could have done a better job of integrating other physiological/transcriptional studies from the literature. The authors responded that because they are focused on "truly extreme" freeze tolerance, other studies are not relevant. I see their point, but I still think there is room to include these studies. Comparisons to other moderately freeze tolerant organisms, and the authors' goals identifying mechanisms of extreme freeze tolerance, are not mutually exclusive endeavors. Chymomyza does freezing better than any other insect that's been studied, so that must either come about by completely new mechanisms, by augmentation of existing mechanisms, or some of both. If you compare the classes of genes identified with extreme freeze tolerance to those involved in other types of freeze tolerance, you could start to address this question and provide more insight into the nature of your phenotype. If word limits are an issue, you probably could cut down the number of references without compromising the paper. I appreciate the thoroughness of your literature search, but in some cases multiple references are cited when 1-2 key references would suffice.

2. I was concerned that only 19% of the transcriptome was represented in the current study, and the authors pointed out that diapause causes a general repression of transcription. While this may be true, to allay my concerns it would be nice to see some read mapping statistics, which I had asked for in my previous review. The percentage of mapped reads is an important indicator of quality for an RNA-seq experiment and is frequently reported (see Conesa et al., 2016, Genome Biology 17:13). If you do indeed have good read mapping percentage (typically above 70% for the studies I've looked at), then I would be satisfied with your response. However, if your mapping statistics are poor, it could indicate the presence of new unassembled transcripts or technical issues.

Additional Minor Comments:

1. Line 97: I also think the sensitivity of RNA-seq is attractive for a study like this. Proteomics and metabolomics are typically 1-2 orders of magnitude less sensitive (at least in terms of the number of resolved features) than RNA-seq. May be less of concern for you if you're interested in high-quantity cryoprotectants, but the ability to measure true genome-wide expression is a nice feature of RNA-seq.

2. Line 108-109: Suggest changing to "…on artificial diet in MIR-154 incubators (Sanyo Electric, Osaka, Japan) as described previously [57]. For experiments, three different malt fly larval phenotypic variants were generated according to…" As written it makes it sound like the incubators are generating the phenotypes, when it's the temperature/light regimes that are more important.

3. Line 215: This is more of a personal qualm, but I don't like it when transcriptomics studies are referred to as "hypothesis-generating" studies. I think it undersells the value of transcriptomics and contributes to the narrative that genomics is a "fishing expedition." That may be true in some cases, but if you are testing clear mechanistic hypotheses and the RNA-seq experiment is properly designed and replicated, it can do more than generate hypotheses. But ultimately this is your work, so it's your decision how you want to frame the study.

4. Line 229: Change "that" to "those"

5. Line 242: Change "were" to "was"

6. Figure 3 and select Supplemental Figures: In heat maps and similar figures, it's advisable to avoid red/green color scales, to make the figures accessible to color blind readers. If it's easy to customize, suggest going with a blue/yellow scale or another palette that is discernible by color blind individuals.

Review form: Reviewer 2 (Craig Marshall)

Recommendation
Accept with minor revision (please list 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?
Excellent

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?
No

Comments to the Author
I'm very happy with the responses to my questions in the previous review and I see almost nothing that new needs attention. This is a nice paper and I look forward to seeing the other analyses mentioned.
There are two things I would like to see changed: 'comprised of' to 'comprising' at L118 and L185, along with the replacement of 'x' for '×', the multiplication sign at L150.

Decision letter (RSPB-2019-2019.R0)

25-Sep-2019

Dear Dr Des Marteaux:

Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewers and the Associate Editor have raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address

all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please 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.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (https://royalsociety.org/journals/ethics-policies/). You should pay particular attention to the following:

Research ethics:
If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:
If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:
It is a condition of publication that you make available the data and research materials supporting the results in the article. Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article (https://royalsociety.org/journals/ethics-policies/data-sharing-mining/). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

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=(Document not available), 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.

For more information please see our open data policy http://royalsocietypublishing.org/data-sharing.

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. Please try to submit all supplementary material as a single file.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Professor Hans Heesterbeek
mailto: proceedingsb@royalsociety.org

Associate Editor Board Member
Comments to Author:
The authors have addressed most of the concerns of the reviewers. The main remaining concern is regarding the transcriptome coverage. Please provide read mapping statistics as requested by Reviewer 1. The other suggestions provided by both reviewers should also be considered.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
I think the authors have made some important improvements to their manuscript. The rationale for the study is improved, and appropriate caveats for RNA-seq are now included. I am mostly satisfied with the authors' responses to my previous review, although there are a couple of cases I thought they could have done a better job. See below:
1. In my first review, I indicated that I thought the authors could have done a better job of integrating other physiological/transcriptional studies from the literature. The authors responded that because they are focused on "truly extreme" freeze tolerance, other studies are not relevant. I see their point, but I still think there is room to include these studies. Comparisons to other moderately freeze tolerant organisms, and the authors' goals identifying mechanisms of extreme freeze tolerance, are not mutually exclusive endeavors. Chymomyza does freezing better than any other insect that's been studied, so that must either come about by completely new mechanisms, by augmentation of existing mechanisms, or some of both. If you compare the classes of genes identified with extreme freeze tolerance to those involved in other types of freeze

tolerance, you could start to address this question and provide more insight into the nature of your phenotype. If word limits are an issue, you probably could cut down the number of references without compromising the paper. I appreciate the thoroughness of your literature search, but in some cases multiple references are cited when 1-2 key references would suffice.

2. I was concerned that only 19% of the transcriptome was represented in the current study, and the authors pointed out that diapause causes a general repression of transcription. While this may be true, to allay my concerns it would be nice to see some read mapping statistics, which I had asked for in my previous review. The percentage of mapped reads is an important indicator of quality for an RNA-seq experiment and is frequently reported (see Conesa et al., 2016, Genome Biology 17:13). If you do indeed have good read mapping percentage (typically above 70% for the studies I've looked at), then I would be satisfied with your response. However, if your mapping statistics are poor, it could indicate the presence of new unassembled transcripts or technical issues.

Additional Minor Comments:

1. Line 97: I also think the sensitivity of RNA-seq is attractive for a study like this. Proteomics and metabolomics are typically 1-2 orders of magnitude less sensitive (at least in terms of the number of resolved features) than RNA-seq. May be less of concern for you if you're interested in high-quantity cryoprotectants, but the ability to measure true genome-wide expression is a nice feature of RNA-seq.

2. Line 108-109: Suggest changing to "…on artificial diet in MIR-154 incubators (Sanyo Electric, Osaka, Japan) as described previously [57]. For experiments, three different malt fly larval phenotypic variants were generated according to…" As written it makes it sound like the incubators are generating the phenotypes, when it's the temperature/light regimes that are more important.

3. Line 215: This is more of a personal qualm, but I don't like it when transcriptomics studies are referred to as "hypothesis-generating" studies. I think it undersells the value of transcriptomics and contributes to the narrative that genomics is a "fishing expedition." That may be true in some cases, but if you are testing clear mechanistic hypotheses and the RNA-seq experiment is properly designed and replicated, it can do more than generate hypotheses. But ultimately this is your work, so it's your decision how you want to frame the study.

4. Line 229: Change "that" to "those"

5. Line 242: Change "were" to "was"

6. Figure 3 and select Supplemental Figures: In heat maps and similar figures, it's advisable to avoid red/green color scales, to make the figures accessible to color blind readers. If it's easy to customize, suggest going with a blue/yellow scale or another palette that is discernible by color blind individuals.

Referee: 2

Comments to the Author(s).
I'm very happy with the responses to my questions in the previous review and I see almost nothing that new needs attention. This is a nice paper and I look forward to seeing the other analyses mentioned.

There are two things I would like to see changed: 'comprised of' to 'comprising' at L118 and L185, along with the replacement of 'x' for '×', the multiplication sign at L150.

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

See Appendix C.

Decision letter (RSPB-2019-2019.R1)

01-Oct-2019

Dear Dr Des Marteaux

I am pleased to inform you that your manuscript entitled "Transcriptional analysis of insect extreme freeze tolerance" 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

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.

