Figure S2 (A-B). Part of a sub-model taken from the *Interaction* model for CEBPB in H1hesc and Helas3 respectively. Each node in the tree is labeled with the TRANSFAC id, corresponding gene name and the threshold at which the feature is split. In each tree, one binding rule is highlighted: they are identical with respect to leading or increasing the probability of leading the binding probability of CEBPB.

Figure S2 (C). Accuracy (ROC-AUC) distribution for 6 choices of EMT feature sets. "<*" (">*") denotes significant difference (one sided Wilcoxon p-value < 0.05) between the two sets of performances and the direction (greater or less). The plot has same color coding as in Fig 1C-D, i.e. same color is used to denote the models using same features.
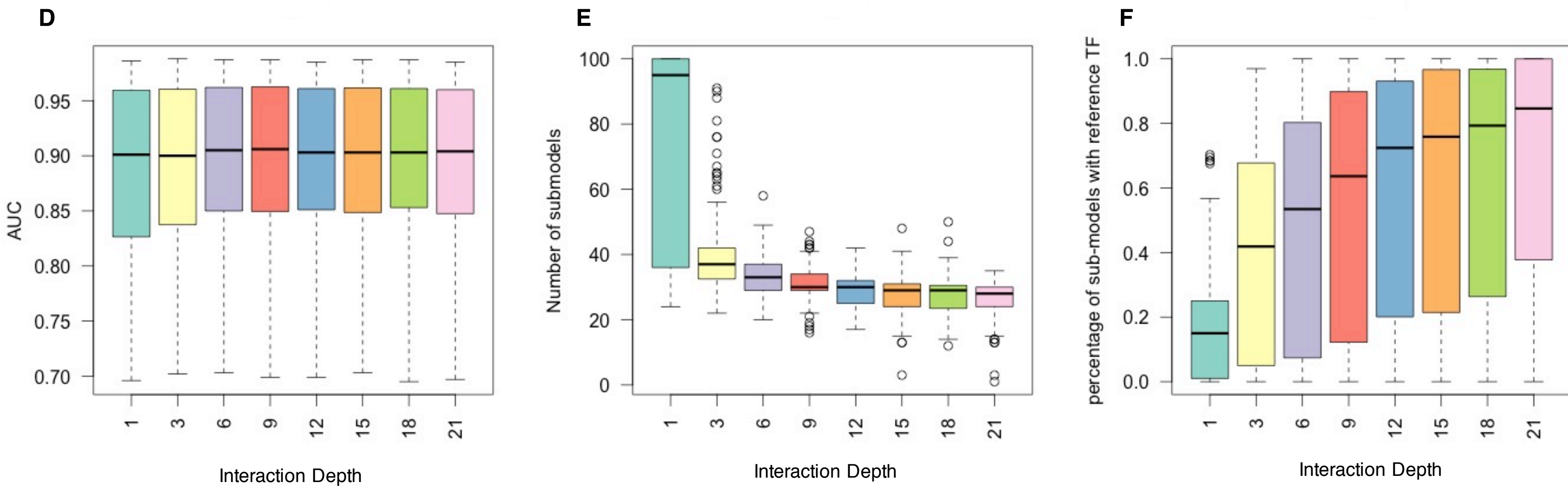

