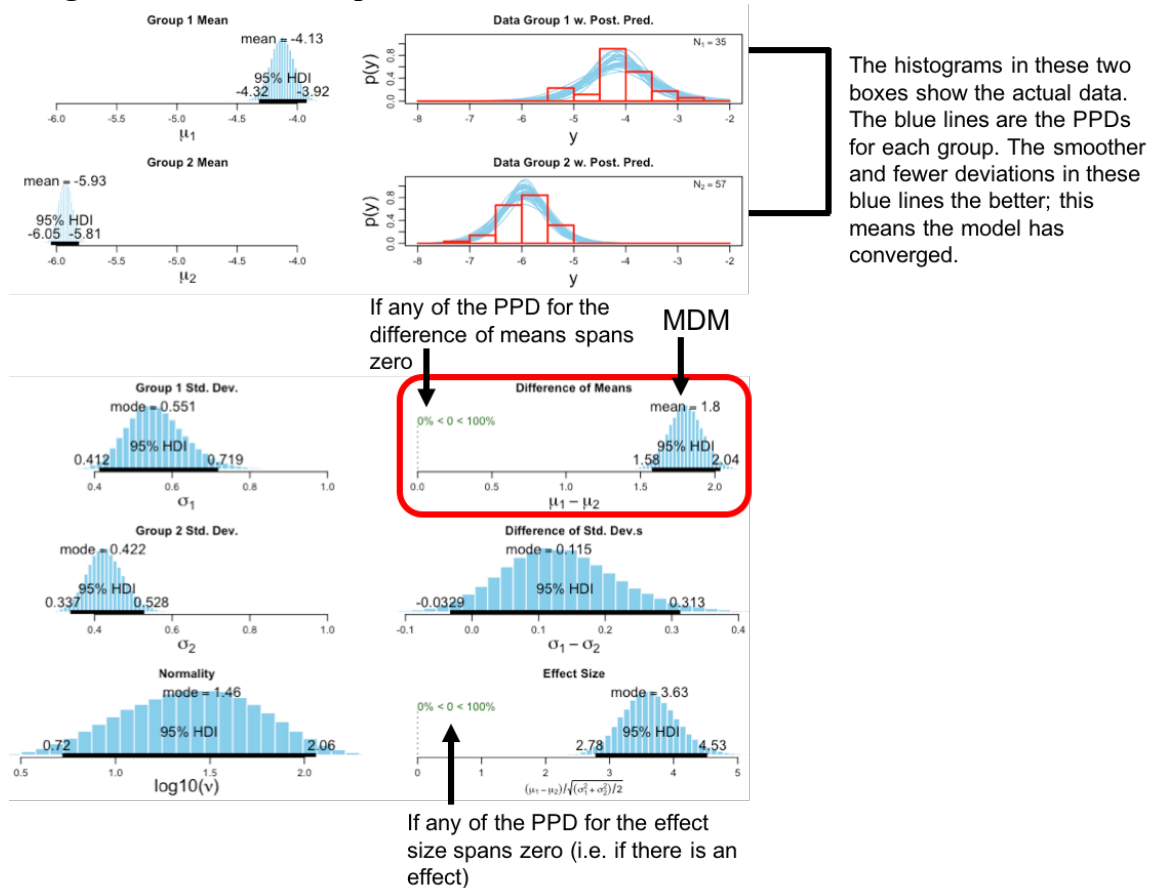**Brief guide on how to interpret BEST Tests:**
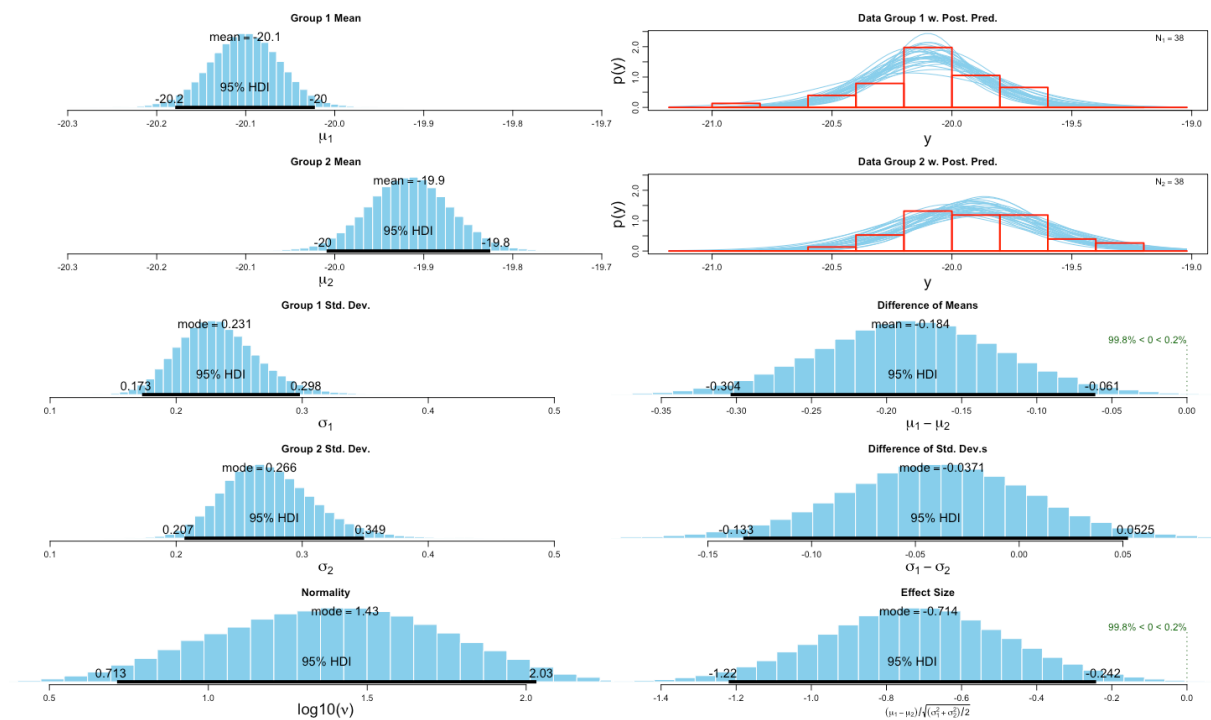