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

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). 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.

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,

Professor Hans Heesterbeek
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Associate Editor:
Board Member
Comments to Author:
The authors have made all changes suggested by the reviewers, and the manuscript is now suitable for publication.

Appendix A

General comments:

This paper addresses a fascinating topic—what is the biochemical response to extreme freezing stress—by taking a transcriptomics approach. The authors are rigorous in their number of biological samples for RNAseq and followup with additional qPCR studies and timepoints. The authors do state early in the introduction that indeed mRNA transcript abundance does not necessarily correlate with protein abundance or physiological process, which I greatly appreciate and is a crucially important distinction (especially in the case of extreme freezing). Overall, this large transcriptomics dataset on various acclimations treatments of a freeze-tolerant insect species a unique contribution to the study of "extreme freeze tolerance" and I find their title and scope to be both engaging and appropriate.

One longstanding point of confusion is, however, that the authors state that a main goal in their study is to identify candidate proteins involved in extreme freeze tolerance (some proteins of which are speculated to be CPs, which could be non-protein cryoprotectants), yet the authors take a purely transcriptomics approach. It may be useful to justify to the reader why a transcriptomics approach was the preferred first-step, rather than proteomics or metabolomics (either independently, or in synergy with transcriptomics). Although I do not feel these experiments would be strictly necessary for the publication of this manuscript, I am curious if the authors could include western blots of at least some of the candidate proteins (such as the HSPs), which may have antibodies known to work in several insect species. Such additional data would better link the author's transcriptomics findings with their goal to identify candidate *proteins.*

As an additional point of confusion: Does any work in the literature suggest that transcription or translation is even biochemically or biophysically possible under instances of extreme freezing (liquid nitrogen)? If so, please cite. If not, what is the biological meaning of DE transcripts (if it unlikely possible for them to be translated in response to extreme and sudden freezing)? I remain somewhat skeptical that *de novo* transcription or translation is possible under such harsh physical conditions.

Otherwise, my main suggestions for this manuscript are calls for clarifications/stylistic changes, as well as providing more context and interpretation in the introduction and discussion/conclusions on several topics.

In the Introduction:

-Would it be possible to please give more context as to why the model species were chosen, or any additional information surrounding the species' habitat or natural history? Is it surprising that this species can survive such extreme cold? Do they experience such temperatures in nature (likely no)? If no, why do you predict this species can survive freezing in liquid nitrogen (is there cross-tolerance with another environmental stress)?

-The introduction might be a good place to both address why a transcriptomics approach was chosen and what evidence you might have from the literature to believe that de novo transcription is possible under conditions of extreme freeze stress.

In the Discussion:
-The authors describe using the Oligo(dT)23 anchored primers to enrich for mRNAs/RNAs with poly(A) tails, in their transcriptomics work. These primers are standard, but a consequence is that the transcriptomics does not identify or consider non-polyadenylated RNAs (some subset of regulatory RNAs). Non-polyadenylated RNAs—such as mature microRNAs—may play important roles in protein processing and gene expression (most often by blocking protein translation). Since the authors are very interested in the role of protein processing in extreme freeze tolerance, I feel this experimental limitation should be mentioned/considered. Sequencing non-polyadenylated RNAs would also be an interesting next step, in future work!
-It would be helpful to consider whether transcriptional regulation vs. other forms of gene regulation beyond transcription, in your interpretation in changes in transcript abundance or predicted physiological processes.

This was a pleasure to read; thank you for your hard work!

Specific comments:
Line 19: The phrase "genes and processes that accompany the physiological transition to extreme freeze tolerance" may be too broad. Perhaps "genes and processes" could be more accurately described as "transcripts."

Line 46: It may be helpful to add one additional sentence of context as to why "These extremely freeze-tolerant animals may hold the key to advancing techniques in cryopreservation of tissues, organs, or even complex organisms [15]." What would one hope to identify from especially freeze tolerant animals, to apply to medical applications?

Line 55: Purely stylistic comment, in the phrase "Both adaptive complexes - diapause and cold acclimation - are based on the massive" the dashes should be em dashes (—) so the phrase is rather "Both adaptive complexes—diapause and cold acclimation—are based on the massive"

Line 98: Question: Is it possible to justify in the text why you took a transcriptomics approach if you write your aim is to identify candidate proteins…" Why not a proteomics or metabolomics approach? It might be helpful to justify to the reader your experimental rationale.

Line 152: The E-value is very relaxed at E-1. A more "standard" E-value would be E-5. Please describe why such a relaxed E-value chosen.

Line 158-162: What is meant by "custom categories"? I personally felt that the "'Protein processing" category was too general/vague and could likely be split into at least 2 more specific categories, in terms of biological function. Likewise, I found it confusing to have both an "Other" and "Unknown" category (although I think I understand the intent).

Line 221/222: The word "true" in "true physiological mechanisms of freeze tolerance" is confusing. Have untrue mechanisms been proposed in the literature? Or what makes information collected from transcriptomics in particular more "true"? Perhaps delete this word to avoid confusion.

Line 229: It would be helpful to remind the reader how you define differential expression (what is the threshold & what is meant by up vs downregulated, and what a sample is DE in comparison to)

Line 233: "Only a single DE gene (Seq109825; downregulated but with no putative identity, i.e. not shown in Fig. 2) overlapped between late diapause and cold acclimation (relative to early diapause), suggesting that extreme freeze tolerance involves a unique set of physiological processes" I am unsure how supported the conclusion that "extreme freeze tolerance involves a unique set of physiological process" is; it is possible that 1) at least some core physiological processes are not controlled by de novo transcription under extreme freezing or 2) transcripts for important physiological processes to survive extreme freezing are constitutively expressed

Line 248: category 'Protein processing' is very general

Line 262: A suggestion regarding the phrase "it seems obvious that extreme freeze tolerance should largely rely on upregulated protein protection": remove the phrase "it seems obvious" with several citations of when process has been observed in other species or systems

Line 272: Regarding the phrase "Cold and freeze-dehydration stresses may hit both nascent and functional proteins" replace the word "hit" with what is predicted to occur to the proteins biochemically (more specifically)

Line 300-301: In the phrase "In extremely freeze-tolerant larvae, six genes encoding chaperones involved in protein stabilization and refolding were enhanced" what does the word enhanced mean? DE? Or something else? Please clarify in the text.

Line 334-336: Regarding "The transcriptional profiles of lethal(2)efl (Fig. S10B) and hsp70 (Fig. S10E) matched relatively well with extreme freeze tolerance and, moreover, showed relatively high ddCT values (suggesting that their protein products are abundantly expressed)." Could you please justify why a "relatively high ddCT" suggested abundant protein products? This may be a controversial claim if not justified.

Line 359-360: The phrase "We hope that this set of unidentified genes contains novel and potentially important cryoprotectants" seems perhaps overly speculatory. Reword?

Line 363: Can the methods of future functional verification be described in any additional detail?

Comments on Figures:

1. Figure 1 is a bit blurry in the proof. Is it possible to use a vector file?
2. Figure 2A - Perhaps put number of genes in each pie slice, to make visual more informative
3. Figure 3 - Although I love the organization of Figure 3, it might be important to emphasize that transcripts do not strictly result in translation of protein or physiological process. This figure may unintentionally convey that these proteins/processes were directly observed.
4. Figure 4A - What is meant by "color areas used to visually enhance?" Confused. Font size in all of Figure 4 is small.

Thank you again!

