Hi all,

Attached is an example where I just transformed the parameter estimates from MANOVA as a proof of concept. You can see it tells a very different story (assume error bars from second figure would be on the first, I don’t know how to transform SE quickly and accurately).

For example, now “Weevils First, Infective Aphids” increases OPR1 expression the most. The rest have reduce expression, particularly those without weevils (yeah this makes sense with the plant-herbivore literature!). “weevils First, sham aphids” has a lower expression level than this, but its still higher than the remaining. This means the virus interaction alters plant response to chewing herbivores. Neat!

To  make sure I’m interpreting correctly, a value of 1 would mean there was *no change* in expression for OPR1 or LOX2 compared to Beta Tubulin (e.g. they were equal). 0-1 indicates less expression than beta tubulin, and 1+ indicates increase in expression compared to this housekeeping gene. W-A- is our control of just clean aphids, and this is the lowest value as it should be. LOX2 is a less exciting story, but I suppose that is why we looked at multiple genes. Some will respond more strongly than others.

I also checked out chitinase. Virus causes a spike in this SA gene, while weevils have a minimal impact.

Saumik and Clare – let me know if the direction of these values *make sense*. I hope this might be the solution we’re looking for.

Dave – after thinking about the analysis, it might make sense to do a back transformation on marginal means and the upper and lower bounds of the SEM in R. Values that scale 0-1 and 1-infinity grossly violate the multivariate normal distribution. Transforming model output is a much more sophisticated approach than what other folks have done in the literature, where they probably also just do a transformation at the end and plot the mean and SE from an excel file (no offense to folks who do that 😊).

Finally, I have to admit that Saumik might’ve been right all along! While fold change in expression is confusing, it might produce the more sensible outcome. I’ll let you all vote on it. However, I’m sticking to my guns that I like the **2^-delta ct** more than **2^- delta delta ct** because it lets us plot the control as a comparison. I’m very excited to get this done!

Hey all,

**Short story:**  I think we’re at a good solution and I’ll make the tables and formal figures soon. I want to make sure you like this approach (see attached for newest version compared to oldest).

I attached the default chitinase example. The control bar should = 1 and is a point of reference. Genes closer to 1 are more similar to the control condition, and including the control in the 2x3 diagram is an aesthetic choice. I suggest keeping it because it shows we’re making a specific hypothesis test comparing treatments to controls, and treatments to each other.

**Long story:**

I tried a few different ways to do this, and originally both Delta CT and Delta Delta CT were giving very strange values; that gene expression increased thousand fold, insanely large or small error bars, etc. etc.

I think the “traditional” way to do this is sketchy from a theory standpoint.

If…

A= treatment beta tubulin ct

B= treatment gene ct

C= control beta tubulin ct

D=control gene ct

e= statistical / experimental error or “noise”

The formula is:

2^-(((A+e)-(B+e))-((C+e)-(D+e)))

The error component in each iteration is added to each step in the calculation. Additive effects of error are not strong, but when you consider this is happening for 20 genes and 5 treatment comparisons, it’s a big problem. If one reaction was even slightly off, the outcome will be biased towards large values.

The solution is to calculate Delta Delta CT *entirely* with parameter estimates from MANOVA.

All Genes(A,B,C,D) = virus\*weevil + error

In other words, I asked R to predict all the CT values and SE for us based on the entire raw data, and then did the relativel expression calculations from there. This is much better than calculating 2^-(((A+e)-(B+e))-((C+e)-(D+e))) *and then* doing MANOVA.

In doing this, we can ask the model to give us any level of gene for Treatment Delta CT for the controls and treatments, and then error is accounted for in the model itself (sort of). Then I subtract the Control Delta CT, plot 2^-x, and voila!

**Some thoughts:** Lots of folks are employing gene expression assays, and they should be evaluated correctly. The cornerstone paper on the topic has a very simplistic approach in my opinion. At this point we could probably write a short methods manuscript by itself. There is even more time I could sink into this, but I think I found a good working solution for now.

TLDR: Now the final figure should look similar to Saumik’s original and what other publications look like, and the stats should be superior to anything in the literature.

**I love doing this work but you’re giving me déjà vu. We definitely already talked about this before and decided to go with Delta CT because the old technique of calculated fold change in expression was confusing.**

I just spent all evening trying to figure this out because I took a lot of time for the analysis. I think I found the key issue:

CQ or CT is the number of cycles PCR has to run before you reach a detection threshold. This means that a HIGHER CT value indicates a SMALLER starting concentration. Therefore a negative Delta CT indicates a positive effect. Yeesh! This was the key piece of information I was missing. If we want to use Raw DELTA CT, we would have to swap the signs in my figure (I think), but that is an arbitrary solution.

If you do fold change in expression, the analysis becomes a very different question from what I did. This is fine and I know it is a standard technique. For example, very high downregulation and upregulation, both *look* like an increase in fold change expression. This means if an aphid with virus increases JA products, we should see a value above 1. If it decreased, we should see a value between 0-1.

<https://bitesizebio.com/24894/4-easy-steps-to-analyze-your-qpcr-data-using-double-delta-ct-analysis/>

2^-Delta Delta CT might make some sense to me now, but it’s still going to be confusing to a non-molecular audience so we have to be careful. To use a metaphor, Dave, this is like computing interaction strength for multiple species impacting a focal population. Positive effects indicate mutualism, negative effects indicate antagonism. Then we do a transformation, and positive effects are scaled 1-infinity, and negative effects are 0-1.

An “easy“ solution (that I like) is to just transform my current analysis by taking all values (x) that went into the MANOVA by 2^-x. This doesn’t mean doing the delta delta cq, but that’s ok according to my preliminary reading. <http://www.gene-quantification.net/livak-2001.pdf>

It’s fold change compared to a housekeeping gene, and you can compare across multiple treatments. Again, to use an ecology metaphor, it’d be like skipping interaction strength calculations and just plotting control and treatment effects.

I will think about it and get back to you guys when I get a chance to redo the “easy solution” first.

Yes, its 1AM. I couldn’t sleep until I’ve figured this out.