A**ppendix S1: Measure of effect size**

The classical Cramer’s criterion, is defined by

,

where is a chi-square test statistics, namely a test statistics which follows a chi-square distribution (or asymptotically) when the null hypothesis is true, is the total sample size, is the number of groups being compared, and is the total number of taxa. However, for the test statistics presented in this work, the Cramer’s criterion has the undesired property that it depends on the sample size and number of reads. Therefore, we propose a modified version of Cramer’s, denoted here by , such that its dependency on the sample size and number of sequence reads is nullify,

,

where is the value of the Chi-square statistics for the maximum difference between the taxa frequency means being compared. For example in the case of a two sample mean comparison the maximum difference between the mean of the taxa frequency is achieved when every group has one taxa each which are non-overlapping across groups. Note that for the one-sample and two-sample mean test comparisons it is straightforward to show that when there is not overdispersion, and since, then is equal to the Cramer’s criterion.