Appendix B

RESPONSE TO REFEREES
Transcriptional analysis of insect extreme freeze tolerance
RSPB-2019-1274 (Resubmission)

Associate Editor

The manuscript has been assessed by three expert reviewers and myself, and we all find it interesting, technically very well done, and of importance to the field of freeze tolerance. There are several important points that have been raised by the reviewers that should be attended to carefully in revision. I will highlight a few here that I consider critical, but please consider carefully all points raised by the reviewers and comment on your decisions. Please incorporate the literature on freeze-tolerance suggested by Reviewer 1, and comment on the potential role of HSPs given that many insects express them but are not extremely freeze tolerant. Please carefully respond to Reviewer 2's query about the low percentage of transcripts mapped to KEGG pathways, and Reviewer 1s comment about the generally low number of transcripts/ quality of transcriptome. As suggested by Reviewer 3, please clarify early in the manuscript that the transcriptional changes occur in response to long-term cold acclimation and not exposure to liquid nitrogen, and provide a rationale for studying the transcriptional regulation. It is not necessary to perform protein quantitation on the key candidate proteins that are overexpressed, but please remind the readers during the discussion that transcript abundance does not necessarily map to protein abundance. You might consider pointing out any key conclusions that rely on that assumption. I would be happy to consider a revised manuscript for publication in Proc B.

R: We thank you and the reviewers for their constructive suggestions and comments, which helped us to improve the manuscript. We have now addressed the reviewer comments (via clarification of the text and/or in this response document), and we hope you find the edits to be sufficient.
Briefly:
1) Regarding A) a low number of KEGG IDs, B) a low number of transcripts: A) Of the contigs with read counts in the present study, approximately 21% were assigned KEGG IDs. Previous wording related the number of IDs (5%) to the full transcriptome previously assembled (Poupardin et al., 2015, BMC Genomics 16, 720) - and this was confusing/not an appropriate comparison, thus it has been reworded. B) The proportion of transcripts captured in the present study (19% of the full transcriptome previously assembled) is expected; here we sequenced only diapausing larvae (which are known to have significantly reduced transcription), while the full assembly was based on both developing and diapausing larvae, sampled at multiple Zeitgebers.
2) We have explained to Reviewer 3 that larvae were not sampled from LN_2. We feel that our methods are entirely clear in this regard, both in the text and Figure 1.
3) We have provided rationale about our approach near the end of the introduction: that this transcriptomic comparison is part of a series (metabolomics manuscript in final stages of preparation, proteomics data in processing), and that it has shortcomings but also clear benefits at this stage of research.
4) We have decided to move 'mRNA abundance does not necessarily equate to protein abundance' to the introduction so that the statement is made clearly, early on. We then do not repeat it (as we are at the word limit, given the space required for our figures).

Referee 1

Summary/General Thoughts:

In this paper, Des Marteaux et al. characterize transcriptional responses to diapause and cold acclimation in a freeze-tolerant fly, focusing on time points with distinct freeze-tolerance phenotypes. The data show strong evidence that upregulation of protein processing genes coincides with the highest levels of freeze tolerance, which suggests this pathway is involved in protection against freezing. The paper is generally well-written, and the figures are nicely presented. The results are somewhat preliminary and speculative, but given the unique system and potential application of these results, I think the paper is of sufficient general interest for Proc B. My main comments regard minor issues with interpretation and some missed opportunities to synthesize the results with the larger body of freeze tolerance literature.

R: We are glad that the reviewer liked the manuscript and found our work to be of sufficient interest - thank you!

Major Comments:

1. Overall the experimental design is good, but I was perplexed by the lack of a freeze-intolerant group in the RNA-seq analysis. The LD group was included in the qPCR validation, but inclusion of this group in the RNA-seq experiment would have potentially strengthened the authors' ability to globally identify genes specific to freeze-tolerant larvae. A stated goal of the paper is to enhance "fundamental understanding of its extreme freeze tolerance," and in my opinion inclusion of a non-freeze tolerant group increases your ability to do so, especially if the ultimate goal is to identify proteins that function as cryoprotectants (i.e., lines 86-90). I don't think this issue is a fatal flaw, partly because you still have variation in freeze tolerance among your SD and SDA groups, and SD is an appropriate "control group" for the SDA group. But perhaps this minor flaw should be addressed in the discussion. For example, how does your interpretation of the hsp expression results change with the knowledge that several of the hsps were expressed more highly in LD larvae (Fig. S10)? This issue was partially addressed in lines 328-334. However, the difference in freeze tolerance between LD and other groups is much more dramatic than the differences between SD and SDA.

R: The transition from LD to SD (i.e. diapause entry) is associated with dramatic alteration of gene transcription (documented for *C. costata* and other insects as well). Leaving the LD larvae out in this study, we hoped to reduce the complexity of analyzing transcriptional changes and focus more on those changes that are linked to gradual building of extreme freeze tolerance phenotype. This phenotype is not yet developed in early diapausing larvae (SD3) and only moderate freeze tolerance develops during diapause maintenance at relatively high temperatures (SD treatment). Here we are thus specifically interested in genes that may be involved in the additional leap from moderate to extreme freeze tolerance associated with cold acclimation (i.e. SD vs SDA). Since both treatments—SD and SDA—share the same physiological background (diapause), we hoped to subtract the changes that occur during diapause maintenance from those which occur during cold acclimation, which is indeed our stated goal.

We are aware of the fact that many genes may play different roles in different physiological contexts (e.g. in diapause vs. direct development). In fact, relatively high expression of *hsp22* and *Lsp2* in LD larvae supports our choice to leave LD larvae aside for the RNAseq analysis; broadly comparing gene

expression in two distinctly different physiological contexts could be potentially confounding. Generally, we do agree with your comment about *hsp22*, and we tried to be cautious about the interpreted relevance of this gene while also not discounting it outright (for reasons discussed in the text and above). We have tried to better justify our interpretation (especially because mitochondria are highly susceptible to freezing in this species):
Lines 325-336: "Although expression of *hsp22* in non-diapause larvae (LD3) was also very high, the gene need not be necessarily removed from the list of candidate cryoprotective macromolecules. Hsp22 localizes to mitochondria [92] and may play very different roles in the two distinct physiological contexts: LD (direct development, rapid growth, high activity) vs. SDA (diapause, deep metabolic suppression, no activity). The rapid metabolism of LD larvae must be supported by high activity of their mitochondria, and (in *D. melanogaster*) *hsp22* expression is upregulated with aging and oxidative stress [93] as well as in response to rising ecdysone titres just prior to pupation [94, 95]. These stimuli for *hsp22* upregulation are absent in the *C. costata* SDA phenotype. We also know that *C. costata* mitochondria are particularly susceptible to cryoinjury in LD larvae, but are robust to freezing challenges in SDA larvae (unpublished results, details in a forthcoming paper). Potentially, Hsp22 could act (synergistically with other cryoprotectants and DE chaperones) in SDA larvae to preserve mitochondrial structure during freeze dehydration insult."

2. There was a missed opportunity to synthesize the current results with previous studies of freeze tolerance. In particular, Toxopeus and Sinclair published a recent review article that summarizes the current physiological knowledge of freeze tolerance, and it would be good to critically assess your results in the framework laid out by that review. For example, to what extent do your results agree with the current understanding of freeze tolerance, and what new insights does your research provide? Furthermore, Toxopeus also published a recent transcriptome for a freeze-tolerant cricket, so it would be good to at least superficially compare these datasets. A transcriptome for a freeze-tolerant worm was also recently published, allowing for some cross-phyla comparisons (de Boer et al., 2018, Ecology and Evolution 8, 3774-3786). These comparisons don't need to be extensive, but a paragraph in the Discussion would be warranted, perhaps at the expense of some of the molecular details of protein processing (which are somewhat speculative and mainly based on mammalian literature). I also realize your work has somewhat different goals (i.e., cryopreservation vs. ecologically relevant freeze tolerance), but I still think a synthesis of the freeze-tolerance literature could strengthen the paper.

