SUPPLEMENTAL FIGURES AND LEGENDS

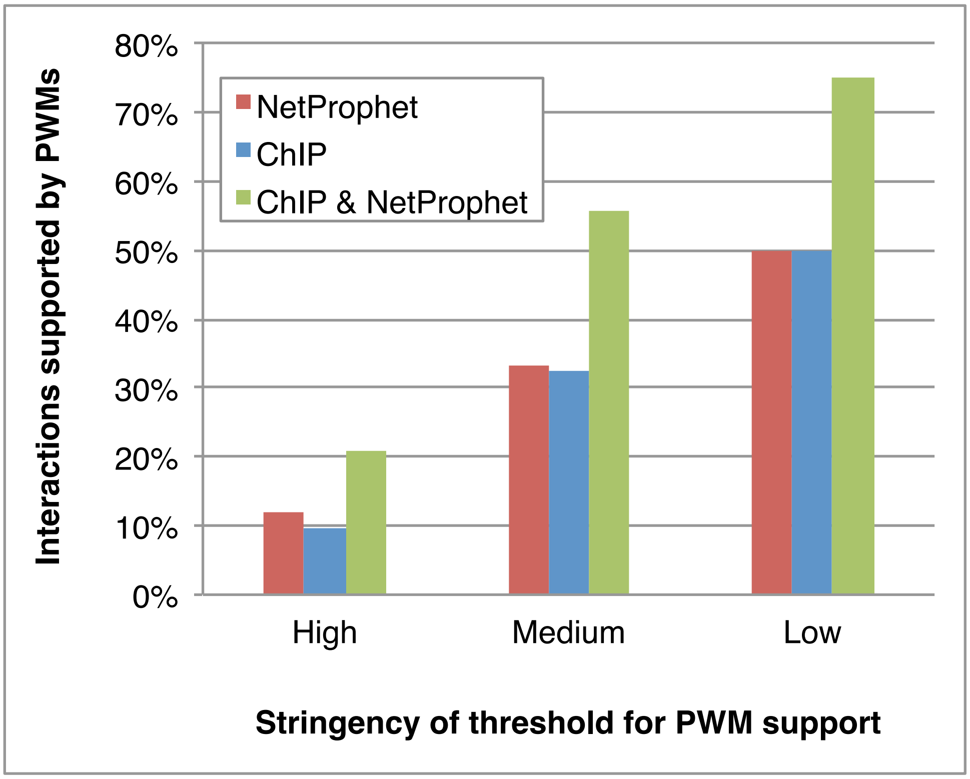

Figure S1. Same as Figure 1 except using PWMs taken from the ScerTF database (Spivak & Stormo, 2012). The PWMs in this curated database are derived primarily from *in vivo* data including ChIP. For each TF, we used the PWMs recommended by ScerTF on the basis of objective predictive power and manual curation. For the high, medium, and low stringency PWM cutoffs, chance inclusion was 5.7%, 20.4%, and 39.5%, respectively.

