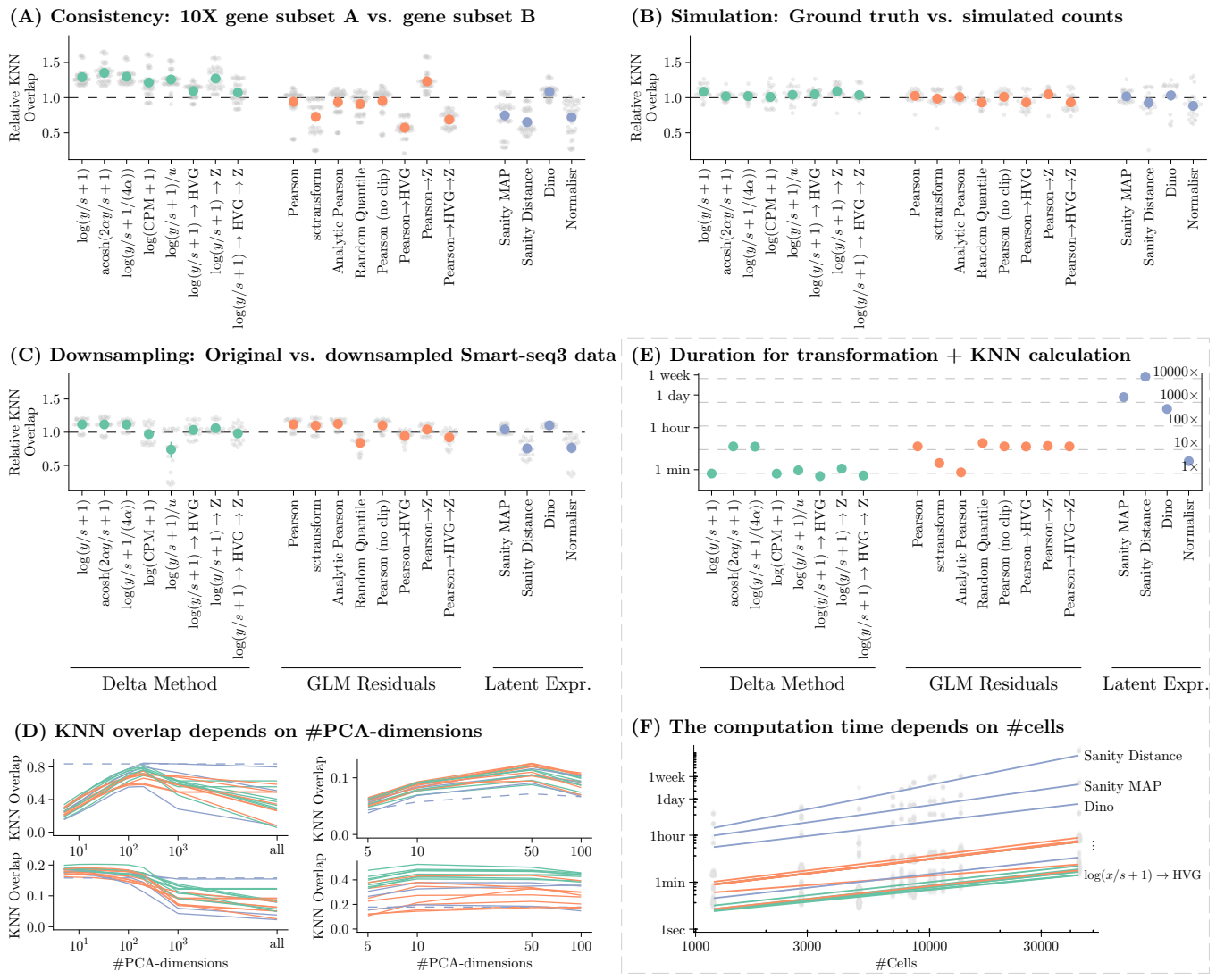
(A) Consistency: 10X gene subset A vs. gene subset B
(B) Simulation: Ground truth vs. simulated counts
(C) Downsampling: Original vs. downsampled Smart-seq3 data
(D) KNN overlap depends on #PCA-dimensions
(E) Duration for transformation + KNN calculation
(F) The computation time depends on #cells
Relative KNN Overlap
KNN Overlap
log(y/s + 1)
acosh(2αy/s + 1)
log(y/s + 1/(4α))
log(CPM + 1)
log(y/s + 1)/u
log(y/s + 1) → HVG
log(y/s + 1) → Z
log(y/s + 1) → HVG → Z
Pearson
sctransform
Analytic Pearson
Random Quantile
Pearson (no clip)
Pearson→HVG
Pearson→Z
Pearson→HVG→Z
Sanity MAP
Sanity Distance
Dino
Normalisr
Delta Method
GLM Residuals
Latent Expr.
#PCA-dimensions
#Cells
1 week
1 day
1 hour
1 min
1sec
10000×
1000×
100×
10×
1×
Sanity Distance
Sanity MAP
Dino
log(x/s + 1) → HVG