R: Thank you for suggesting these papers. As you also point out, here we are interested in application and focus on a single insect species that is relatively exceptional in its extreme freeze tolerance (i.e. a reductionist approach). A synthesis in the context of ecological relevance or phylogeny represents a different, holistic approach. Our intent is to identify candidate molecules specifically associated with truly extreme freeze tolerance. *Gryllus veletis* and *Enchytraeus albidus* may be considered 'freeze tolerant', however their LT_{50s} (-15 to -18°C) do not even approach that of the 'moderately freeze-tolerant' SD malt fly larvae (LT_{50} of roughly -80°C). We are interested not even in those SD larvae, but in what differs between them and SDA larvae that survive at -196°C. We therefore struggle with the relevance / what would be gained by a comparison (superficial or otherwise) to arthropods with such "weak" freeze tolerance. We do, however, appreciate that a metanalysis of insect freeze tolerance (based on transcriptomics or similar) would be highly useful in general (and this is apparently in the works by our colleagues!), but this is not our intent here. Considering our precise aims (as well as page limits and the fact that we are approaching 100 references…), we would prefer that such comparisons be addressed by other papers. We have, however, added additional citations of the Toxopeus and Sinclair (2018) review to our discussion to put our data in some context (lines 232,

260).

Minor Comments:

1. Line 13: “exploit” is a strange word choice here, perhaps reword?
R: Now changed: "Here we turn to larvae of the drosophilid *Chymomyza costata*…"

2. Line 24-27: This sentence is kind of vague, I suggest ending the Abstract with a stronger concluding sentence
R: We considered various concluding sentences that would be more concrete, listing some candidates, etc. However, we found that such alternatives were all too long and, mainly, too speculative. Our original concluding sentence is somewhat vague, but indeed we only identified some candidate proteins of freeze tolerance (and those are now the subject of functional assays in our ongoing work, which is another story as indicated in the Discussion lines 363-367).

3. Line 38: Saying cryoprotective dehydration results in loss of “most” of body water is somewhat misleading, because that makes it seem like >50% is lost. That may be true in some species, but water is lost until the melting point of hemolymph matches that of surrounding ice, so it may not be “most” of the water
R: We generally agree with the referee that many insects show only partial loss of water during cold acclimation, and have reworded as "loss of a substantial portion of the body water". Note that the major force driving the water out of the insect body in the presence of surrounding ice is, however, not the difference in melting points but rather the difference in partial water vapor pressures above the liquid and solid phases. This force is sufficient to drive truly most of the liquid water from the insect (or other invertebrate) with water-permeable integument, which was the original description of the "cryoprotective dehydration" strategy by Holmstrup and Westh, 1994. In a broader context, however, we agree with the reviewer that even minor loss of body water contributes to decreasing the ice fraction at a given subzero temperature and, this way, may contribute to freeze-tolerance.

4. Line 40: insert “of” before “insects”
R: Added!

5. Line 56: Ragland and Keep’s comprehensive metanalysis of diapause would be a good paper to cite here, and you could also revisit this paper in the Discussion to address the extent to which Chymomyza diapause is similar/different transcriptionally to other diapauses (although I realize diapause per se isn’t the primary focus of your paper) Ragland, G. J. and Keep, E. (2017). Comparative transcriptomics support evolutionary convergence of diapause responses across Insecta. Physiological Entomology 42, 246-256.
R: Thank you for pointing out this paper, which we have now cited on line 55. Here we aimed to uncover the DE genes associated with cold acclimation, rather than diapause. In addition, diapause-related transcription in *C. costata* was already covered by Poupardin et al 2015 (which was included in the Ragland and Keep metanalysis) and Kostal et al., 2017.

6. Line 74: I’m not sure a discussion of LEA proteins is warranted, because they seem to be rare in insects. To my knowledge, only one insect (Polypedilum vanderplanki) has LEAs, although they do function in lots of other desiccation/freeze tolerant taxa.
R: We agree that current knowledge of LEA proteins indicates that they are rare in insects (they occur in some collembolans in addition to *P. vanderplancki*). However, the identification of LEA proteins is

challenging due to weak sequence conservation (rather, they are functionally conserved) and this is most likely why they were not often sought out. They are at least seeing some light in the context of freeze tolerance now (e.g. see Sinclair and Toxopeus 2018, Biological Reviews). We have added citations for the protective effects of LEA proteins in desiccation stress for collembolans (and when applied to human cells!), and we indeed do find LEA-like proteins in malt flies (however the details are in a forthcoming paper).

7. Line 137-141: These numbers (19% of full transcriptome, 6,800 remaining contigs for gene expression analysis) seem low, especially since the previous transcriptome was also built with larvae. I don't have a lot of experience aligning reads to de novo transcriptomes, but do these results suggest that perhaps you should reassemble your transcriptome with the new reads? What percent of your reads aligned to the previous trnscriptome? It seems like you might be missing a lot.
R: The original malt fly *de novo* transcriptome assembly by (Poupardin et al. 2015) was derived from third instar larvae of both diapause and non-diapause variants, each collected at 6 different Zeitgeber times (to capture as many genes as possible). Considering that we sampled only diapausing / diapausing & cold-acclimated larvae (which we know have substantially suppressed gene expression), we are not surprised to have captured only 19% of the transcripts! We have added to the text "…to align sequences to a previously-assembled reference malt fly transcriptome *derived from both LD and SD larvae*…". We filtered the low-count transcripts further to avoid false-positive 'DE genes' based on, for e.g., 1 count vs 0 counts (which may have questionable biological relevance and which do not fit the profile of a functional cryoprotectant).

8. Line 145-148: Perhaps briefly justify your choice of a fold change cutoff. I'm personally against fold change cutoffs, because the biological consequences of "fold change" depend a lot on the particular gene. For example, small fold changes of metabolic or regulatory genes could have big effects on physiology due.
R: While we agree that some genes have large impacts even with small fold changes, the focus of this study is on cryoprotectants/molecular shields, which typically exert their effects in high concentrations (as we mention in the last paragraph of introduction). While we cannot say that high mRNA expression necessarily translates to high protein abundance, we can at least likely exclude those with very low expression. We therefore chose a typical 'strict' fold change cutoff, and we understand that these (like values for alpha) are somewhat arbitrary.

9. Line 164-173: A bit more information (either here or in the Supplemental Information) is needed to evaluate the KEGG analysis. What data went into these KEGG analyses, and what reference set of genes was used to test for enrichment analysis?
R: We have clarified the statement lines 164-166: "Differential pathway expression among malt fly variants was analyzed *based on read counts for transcripts with KEGG identities* using the Generally Applicable Gene-set Enrichment (GAGE) and Pathview Bioconductor packages [71, 72] in R.". If we understand correctly what you mean by 'reference set of genes': the larval variant used as a reference depended on the particular pair-wise comparison (all comparisons given in the Supplementary Figures).

10. Line 227: "As may be expected" is subjective, so I suggest removing this phrase
R: Agreed, and now removed.

11. Line 254: The statement "unlikely to be biologically relevant" is also subjective
R: We have reworded in an attempt to convey our skepticism without being overly subjective: "Some

pathways appeared to be statistically- rather than biologically-upregulated with progression from early to late diapause. For example, upregulation of KEGG pathway ko04141 ('Protein processing in endoplasmic reticulum') in late relative to early diapause was driven primarily by only a few genes; *secretory 61* (*Sec61*), *translocon-associated protein* (*TRAP*), and *Hsp70* (Fig. S5). Similarly, upregulation of pathway ko03050 ('Proteasome') during late diapause was extremely weak (based on two genes, each in only one of the three biological replicates; see Fig. S6)".

12. Line 269: The "anticipatory" hypothesis is supported by previous diapause transcriptome studies that show photoperiodic induction of heat shock proteins even in the absence of temperature change (e.g., Ragland, G. J., Denlinger, D. L. and Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. Proceedings of the National Academy of Sciences of the United States of America 107, 14909-14914.)
R: Thank you for suggesting this paper, which we now cite here (now line 267).

13. Line 287-289: This sentence is overly speculative, because many of these genes are kinases, so they are likely post-translationally regulated.
R: The reviewer is correct that current knowledge suggests that the three proximal sensors of ER stress (PERK, ATF6, and IRE1) are regulated mainly by post-translational modifications (binding to BiP, phosphorylation, dimerization, etc.). There is some evidence, however, that their gene expression may change developmentally, for instance with aging (Paz Gavilan et al. 2006, Neurobiol. Aging), which is reflected in a decrease of the pathway activity.
We have reworded whole sentence in order to make clear that we understand that the lack of differential expression in three ER stress-sensing pathways constitutes only indirect support for "anticipatory" hypothesis (lines 286-288).