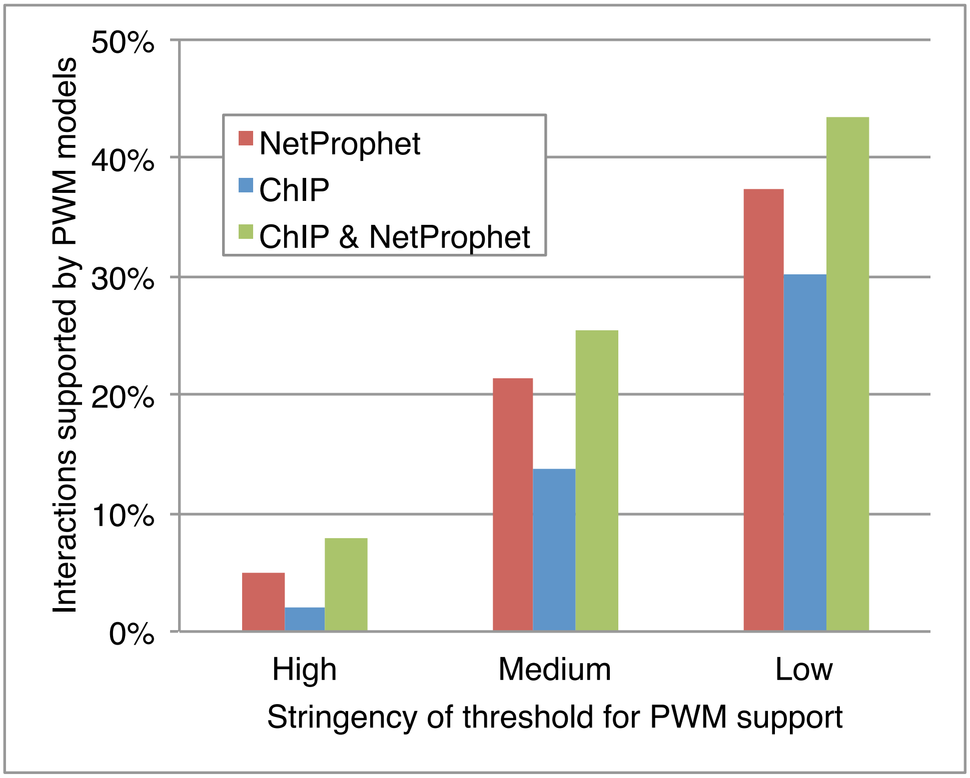

Figure S2. Conserved TF-promoter binding potential for the top 4,000 NetProphet predictions (red), all direct targets implicated by ChIP hits in the Yeastract or Tnet databases (blue), and targets implicated by ChIP hits that are also predicted by NetProphet (green). The supported interactions are those whose binding potential scores meet or exceed the stringency threshold in *S. cerevisiae* and in at least two of three other *sensu strictu* species (*S. bayanus*, *S. mikatae* and *S. paradoxus*). Support within each species was determined using the same methods and thresholds as in Figure 1.

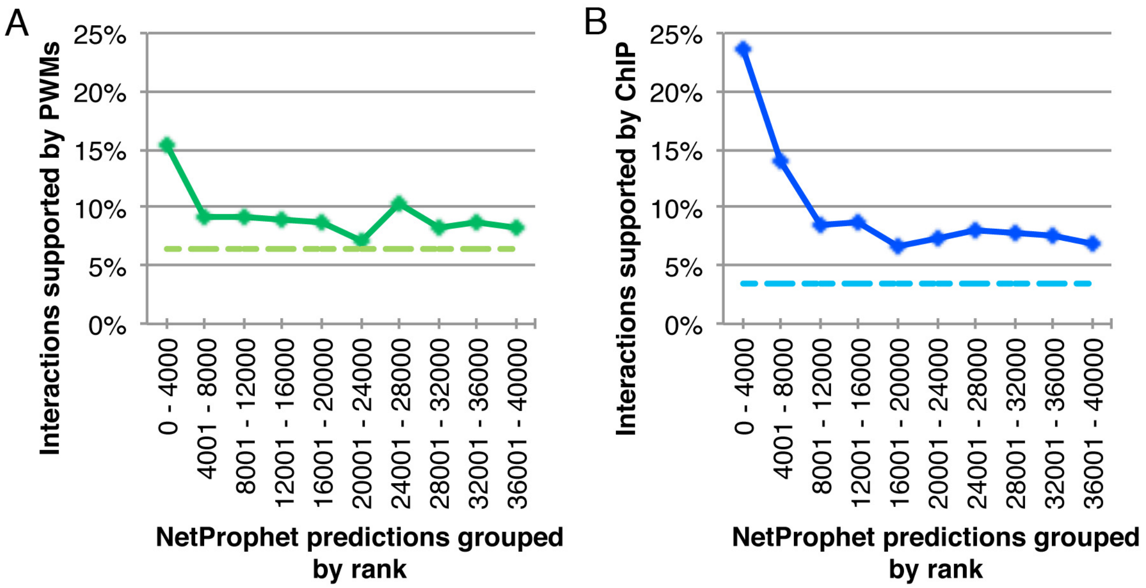

Figure S3. Same as Figure 2 except that additional expression profiles from stress conditions and overexpression of 55 TFs are included in the analysis along with the TF deletions. LASSO regression was carried out on each of the three data sets separately and the resulting regression coefficients were averaged. For differential expression analysis, data from TF overexpression was used only for those TFs for which no deletion data were available. Several other methods of combining the data were tried but all were less accurate.