Supplementary Figure 1:Example BEST test output with explanatory notations.

What aids us in group comparisons in a BEST test are their respective PPD (posterior predictive distribution) means shown in the top left, and the mean difference of means (referred in text as MDM) which is the average of the group means from the PPD. If the MDM graph (highlighted by the red box in Supplementary Figure 1) spans zero (on the x axis underneath the distribution, also shown helpfully by green text in the graphic) the two groups were found by the model to have the same mean, at least some of the time, so we cannot be confident in any difference between their means. Their PPD standard deviations and difference between these are to be taken the same way. This highlights if the two groups might have large differences in their distributions. The BEST test also looks at the normality of the two groups combined (bottom left) and gives the effect size of any differences between the group means (bottom right) showing the strength and directionality of any trend. For a more thorough explanation please see the R package notation and Kruschke, John K. "Bayesian Estimation Supersedes the t Test." *Journal of Experimental Psychology: General* 142, no. 2 (2013): 573–603. https://doi.org/10.1037/a0029146.
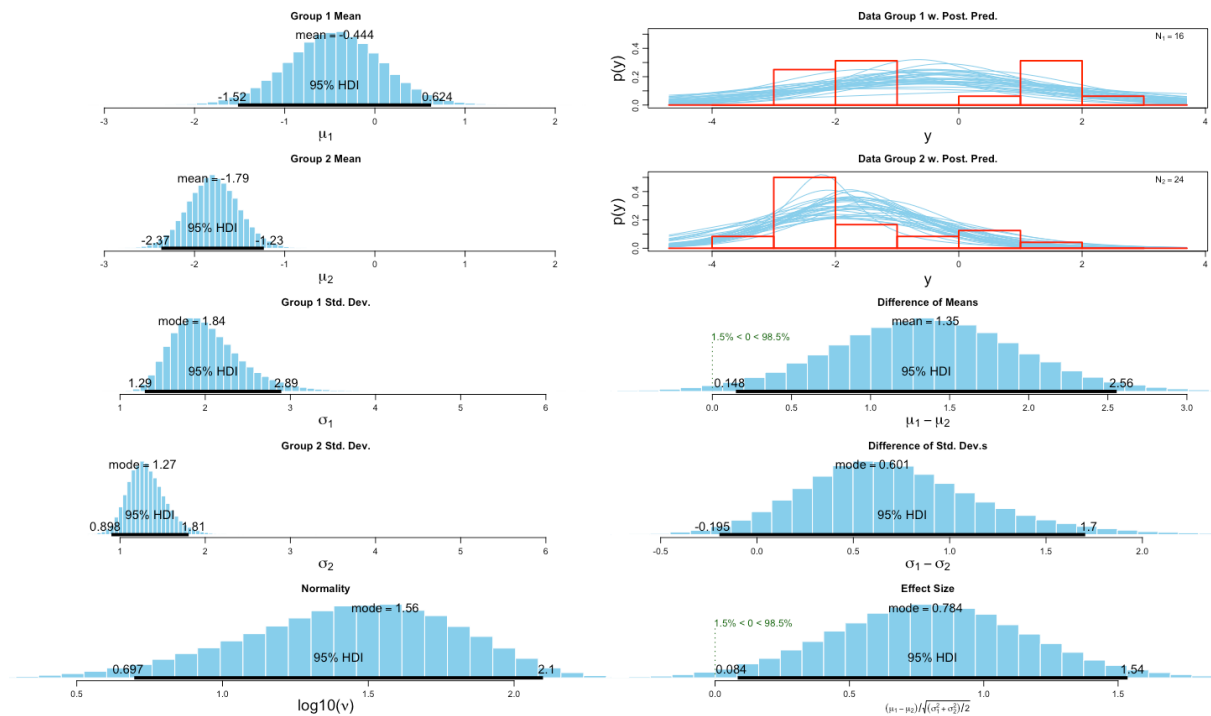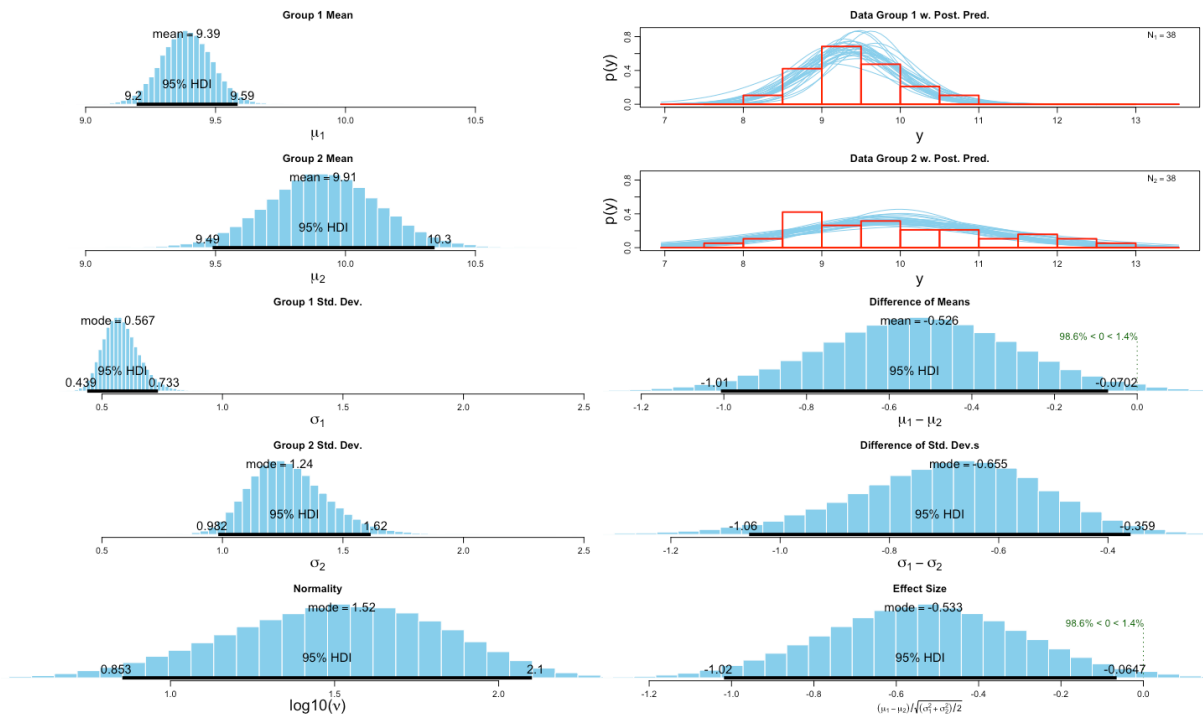
BEST Test Outputs

Supplementary Figure 2: BEST test output for females (group 1) vs males (group 2) for $\Delta^{18}O_{dw\text{-}MAP\ (Chenery)}$ values in Finglesham.

Supplementary Figure 3: BEST test output for matched bone (group 1) vs dentine (group 2) for $\delta^{13}C_{coll}$ values in Finglesham.

Supplementary Figure 4: BEST test output for matched bone (group 1) vs dentine (group 2) for $\delta^{15}N_{coll}$ values in Finglesham.