A

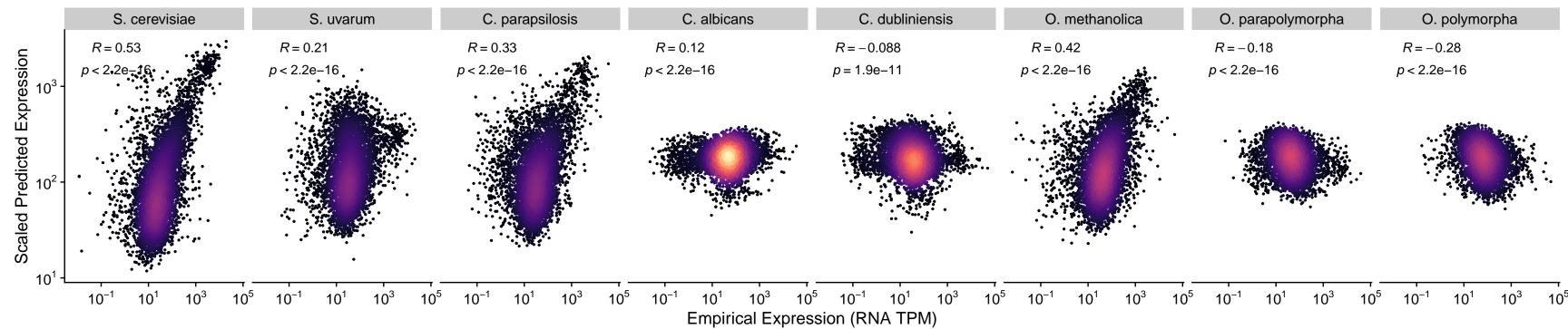**B** ROC−SEMPPR Predictions of Gene Expression
Compared to Empirical RNA−seq Data

S. cerevisiae: $R = 0.53$, $p < 2.2e{-}16$
S. uvarum: $R = 0.21$, $p < 2.2e{-}16$
C. parapsilosis: $R = 0.33$, $p < 2.2e{-}16$
C. albicans: $R = 0.12$, $p < 2.2e{-}16$
C. dubliniensis: $R = -0.088$, $p = 1.9e{-}11$
O. methanolica: $R = 0.42$, $p < 2.2e{-}16$
O. parapolymorpha: $R = -0.18$, $p < 2.2e{-}16$
O. polymorpha: $R = -0.28$, $p < 2.2e{-}16$

Scaled Predicted Expression

Empirical Expression (RNA TPM)

C Distribution of GC3% Across Genes

GC3

S. cerevisiae: 0.08
S. uvarum: 0.17
C. parapsilosis: 0.03
C. albicans: 0.13
C. dubliniensis: 0.16
O. methanolica: 0
O. parapolymorpha: 0.16
O. polymorpha: 0.16

Gene Cluster: Genome, Lower GC3%, Higher GC3%