Challenge 2 Evaluation **Predictions Ground truth** Entry Entry Available before Entry submission 4 submission 7 submission submission 1 SIRV Set 4 Entry Entry submission 5 Available after submission 2 submission Entry **Entry** Entry ground truth of simulated submission 6 submission 3 submission N data Samples: wTC11 + SIRV4, H1 mix + SIRV4, human simulation, mouse simulation Ground truth given No ground truth H1 mix, WTC11 human simulation, mouse simulation, SIRV4 in each sample Multiple replicates SIRV4 (Log fold change of SIRV = 0) Irreproducibility Consistency Single sample Two conditions Correlation Precision Abundance Recovery Rate (ARR) Recall Single sample Median Relative Difference (MRD) Accuracy Resolution Entropy (RE) Normalized Root Mean Square Error (NRMSE) F1-score ROC and AUC

