# Supplementary File 2

# *Comparative transcriptomics of two temperate C3 bioenergy and forage grasses under water stress*

# Experimental evaluation of RoDEO performance on varying sequencing depths

Here we present results from evaluating the parameter *P* selection for RoDEO. The parameter selection approach is described in the main manuscript. We validated the methodology using human RNA-seq samples (from the PRO dataset), with known DE answers from q-PCR, that were originally used for RoDEO validation (Haiminen *et al* 2014). We evaluated experiments with either 1 or 3 replicates per sample.

In order to simulate the comparison of small to large datasets, we diluted the read counts from original samples in order to reduce the total read count, and to increase the number of genes with zero count. We subsampled down to 10% of the original read counts by dividing the counts by 10 and rounding any values <1 to zero. In an exemplar replicate the number of genes with zero count went from 5,731 to 14,183, out of the total 33,035 human genes.

Parameter *P* was selected as 20 as in the original validation setting, while the parameter *P’* was set to 15 according to the novel methodology, see Figure S2.1 for an illustration of the segment boundaries and cumulative curve of read counts.

The evaluation against q-PCR results was performed as in (Haiminen *et al* 2014), the resulting ROC curves are shown in Figure S2.2 for the case of A and B having each only 1 replicate per sample. Note that the ROC is nearly identical for the comparison of full A vs. full B, and for small A vs. full B or full A vs. small B. Only when small A is compared to small B, the ROC is worse. This demonstrates that even when one sample is very small, the methodology correctly aligns them and is able to yield accurate DE gene detection.



Figure S2.1: Illustration of segment boundaries for P=20 for A,B and P'=15 for subsampled A,B.

The evaluation on samples with 3 replicates is shown in Figure S2.3. Here one of the 3 replicates was subsampled to 10% while other two were kept full size. The ROC are more consistent than in the single replicate case, the performance of detecting DE genes is not hindered by including one small replicate among two other full size replicates.

C:\work\MATLAB\DE\RoDEO_scaling\ROC_k15-20.tif

Figure S2. 2: Evaluation of q-PCR vs. RoDEO DE results on human data with 1 replicate per sample. Full data (blue) and small vs. large dataset (yellow, purple) yield similar ROC.

C:\work\MATLAB\DE\RoDEO_scaling\roc-3rep.tif

Figure S2. 3: Evaluation of q-PCR vs. RoDEO DE results on human data with 3 replicates, 1 of which is small in some comparisons.