14. I'm not sure how to articulate this, but I think you may be on to something by paying attention to the species-specific, unannotated genes. Pretty much every species that diapauses or undergoes cold acclimation upregulates heat shock proteins, but none of them other than Chymomyza survive cryopreservation. So that means Chymomyza either has "better" heat shock proteins, or perhaps they are not really an important piece of the puzzle. I have no doubt they're important for deep freezing, but I think the Discussion might be better served by focusing on the new findings. In other words, most of the Discussion focuses on heat shock protein/proteasome/ER, which we already know is important for cold hardiness. So I would perhaps shift the emphasis off heat shock proteins a little, even though the results are compelling, because heat shock proteins seem to be involved in every insect stress response and likely don't explain why Chymomyza can survive in liquid N2 (unless you have reason to believe they have super heat shock proteins).
R: We do intend to put most of our hopes into "novel" candidates, however the task of this paper was to identify the candidates; the task for ongoing work is to subject them to functional analysis. We had tried to synthesize and highlight the changes in protein processing, and now include additional justification for our discussion of Hsps as candidates (focusing on Hsp22, lines 322-336), and provide greater detail about our ongoing work that addresses the unannotated genes (lines 363-367).

Referee 2:

Comments to the Author(s)
This is a very nice paper indeed that was a pleasure to assess and there is very little that needs

comment.
R: We are glad that the referee enjoyed our paper!

I particularly like the use of four biological replicate populations for the validation of expression by qPCR. It would also be useful for a brief description of how the four reference genes were chosen for qPCR. This is referenced, but the details are buried there and there is little discussion as to why these genes were chosen over others. They are all ribosomal protein genes and I wonder if their expression is entirely independent.
R: The referee is probably right that the expression of different ribosomal genes is somehow interlocked. We selected these four genes based on our previous analysis of diapause-related gene expression patterns (Kostal et al., 2017) [Ref#27]. These genes showed minimum variation across many different phenotypes ranging from non-diapause, diapause warm- and cold-acclimated, and post-diapause (exact data presented in Kostal et al., 2017). Our personal experience (and general logic), however, indicates that no ideal reference gene exists. We hope that the major assumption of the qPCR analysis—that is to have equal, firmly established (even if not perfectly constant), background for all target genes—was fulfilled in our study.

The assignment of only 5% of the contigs in the full transcriptome to KEGG pathways seems low and has the potential to skew the assessment of gene differential expression. It would be helpful to include a little discussion as to what might underlie this low rate of assignment given that the proportion of transcripts matched to homologues was much higher. Does this reflect an absence of insect data in KEGG or an abundance of novel differentially-expressed sequences in C. costata?
R: We realize that this statement was confusing as originally worded. Roughly 5% of contigs were assigned KEGG IDs relative to the full, previous transcriptome assembly (by Poupardin et al 2015). Firstly, our samples (diapausing larvae) only covered 19% of the total transcriptome (which was based on both diapausing and non-diapausing larvae, at multiple Zeitgeber times). Of those 19% of contigs with read counts, 21% were assigned KEGG IDs - which is a clearer way of stating our result. We have reworded as: "Approximately 21% of the contigs with non-zero read counts were assigned to KEGG pathways" (lines 163-164).

In particular, I wonder what was seen in pathways likely to lead to proline synthesis (and degradation) given proline's role as a compatible osmolyte and in cryoprotection in C. crostata. One limitation in using differentially-expressed sequences in these sorts of studies is that sequences constitutively expressed in large amounts might be missed: perhaps sequences associated with proline metabolism might fall into this category and it might be worth looking at the abundance of these sequences even if they are not differentially-expressed.
R: Yes, we also expected a more clear reflection of massive proline biosynthesis and accumulation (linked with SDA phenotype) in our dataset. Although we found some DE genes putatively related to proline (Seq97173, Seq93777, Seq93284, Seq99387, Seq113370, Seq81662), and even more DE genes putatively related to trehalose, these changes were not systematically reflected in our KEGG pathway mapping. The reasons why this was not the case might be different. One reason (indicated by reviewer) could be constitutively high expression of proline metabolism-related genes during diapause. An alternative explanation: our sampling for RNAseq had only limited time-resolution (three different phenotypes), meaning that we potentially missed DE genes associated with proline synthesis that might occur in a specific (prior) stage of cold acclimation. Whatever was the reason, we have chosen to focus on findings that were supported by data (this is also dictated by word limit in the journal).

The finding of a large number of differentially-expressed but unrecognised sequences is intriguing and I would like to know more about these. Although some basic blast-based searches have been used to try and find homologues, I wonder if searches for open reading frames were considered along with a search of these against protein and motif databases. It is reasonable that this work might form the basis of another paper.
R: Yes, as suggested by reviewer, a more detailed functional analysis of select candidate sequences is on the way in our lab. We are sequencing the putative ends of genes, finding homologues by BLAST or by searching for putative open reading frames in unknown sequences, identifying localization signals and functional domains using number of different tools (collected in Expasy portal). Further, we are raising antibodies (and plan to use them for localization of gene products and their quantification in tissues). We are also expressing the candidate proteins in bacterial vectors in order to use them for *in vitro* analysis of their protective effects on enzyme activities. Next, we plan to transfect S2 cells and express the candidate proteins in them to see whether they can influence their freeze tolerance (in synergy with small molecular mass cryoprotectants). We now include a brief mention of some in-progress studies on lines 363-367.

I wonder also if there is room for discussion of ways in which interesting transcripts might be tested for their role in freezing tolerance. C. costata is not a model organism and techniques such as RNAi are not readily available, but it would be interesting to know how key candidates for freezing tolerance could be investigated further even if such work might not comprise part of this paper.
R: We have expanded on our 'future directions' sentence, while aiming to keep within word limits (lines 363-367): "For select unannotated candidates we are now conducting more comprehensive searches for their putative identities, raising antibodies for localization and quantification of gene products, and performing functional validation assays (e.g. for enzyme activity *in vitro* and survival of transfected cells post-freezing, results in a forthcoming study).

Referee 3:

General comments:
This paper addresses a fascinating topic—what is the biochemical response to extreme freezing stress—by taking a transcriptomics approach. The authors are rigorous in their number of biological samples for RNAseq and followup with additional qPCR studies and timepoints. The authors do state early in the introduction that indeed mRNA transcript abundance does not necessarily correlate with protein abundance or physiological process, which I greatly appreciate and is a crucially important distinction (especially in the case of extreme freezing). Overall, this large transcriptomics dataset on various acclimations treatments of a freeze-tolerant insect species a unique contribution to the study of “extreme freeze tolerance” and I find their title and scope to be both engaging and appropriate.

R: Thank you for your positive feedback!

One longstanding point of confusion is, however, that the authors state that a main goal in their study is to identify candidate proteins involved in extreme freeze tolerance (some proteins of which are speculated to be CPs, which could be non-protein cryoprotectants), yet the authors take a purely transcriptomics approach. It may be useful to justify to the reader why a transcriptomics approach was the preferred firststep, rather than proteomics or metabolomics (either independently, or in

synergy with transcriptomics). Although I do not feel these experiments would be strictly necessary for the publication of this manuscript, I am curious if the authors could include western blots of at least some of the candidate proteins (such as the HSPs), which may have antibodies known to work in several insect species. Such additional data would better link the author's transcriptomics findings with their goal to identify candidate proteins.