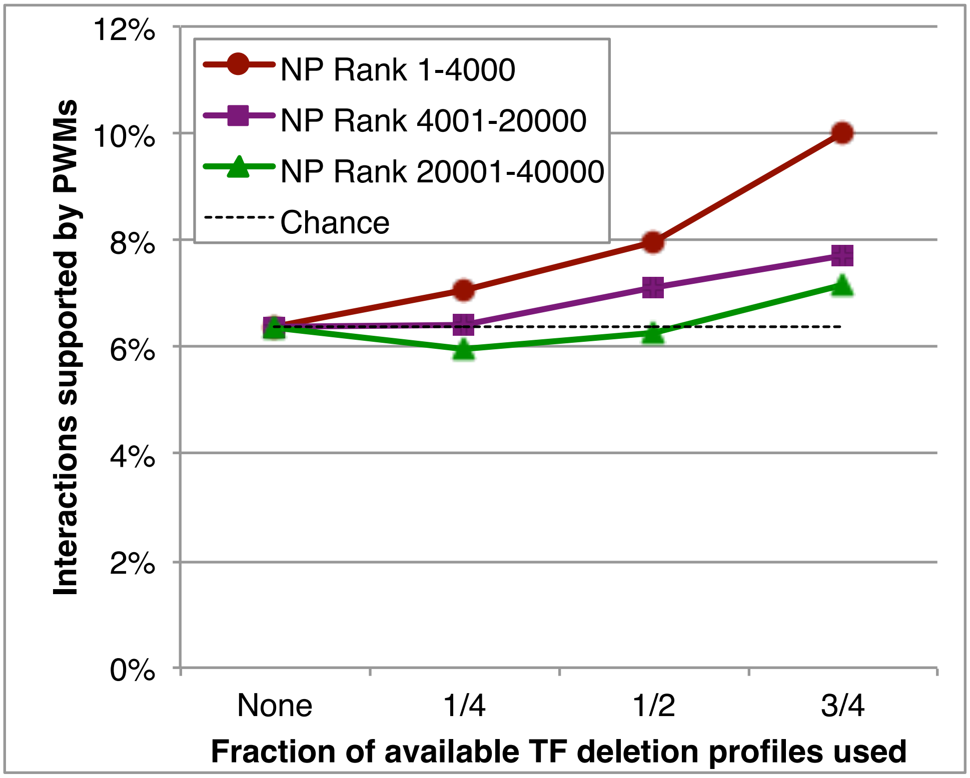

Figure S4. Rate of PWM support for NetProphet predictions on targets of TFs whose deletion profiles are not included in the data, as a function of the number fraction of the total deletion profile set used.

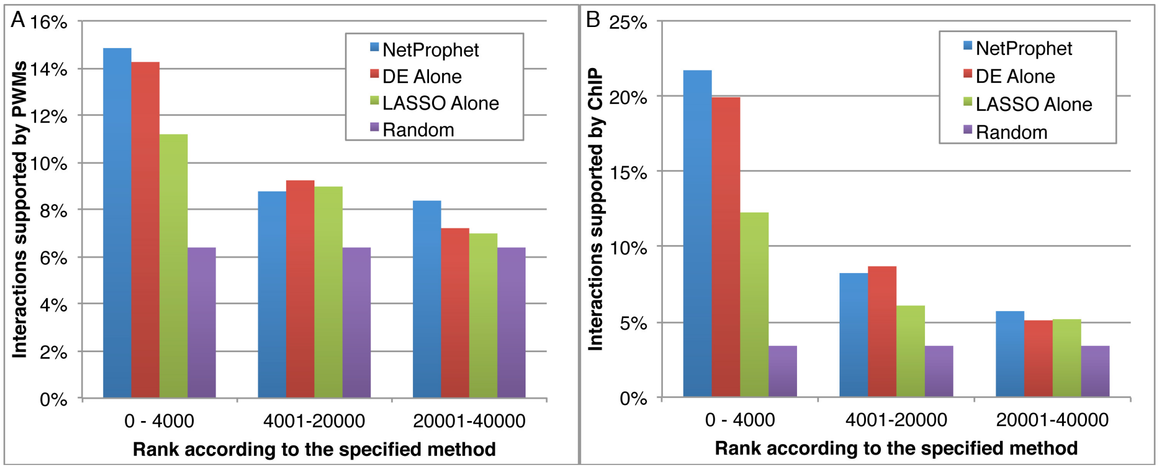

Figure S5. Accuracy of NetProphet and its components in various rank ranges according to PWM support (A) and ChIP support (B).

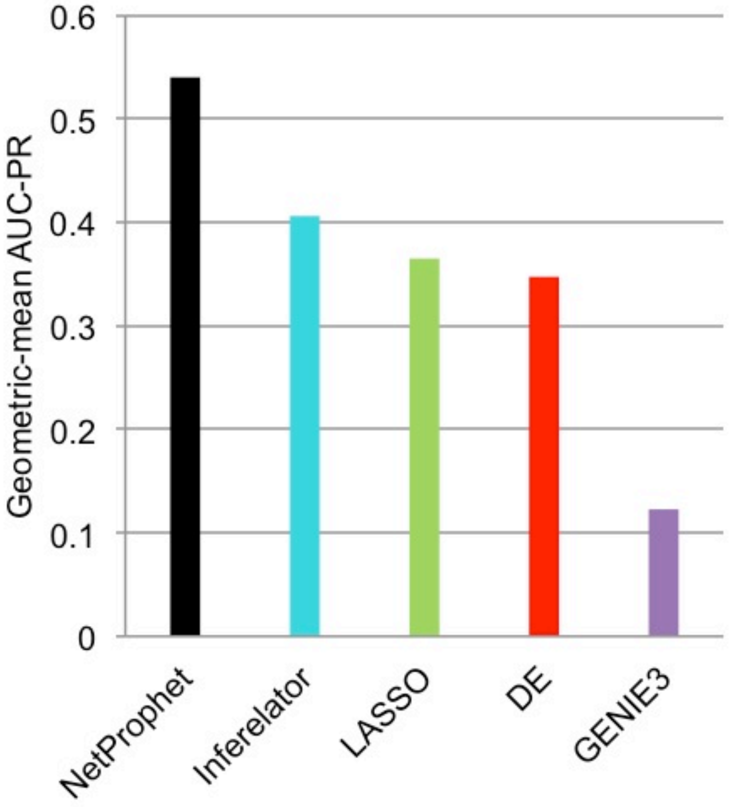

Figure S6. Evaluation of reconstruction accuracy on the 5 100-gene networks from DREAM4. NetProphet (black) is compared against the two baseline methods LASSO (green) and DE rank (red) as well as Inferelator (turquoise) and GENIE3 (purple). Accuracy is summarized as the geometric mean of the area under the precision recall curve statistic (AUC-PR) for the 5 networks.