(A) Consistency: 10X gene subset A vs. gene subset B GSE130931 GSE150068 GSE163505 GSE158941 $\begin{array}{c} \log(y/s+1)\\ \operatorname{acosh}(2\alpha y/s+1)\\ \log(y/s+1/(4\alpha))\\ \log(\operatorname{CPM}+1)\\ \log(\operatorname{CPM}+1)\\ \log(y/s+1)\rightarrow \operatorname{HVG}\\ \log(y/s+1)\rightarrow \operatorname{HVG}\rightarrow \operatorname{Z} \end{array}$ Pearson sctransform **GLM Residuals** Analytic Pearson Random Quantile Pearson (no clip) Pearson→HVG Pearson→Z Pearson→HVG→Z Expr. Sanity MAP Sanity Distance Dino Latent E Normalish GSE164017 GSE178765 GSE179714 GSE179831 GSE184806 $\begin{array}{c} \log(y/s+1)\\ \operatorname{acosh}(2\alpha y/s+1)\\ \log(y/s+1/4\alpha))\\ \log(\operatorname{CPM}+1)\\ \log(\operatorname{CPM}+1)\\ \log(y/s+1)\rightarrow \operatorname{HVG}\\ \log(y/s+1)\rightarrow \operatorname{HVG}\rightarrow \operatorname{Z} \end{array}$ **GLM Residuals** sctransform Analytic Pearson Random Quantile Pearson (no clip) Pearson→HVG Pearson→Z Pearson→HVG→Z Latent Expr Sanity MAP Sanity Distance Dino Normalisr 300 10 10 20 k-NN Overlap (B) Simulation: Ground truth vs. simulated counts Linear Walk Random Walk scDesign2 Dyngen muscat $\begin{array}{c} \log(y/s+1)\\ \operatorname{acosh}(2\alpha y/s+1)\\ \log(y/s+1)(4\alpha)\\ \log(\operatorname{CPM}+1)\\ \log(y/s+1)>1\\ \log(y/s+1)\to \operatorname{HVG}\\ \log(y/s+1)\to \operatorname{DG}(y/s+1)\to \operatorname{AVG} \to \operatorname{CMG}(y/s+1) \to \operatorname{AVG}(y/s+1) \to \operatorname{AVG$