R: Comparative transcriptomics is broad and technically feasible early-stage means of approaching the topic of extreme freeze tolerance (which we know very little about), and useful for generating hypotheses about processes as well as proteins that may be involved. Metabolomics and proteomics studies to accompany this first step are in fact in the works (at various stages of completion). We have attempted to make our approach here more clearly justified by adding to the introduction (lines 96-102): "The transcriptomics approach is a rapid and technically feasible means of comparing among multiple phenotypes and time points, and will be complemented by forthcoming metabolomic and proteomic studies (in preparation). Cryoprotective proteins must be present in high abundance to act as molecular shields, and although mRNA transcript expression does not necessarily reflect protein abundance [55], this rough estimate of relative protein abundance allows us to focus on transcripts with large fold changes across larval variants.".

We are also performing a number of subsequent experiments for our candidates and have now mentioned this in more detail (lines 363-367): "For select unannotated candidates we are now conducting more comprehensive searches for their putative identities, raising antibodies for localization and quantification of gene products, and performing functional validation assays (e.g. for enzyme activity *in vitro* and survival of transfected cells post-freezing, results in a forthcoming study).

As an additional point of confusion: Does any work in the literature suggest that transcription or translation is even biochemically or biophysically possible under instances of extreme freezing (liquid nitrogen)? If so, please cite. If not, what is the biological meaning of DE transcripts (if it unlikely possible for them to be translated in response to extreme and sudden freezing)? I remain somewhat skeptical that de novo transcription or translation is possible under such harsh physical conditions.
R: The referee appears to have misunderstood our sampling conditions. None of the larvae sampled for sequencing had been exposed to LN_2; rather we sampled SD and SDA larvae at their respective acclimation conditions (i.e. prior to any freezing stress, but with the knowledge that SDA larvae are capable of surviving in LN_2, while SD larvae are not). These methods were clearly explained in both the text and Figure 1.

Otherwise, my main suggestions for this manuscript are calls for clarifications/stylistic changes, as well as providing more context and interpretation in the introduction and discussion/conclusions on several topics.

In the Introduction:
-Would it be possible to please give more context as to why the model species were chosen, or any additional information surrounding the species' habitat or natural history? Is it surprising that this species can survive such extreme cold? Do they experience such temperatures in nature (likely no)? If no, why do you predict this species can survive freezing in liquid nitrogen (is there cross-tolerance with another environmental stress)?
R: Considering that the lowest temperature recorded on Earth was -90°C, we can safely say that surviving in LN_2 is not part of the malt fly's natural history. Our intent in this paper is not to focus on the evolution/biological relevance of its extreme freeze tolerance. Rather, we know that the species

can survive deep cryopreservation and this is framed (in our Introduction) in the context of application. Considering our aims and given the strict word limit for the paper, we would therefore prefer not to comment on the malt fly habitat/natural history. We would like to stress that there is no logical reason 'why' *C. costata* larvae survive temperatures that never existed on Earth, we simply want to know 'how' they do so. We provide a number of references to papers about the growing story of 'how' in this model [ref #s 13, 22, 23, 27, 35, 56, 57, 63, 87].

-The introduction might be a good place to both address why a transcriptomics approach was chosen and what evidence you might have from the literature to believe that de novo transcription is possible under conditions of extreme freeze stress.
R: This transcriptome comparison is only part of an investigation (the metabolomics study is forthcoming), and we have added this (plus justification of our approach) in the introduction lines 93-99 (written out in the response to your second general comment, above). Note that we did not assemble a transcriptome *de novo* (that was done previously by Poupardin et al 2015; we aligned our reads to those transcripts). Regarding transcription under extreme freeze stress, please see response to your second general comment, above.

In the Discussion:
-The authors describe using the Oligo(dT)23 anchored primers to enrich for mRNAs/RNAs with poly(A) tails, in their transcriptomics work. These primers are standard, but a consequence is that the transcriptomics does not identify or consider non-polyadenylated RNAs (some subset of regulatory RNAs). Non-polyadenylated RNAs—such as mature microRNAs—may play important roles in protein processing and gene expression (most often by blocking protein translation). Since the authors are very interested in the role of protein processing in extreme freeze tolerance, I feel this experimental limitation should be mentioned/considered. Sequencing non-polyadenylated RNAs would also be an interesting next step, in future work!
R: Thank you for this very good suggestion as a next step and which is definitely a 'hot topic' of current research. We now mention this in the opening of the discussion: "The comparative transcriptomic approach is an hypothesis-generating first step in understanding the physiological mechanisms of extreme freeze tolerance, however we acknowledge that other forms of gene and protein regulation which may be important (e.g. miRNA or post-translational modifications) [73] are not captured by our methods and therefore warrant investigation in future."

-It would be helpful to consider whether transcriptional regulation vs. other forms of gene regulation beyond transcription, in your interpretation in changes in transcript abundance or predicted physiological processes.
R: We agree, and although we cannot speculate on these other regulatory processes throughout the discussion, we have attempted to convey the sentiment in the opening paragraph of the discussion (see response directly above).

This was a pleasure to read; thank you for your hard work!
R: We are glad to hear it!

Specific comments:
Line 19: The phrase “genes and processes that accompany the physiological transition to extreme freeze tolerance” may be too broad. Perhaps “genes and processes” could be more accurately described as “transcripts.”

R: We intended to use KEGG analysis specifically to identify processes involved in extreme tolerance, as individual transcripts do not necessarily provide the full functional picture in the pathway. We have now reworded as: "We identify *mRNA transcripts representing* genes and processes that accompany…".

Line 46: It may be helpful to add one additional sentence of context as to why "These extremely freeze tolerant animals may hold the key to advancing techniques in cryopreservation of tissues, organs, or even complex organisms [15]." What would one hope to identify from especially freeze tolerant animals, to apply to medical applications?
R: Sentences about what we hope to gain (i.e. involvement of known macromolecules in extreme freeze tolerance and/or novel cryoprotectants) do appear shortly after, on lines 81-86 (we felt this location was better for flow, as we wanted to keep the opening paragraph general).

Line 55: Purely stylistic comment, in the phrase "Both adaptive complexes - diapause and cold acclimation - are based on the massive" the dashes should be em dashes (—) so the phrase is rather "Both adaptive complexes—diapause and cold acclimation—are based on the massive"
R: Thank you for pointing this out. We have made the change.

Line 98: Question: Is it possible to justify in the text why you took a transcriptomics approach if you write your aim is to identify candidate proteins…" Why not a proteomics or metabolomics approach? It might be helpful to justify to the reader your experimental rationale.
R: Unknown principles require untargeted approaches – such as global transcriptomics by RNAseq (not by microarrays, for instance), which is also the most easy of all -omics. We have now added a justification to the text on lines 96-102 (details in the response to your second general comment, above)

Line 152: The E-value is very relaxed at E-1. A more "standard" E-value would be E-5. Please describe why such a relaxed E-value chosen.
R: This additional BLAST was specifically not standard, and performed with the aim of assigning at least some putative identities to uncharacterized/predicted proteins (which were based on the default Blast2GO E value of -3). Please, note that we specifically indicated that this BLAST was less-stringent: "We attempted to identify the uncharacterized or predicted proteins via a less-stringent BLASTx search against the nr database (E-value threshold = 1,…)".

Line 158-162: What is meant by "custom categories"? I personally felt that the "'Protein processing" category was too general/vague and could likely be split into at least 2 more specific categories, in terms of biological function. Likewise, I found it confusing to have both an "Other" and "Unknown" category (although I think I understand the intent).
R: It is always a challenge to organize such large datasets in meaningful ways, and we did this 'by hand' to identify relative trends (we make no specific/strong conclusions from this categorization). Of course there is always a subjective element here, but we feel that it helped to visualize the data without the burden of many specific subcategories. "Unknown" genes had no putative identities, while "Other" genes had identities but simply did not fit unambiguously into the major categories we chose.

Line 221/222: The word "true" in "true physiological mechanisms of freeze tolerance" is confusing. Have untrue mechanisms been proposed in the literature? Or what makes information collected from transcriptomics in particular more "true"? Perhaps delete this word to avoid confusion.

R: We agree, and have removed this word.

