Positive Darwinian Selection Promotes Diversity Amongst Members of the Antifreeze Protein Multigene Family. W.J. Swanson and C. F. Aquadro.

This is the second time I am asked to review this manuscript. The first version was through Mol. Biol. and Evol. which had declined acceptance, and the current version through J. Mol. Evol.. In order not to repeat myself, I have attached my comments to the first version as reference for the JME editor. I have indicated to the JME editor that I have reviewed a prior version of this ms, and thus a comparison to that is unavoidable to determine if suggested improvements or corrections have been made.

The organization of the current version remains the same as the MBE version - the statistical analyses of the AFP (antifreeze protein) sequences to demonstrate positive selection, and the functional inference of the putative residues under selection. The authors's functional inference was very problematic in the MBE version and I listed some of the reasons why I think it's wrong (see comments to MBE ms). In the JME version, the authors have toned down their previous emphatic statements about the functional significance or the adaptive value of the selected residues, but the gist of their hypothesis remains the same.

The amino acid residues the authors identified to be under positive selection are outside of the ice binding surface (structurally identified in the fish type III AFP). The two related inferences the authors make are (1) these neighboring residue changes allow protein specificity to evolve, and (2) they result in a heterogeneous suite of antifreeze molecules that may interact more efficiently with the morphologically heterogeneous ice surface (page 10 middle paragraph).

As I have pointed out in my review for the MBE version, the first inference is erroneous because the specificity of an antifreeze (i.e. which crystallographic plane in ice it binds to) as far as we know is determined by the residues that comprise the ice-binding surface, and these residues are invariant among paralogous isoforms (at least in type III AFP). Changing the residues outside of the ice binding surface does not affect specificity. The changes may affect how strongly the antifreeze binds (I call that affinity) and not where it binds (or specificity).

The problem with the second inference is the authors's fixation on coupling heterogeneous AFPs with morphologically heterogeneous ice surface. The ice surface is "rough", and in the condensed matter physics sense, gentle or broad. It is carefully organized to minimize surface free energy and in no way morphologically diverse (heterogeneous in the JME ms). In fact I do not know what the authors mean by the ice surface being Amorphologically heterogeneous≅ and I doubt if they do. Again Delzeit et al 1997 was cited to be the source information for morphologically heterogeneous ice surface and I have previously pointed out that's the wrong reference (see comments to MBE ms). The issue of the ice-water interface in the context of AFP binding at physiological conditions has been carefully addressed by Haymet and his co-workers (eg. Eur. J. Biochem. (1999) 264, 653-665 and references therein to his related work) and the authors would do well consulting a proper reference such as this.

I could agree that the selected changes in the neighboring residues may help the AFP molecule positioning more effectively at the ice-water interface for binding by the ice-binding residues to occur. But this is only one of many equally probable hypotheses that are well articulated by the authors in the second paragraph of the Discussion section (page 9). The matter of AFP solubility is as critical as any because of their high physiological concentrations (>10 mg/ml in blood) and some types of AFP indeed are difficult to re-dissolve when purified. Why the authors must insist on their chosen inference is difficult to understand, except for allowing them to draw parallels to other examples of protein-ligand interaction.

To support their argument, the authors said many of the adaptive residues identified involved charge changes (page 10 middle paragraph). It is not indicated what these changes are - uncharged to charged, non-polar to polar, or vice versa? Indeed this depends on which residue is considered to be the ancestral state. I honestly do not see a clear pattern as the authors claim they do. I should also point out that two different groups (P.L. Davies and A. Haymet) have recently experimentally shown that hydrophobic interactions may play a role in antifreeze binding as opposed to (or in addition to) the widely accepted hydrogen bonding. Thus a change from an uncharged to charged residue may not necessarily be adaptive.

There are a couple of other important issues. For type III fish AFP, the JME ms included an additional type III AFP sequence, from the Antarctic eel pout (which I suggested), and the calculations were redone. Interestingly, half of (6 out of 12) the residues identified to be under selection in the last data set (without Antarctic eel pout AFP) did not make the cut statistically in the recalculation. What would happen if a bigger yet data set is used? Of course if a reviewer hasn't seen the MBE ms, this question would not be raised. But since I have, it is a legitimate question to ask. The same question can be asked of the insect AFPs. The authors used the AFPs of the meal worm, *Tenebrio molitor*. The AFPs of another insect *Dendroides canadensis* (Jack Duman's work) have been published prior to the authors's MBE ms and more cDNA sequences were published subsequently. As far as I can tell, the AFPs of these two taxa are homologous as they have the same repetitive structure and share substantial sequence identity. So why haven't the *D. canadensis* sequences been included in the analysis.

There are also a couple of persistent minor errors that are annoying and reflect sloppiness. A wrong accession number for the type III in the MBE ms is replaced by another wrong one in the JME ms (J09323 does not exist), even though I penciled in the corrections in the MBE ms.. I also pointed out in the MBE review that the *T. molitor* AFP sequences have been revised and there were changes in the protein sequence, but the changes weren't made in the sequences in the JME ms, and I have penciled in the residues in the MBE ms, and I have asked the MBE editor to return the ms to the authors.