Line 229: It would be helpful to remind the reader how you define differential expression (what is the threshold & what is meant by up vs downregulated, and what a sample is DE in comparison to)
R: This is meant to indicate the general trend, and in the interest of keeping within word limits we would prefer not to reiterate lines 143-145, i.e. "…genes were considered to be differentially expressed (DE) if the P-value (adjusted for false discovery rate; FDR) was < 0.05 and the absolute fold change between treatments was ≥ 4 (log2 fold change ≥ 2).". What the DE is in comparison to is given in the sentence "between early and late diapause" vs "between early diapause and cold acclimated". We have attempted to clarify somewhat by moving the number of DE transcripts such that it follows each comparison: "We observed fewer DE genes between early and late diapause phenotypes (190 transcripts) than between early diapause and cold acclimated phenotypes (776 transcripts).".

Line 233: “Only a single DE gene (Seq109825; downregulated but with no putative identity, i.e. not shown in Fig. 2) overlapped between late diapause and cold acclimation (relative to early diapause), suggesting that extreme freeze tolerance involves a unique set of physiological processes” I am unsure how supported the conclusion that “extreme freeze tolerance involves a unique set of physiological process” is; it is possible that 1) at least some core physiological processes are not controlled by de novo transcription under extreme freezing or 2) transcripts for important physiological processes to survive extreme freezing are constitutively expressed
R: Please note that we were careful to word this as "*suggests*", as we cannot make conclusions stronger than that. We have attempted to tone this down further, and make the sentence more precise: "…*suggesting* that cold acclimation and development of extreme freeze tolerance *may* involve a unique set of physiological processes distinct from that linked to diapause progression.". For your comment 1): There is some confusion on the part of the referee about *de novo* transcription/transcription while in LN_2 (which was not investigated here). For comment 2): If you mean that particular genes are constitutively expressed in all larvae, then on the whole they are unlikely to explain differences in freeze tolerance among larval variants. If you mean that particular genes are constitutively expressed in SDA larvae, then they should show up as DE relative to SD larvae. However, we do agree that there may be processes either constitutively expressed in all diapause phenotypes (SD and SDA) or those not captured/represented by transcriptome profiling at the time of sampling.

Line 248: category 'Protein processing’ is very general
R: Please see comment above regarding the "Protein processing" category.

Line 262: A suggestion regarding the phrase “it seems obvious that extreme freeze tolerance should largely rely on upregulated protein protection”: remove the phrase “it seems obvious” with several citations of when process has been observed in other species or systems
R: We have now removed "it seems obvious".

Line 272: Regarding the phrase “Cold and freeze-dehydration stresses may hit both nascent and functional proteins” replace the word “hit” with what is predicted to occur to the proteins biochemically (more specifically)
R: We have changed this to: "Cold and freeze-dehydration stresses may impact the process of nascent proteins folding or mature proteins higher order structures, and cause their misfolding or unfolding respectively.".

Line 300-301: In the phrase “In extremely freeze-tolerant larvae, six genes encoding chaperones involved in protein stabilization and refolding were enhanced” what does the word enhanced mean? DE? Or something else? Please clarify in the text.
R: We agree this was not clearly worded. Now changed to "upregulated".

Line 334-336: Regarding “The transcriptional profiles of lethal(2)efl (Fig. S10B) and hsp70 (Fig. S10E) matched relatively well with extreme freeze tolerance and, moreover, showed relatively high ddCT values (suggesting that their protein products are abundantly expressed).” Could you please justify why a “relatively high ddCT” suggested abundant protein products? This may be a controversial claim if not justified.
R: As with any untargeted transcriptome comparison, we cannot make any claim about protein abundance from mRNA abundance alone, thus we were careful to only "suggest" that higher ddCT values (i.e. early accumulation of amplicons during qPCR procedure caused by relatively high initial abundance of mRNA transcripts in the sample) are likely linked with more abundant protein products. However, as the reviewer is rightfully uncomfortable with the assumption we have removed this part of the sentence.

Line 359-360: The phrase “We hope that this set of unidentified genes contains novel and potentially important cryoprotectants” seems perhaps overly speculatory. Reword?
R: Now reworded as "Ideally, this set…".

Line 363: Can the methods of future functional verification be described in any additional detail?
R: We have many things in the works, however are exceptionally limited for space based on the journal requirements. So, we have briefly expanded on our 'future directions' sentence (lines 363-367): "For select unannotated candidates we are now conducting more comprehensive searches for their putative identities, raising antibodies for localization and quantification of gene products, and performing functional validation assays (e.g. for enzyme activity *in vitro* and survival of transfected cells post-freezing, results in a forthcoming study).

Comments on Figures:
1. Figure 1 is a bit blurry in the proof. Is it possible to use a vector file?
R: The first proof included lower-res figures with their captions first, followed by the full-resolution/vector images.

2. Figure 2A - Perhaps put number of genes in each pie slice, to make visual more informative
R: We had considered this, but given so many tiny pie slices this would end up looking rather crowded / messy! We hoped that having the total gene numbers in brackets below each pie would suffice (as it is all relative).

3. Figure 3 - Although I love the organization of Figure 3, it might be important to emphasize that transcripts do not strictly result in translation of protein or physiological process. This figure may unintentionally convey that these proteins/processes were directly observed.
R: We agree. We have clarified by rewording the first sentence of the caption as: "Upregulation of transcripts involved in protein processing machinery associated with acquisition of extreme freeze tolerance in malt fly larvae".

4. Figure 4A - What is meant by “color areas used to visually enhance?” Confused. Font size in all of Figure 4 is small.

R: We have now increased the font sizes and reduce visual clutter in Figure 4. We have also defined the colored areas in a way that we hope is more clear: "Shaded areas behind the bars indicate survival in liquid nitrogen for SD larvae (grey shading) and SDA larvae (blue shading), as derived from panel A (note that survival was zero for non-diapause LD larvae and also SD3 larvae).".

Appendix C

RESPONSE TO REFEREES
Transcriptional analysis of insect extreme freeze tolerance
RSPB-2019-1274 (Revision)

Associate Editor Board Member

The authors have addressed most of the concerns of the reviewers. The main remaining concern is regarding the transcriptome coverage. Please provide read mapping statistics as requested by Reviewer 1. The other suggestions provided by both reviewers should also be considered.
R: We have made the changes suggested by the referees (below), including the mapping statistics, and hope that the manuscript is now publishable in Proceedings.

Referee 1

I think the authors have made some important improvements to their manuscript. The rationale for the study is improved, and appropriate caveats for RNA-seq are now included. I am mostly satisfied with the authors' responses to my previous review, although there are a couple of cases I thought they could have done a better job. See below:

1. In my first review, I indicated that I thought the authors could have done a better job of integrating other physiological/transcriptional studies from the literature. The authors responded that because they are focused on "truly extreme" freeze tolerance, other studies are not relevant. I see their point, but I still think there is room to include these studies. Comparisons to other moderately freeze tolerant organisms, and the authors' goals identifying mechanisms of extreme freeze tolerance, are not mutually exclusive endeavors. Chymomyza does freezing better than any other insect that's been studied, so that must either come about by completely new mechanisms, by augmentation of existing mechanisms, or some of both. If you compare the classes of genes identified with extreme freeze tolerance to those involved in other types of freeze tolerance, you could start to address this question and provide more insight into the nature of your phenotype. If word limits are an issue, you probably could cut down the number of references without compromising the paper. I appreciate the thoroughness of your literature search, but in some cases multiple references are cited when 1-2 key references would suffice.
R: Here we do not wish to take the 'classic' approach of tackling all classes of genes in our transcriptome and relating them to the literature on weakly to moderately freeze-tolerant insects. Rather, we wanted to focus on genes specific to *C. costata* in the extremely freeze-tolerant state. Still, as the referee feels strongly about this point we have replaced the discussion paragraph (original lines 229-240) with the following paragraph that attempts to put our results in better context with literature on general mechanisms of insect freeze tolerance (lines 233-261):