In sum, my overall view on the JME ms is similar to the MBE ms. The authors have identified some residues that may be under positive selection, but no plausible suggestion as to why the selection occurs. Their emphasis on the Aselection≅ of AFP heterogeneity to counteract Aheterogeneous surface of an ice crystal≅ is purely

speculative, thus I am not sure if this ms is suitable for a journal like JME. I have previously suggested the authors to consult the antifreeze literature in greater depth or consult people in the antifreeze field, which appeared not to have been heeded. The acknowledgments stated that some of these ideas (in the ms) were explored by WJS during a NSF sponsored course on biological adaptations of Antarctic marine organisms, so why not consult Art DeVries, one of the instructors in that course, and the discoverer and leading authority in fish antifreeze proteins.

Review of "Positive Darwinian selection promotes heterogeneity amongst members of the antifreeze protein multigene family" by Swanson and Aquadro.

Swanson and Aquadro examine a number of (putatively) paralagous sequences of antifreeze proteins. They perform a number of statistical analyses of these sequences and conclude that there is evidence of positive selection acting on the sequences. I was asked to review this paper in the light of the previous two reviews. I short, I like the paper and hope that it will be published in JME.

The phylogeny-based tests of positive selection used by the authors are first rate. First, they use a codon model in which a site can be in one of a number of selection classes. The likelihood is calculated by summing probabilities over the categories (weighted by the prior probability that the codon is in each category). They can do two things with this type of model. First, they can perform a test of a model that does not include a class of positively selected sites (omega > 1). Second, they can identify sites that are under positive selection by examining the posterior probability that a site is in the positively selected class. Their analyses indicate that a model that includes a class of positively selected sites is significantly better than one that doesn't (using a likelihood ratio test). They responded to the first reviewers criticism that the chi-square distribution may not be a good approximation in this case by simulating the null distribution. They come to the same conclusion. They were also able to identify a number of sites that have a high (posterior) probability of being under positive selection. They find that these sites surround the putative ice-binding surface.

They also examine a model that allows omega (the nonsynonymous/ synonymous rate ratio) to vary across lineages. They compare

this model to a null model in which omega is constant across the tree. They reject the null, indicating that selection has been acting heterogeneously across the tree.

I think that Swanson and Aquadro adequately responded to the criticisms of the first reviewer (the one who had more statistically-based criticisms). It is not clear to me what they could or should do to appease the second reviewer. I have no problem with a bit of speculation about cause in the discussion of a paper.

Review of "Positive darwinian selection promotes heterogeneity amongst members of the antifreeze protein multigene family" by Swanson and Aquadro.
Many organisms have independently evolved antifreeze proteins (AFPs), which bind ice nuclei and reduce the freezing point of water, thus allowing the organisms to live in sub-zero temperatures. The authors used the fish and insect AFP sequences available in the GenBank to test the hypothesis that the paralogous AFPs have diverged under positive darwinian selection. Although the subject is interesting, their analysis is incomplete. My detailed comments follow.
1. They stated in the manuscript that they only studied paralogous AFPs. But the fish sequences used were apparently from multiple species (end of page 3). It is unclear how they determined that the genes are paralogous, rather than orthologous. In fact, if orthologous gene sequences are available, it will be more interesting to examine if positive selection also operates in orthologous genes. Thus, they may know whether the selection was for diversity among genes or among species, which will be useful in the search of the selective agent.
2. The sequences used here are relatively short (62-81 codons). A previous study showed that the positive selection tests based on the large-sample assumption may give false positive results when applied to short sequences with low divergence (Zhang et al. 1997). The likelihood ratio test also assumes large samples when the chi-square test is used. I would like to see if positive selection is substantiated for the AFPs when small-sample tests are used.
3. In contrast to what the authors claimed, the likelihood method they used for detecting selection does NOT assume a constant rate of synonymous substitution among sites. A high dn/ds ratio at a site can be due to either a reduction in ds or an increase in dn. Since fluctuations of the synonymous rate among sites happen in evolution (e.g., due to codon usage bias), the likelihood method has the potential of making mistakes in such cases.

4. The authors assumed a number of models in the likelihood ratio test of selection. Are these models biologically meaningful? Is there any empirical evidence for the beta distribution of dn/ds among sites? I am skeptical of the exceptionally high values of w (32.1 in beetle AFPs). This w value means that for every synonymous substitution, there will be about 32.1*3=96 nonsynonymous substitutions! If this is the case, one cannot even correct multiple hits.
5. The discussion part of the manuscript relies heavily on the inferred sites that are under selection. Are these estimates reliable? Since likelihood is known to be sensitive to models and the authors did use some simplified models, the estimates may not be robust. After all, no one has shown that the method can accurately predict those sites even in simple computer simulations. I wonder (1) if the identified sites are also the most variable sites and (2) if ice-binding sites are conserved. Also, if the authors include all those sites that have posterior probability of >0.9 or >0.8, will their conclusion remain unchanged?
6. The authors claimed that there are many amino acid substitutions involving charge changes, but they did not perform any tests to demonstrate that. They should compare the rates of conservative and radical amino acid changes with regard to charge profile.
7. Although the authors proposed that the paralogous AFPs are for binding diverse ice surfaces, no direct evidence is provided. The analogy between ice and pathogen is not appropriate, because pathogens are under selection so that there is coevolution between host defense genes and antigenic genes (the red queen scenario). But for ice, the ice types never change, so presumably after enough AFP genes are generated, no more diversifying selection is needed. The authors may want to clarify this difference in their analogy.