"In a recent review of insect freeze tolerance, Toxopeus and Sinclair [7] hypothesized that five broad mechanisms are involved: (1) control of ice formation; (2) reduction of ice content; (3) stabilization of macromolecules; (4) management of biochemical processes / reduction of harmful metabolite damage; and (5) post-thaw repair and recovery. We did not expect to see a reflection of mechanism 1 in this study, as no gene annotations in the *C. costata* transcriptome [63] had identifiers such as ice binding, nucleation, or thermal hysteresis. Moreover, our knowledge suggests that *C. costata* has limited capacity to control ice formation: neither supercooling capacity nor vulnerability to ice inoculation changed much with acclimation [23, 35], and there is no sign of thermal hysteresis

activity in larval hemolymph (Koštál and Rozsypal, unpublished observations). Ice content (related to mechanism 2) differs only slightly among *C. costata* larval variants with decreasing temperatures [described previously by 23]. Mechanism 3 likely involves cryoprotectant accumulation as a non-colligative means of stabilizing macromolecular structures, and despite that cryoprotectant accumulation is a hallmark of *C. costata* freeze tolerance [23, 35], we found little direct reflection of it in transcriptome. Similar lack of evidence for direct transcriptional control of cryoprotectant synthesis (myo-inositol, proline, and trehalose) was reported for the freeze tolerant cricket, *Gryllus veletis* [43]. The reasons for such results may relate to the importance of post-transcriptional control mechanisms [43], general problems with interpretation of metabolism using –omics approaches [74], and/or technical limitations such as insufficient resolution at time and tissue levels. Mechanism 3 also includes upregulation of molecular chaperones [a common phenomenon of insect thermal tolerance, e.g. 75, 76], which was clearly supported in the *C. costata* transcriptome (discussed further in the next section). Mechanism 4 was reflected in the form of global downregulation of processes in *C. costata* larvae linked to active metabolism, including oxidative phosphorylation (Fig. 1B; a full list of the pathways provided in Fig. S3). Still, ribosomal transcription was generally maintained during cold acclimation (Figs. 1, S3-S5). The aspect of mechanism 4 relating to reduction of damage from harmful metabolites showed rather a trend of downregulation (Spreadsheet S1) and no systematic reflection in KEGG pathways. Mechanism 5 includes some elements overlapping with mechanism 3 (e.g. protein processing machinery), which appears to be a central theme of the transcriptomic transformation in extremely freeze tolerant *C. costata* larvae."

2. I was concerned that only 19% of the transcriptome was represented in the current study, and the authors pointed out that diapause causes a general repression of transcription. While this may be true, to allay my concerns it would be nice to see some read mapping statistics, which I had asked for in my previous review. The percentage of mapped reads is an important indicator of quality for an RNA-seq experiment and is frequently reported (see Conesa et al., 2016, Genome Biology 17:13). If you do indeed have good read mapping percentage (typically above 70% for the studies I've looked at), then I would be satisfied with your response. However, if your mapping statistics are poor, it could indicate the presence of new unassembled transcripts or technical issues.
R: We apologize for missing the previous request for proportion of mapped reads. When we aligned the reads to the reference transcriptome with Bowtie2, we had kept some default settings to (1) not separate out the aligned and unaligned reads, and (2) to not print the mapping statistics (lesson learned!). We therefore repeated our alignment step exactly as before with each library to obtain the mapping statistics, and just over 99% of reads align in each case (range: 99.11 - 99.26%). On average, 54% of the reads aligned once and 45% of the reads aligned more than once. Most importantly, based on the referee's comment we went back and critically thought about how we arrived to our statistic of 19% transcript representation. The original transcriptome assembly (by Poupardin et al. 2015), which we were advised to align to, had a total of 113,447 contigs. Our library (i.e. 22,872 contigs with at least 1 read count) represented 19% of that total 113,447 contigs. However, as that full reference assembly had many redundancies, Rodolphe Poupardin had afterwards used CD-HIT to cluster and refine it down to 21,326 contigs (which was then used going forward in their 2015 manuscript). When we matched our 22,872 contigs to Rodolphe's refined library of 21,326 contigs, 18,357 of the contigs were shared. In other words, our dataset captured 86.08% of the 'refined' transcriptome. This is a much more relevant comparison! So, thank you very much for pointing this out again. We are glad to make this correction (and it makes a lot more sense now to us as well!).

We have now reworded our methods as follows (lines 131-137): "Through the Galaxy web service we used Bowtie2 [62] to align reads to a previously-assembled reference malt fly transcriptome derived

from both LD and SD larvae [a full assembly of 113,447 contigs, which was further refined to 21,326 non-redundant contigs; 63]. Just over 99% of our reads aligned to the full reference assembly. We then used Cufflinks [64] to assemble transcripts (contigs) and estimate read counts. These normalized libraries contained a total of 22,872 contigs with non-zero read counts [representing 86.1% of the refined reference assembly; 63].".

Additional Minor Comments:

1. Line 97: I also think the sensitivity of RNA-seq is attractive for a study like this. Proteomics and metabolomics are typically 1-2 orders of magnitude less sensitive (at least in terms of the number of resolved features) than RNA-seq. May be less of concern for you if you're interested in high-quantity cryoprotectants, but the ability to measure true genome-wide expression is a nice feature of RNA-seq.
R: We thank the reviewer for pointing out this beneficial aspect of transcriptomics and have now included that on lines 96-99: "The transcriptomics approach is a rapid, sensitive, and technically feasible means of comparing…".

2. Line 108-109: Suggest changing to "…on artificial diet in MIR-154 incubators (Sanyo Electric, Osaka, Japan) as described previously [57]. For experiments, three different malt fly larval phenotypic variants were generated according to…" As written it makes it sound like the incubators are generating the phenotypes, when it's the temperature/light regimes that are more important.
R: We have now made this change as suggested.

3. Line 215: This is more of a personal qualm, but I don't like it when transcriptomics studies are referred to as "hypothesis-generating" studies. I think it undersells the value of transcriptomics and contributes to the narrative that genomics is a "fishing expedition." That may be true in some cases, but if you are testing clear mechanistic hypotheses and the RNA-seq experiment is properly designed and replicated, it can do more than generate hypotheses. But ultimately this is your work, so it's your decision how you want to frame the study.
R: Indeed we do not wish to undersell the value of transcriptomics, however in our case we are truly 'fishing/hypothesis-generating' as we hoped to identify theoretical cryoprotectants. Still, we agree with the reviewer that transcriptomic studies need not be only hypothesis-generating in general, so we have reworded the sentence to be more specific about how we used the approach and why (lines 215-220): "We used the comparative transcriptomic approach as an hypothesis-generating first step in seeking new candidate cryoprotectants and to further understand the physiological mechanisms of extreme freeze tolerance,…".

4. Line 229: Change "that" to "those"
R: Now changed to "those".

5. Line 242: Change "were" to "was"
R: Because we refer to genes (plural), we would like to keep the use of "were" ("Expression of genes…were").

6. Figure 3 and select Supplemental Figures: In heat maps and similar figures, it's advisable to avoid red/green color scales, to make the figures accessible to color blind readers. If it's easy to customize, suggest going with a blue/yellow scale or another palette that is discernible by color blind individuals.

R: We thank the referee for pointing this out. Our second author is in fact red-green colorblind, so we had attempted to make the figures color-blind friendly by including up or down arrows next to genes in Figure 3 (now Fig. 2), and by separating up- and down-regulated pathway sets in the supplemental KEGG Heatmap. However, we now realize that the individual KEGG pathways indeed needed correction, so we have changed the greens to blues throughout all the figures.

Referee 2:

I'm very happy with the responses to my questions in the previous review and I see almost nothing that new needs attention. This is a nice paper and I look forward to seeing the other analyses mentioned.
R: We are glad that the referee was satisfied with our revised manuscript!

There are two things I would like to see changed: 'comprised of' to 'comprising' at L118 and L185, along with the replacement of 'x' for '×', the multiplication sign at L150.
R: We have changed the word to "comprising", and replaced the x for a proper multiplication sign on new lines 150, 151.