Figure S2 (D-F). ROC-AUC (D), number of sub-models (E), and fraction of sub-models (F) that included reference TFs, for different interaction depth of the models. Interaction depth defines maximum allowable features in a sub-model. The same colors denote model interaction depth for plots D-F.
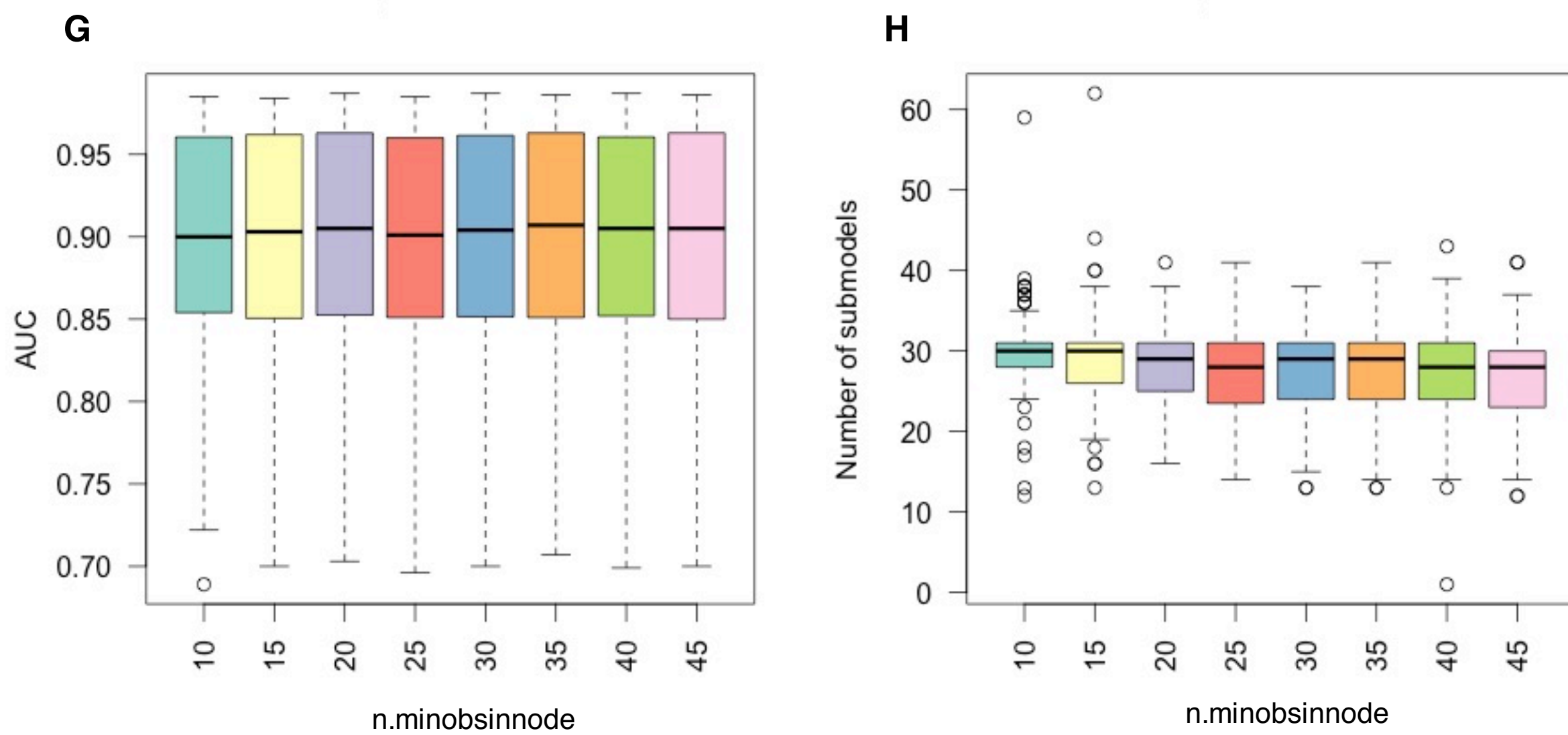

Figure S2 (G-H). ROC-AUC (G), and number of sub-models (H) for different n.minobsinnode of the models. 'n.minobsinnode' denotes minimum number of observations made at each node while building the decision tree. Plots G and H use same the colors to indicate models using 'n.minobsinnode'.
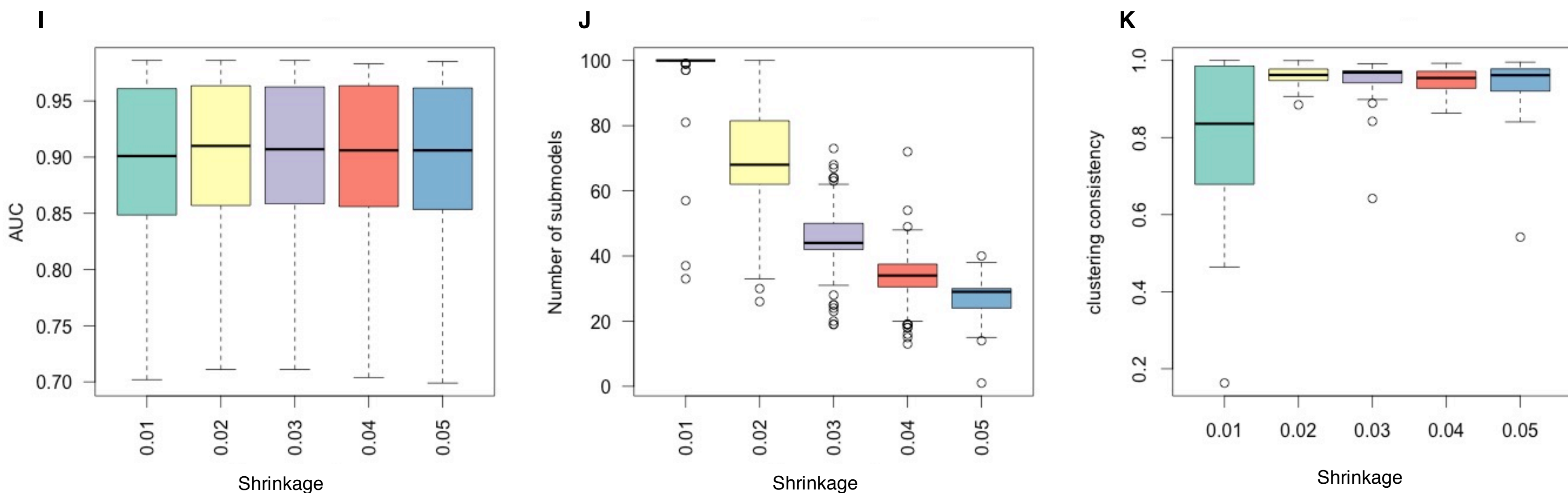

Figure S2 (I-K). ROC-AUC (I), number of sub-models (J), and clustering consistency (K) of the sequences (percentage of sequence-pairs that fall in same clusters) for different values of the shrinkage parameter. Shrinkage indicates the learning rate of the model. Plots I-K the use same colors to denote models using same learning rate.
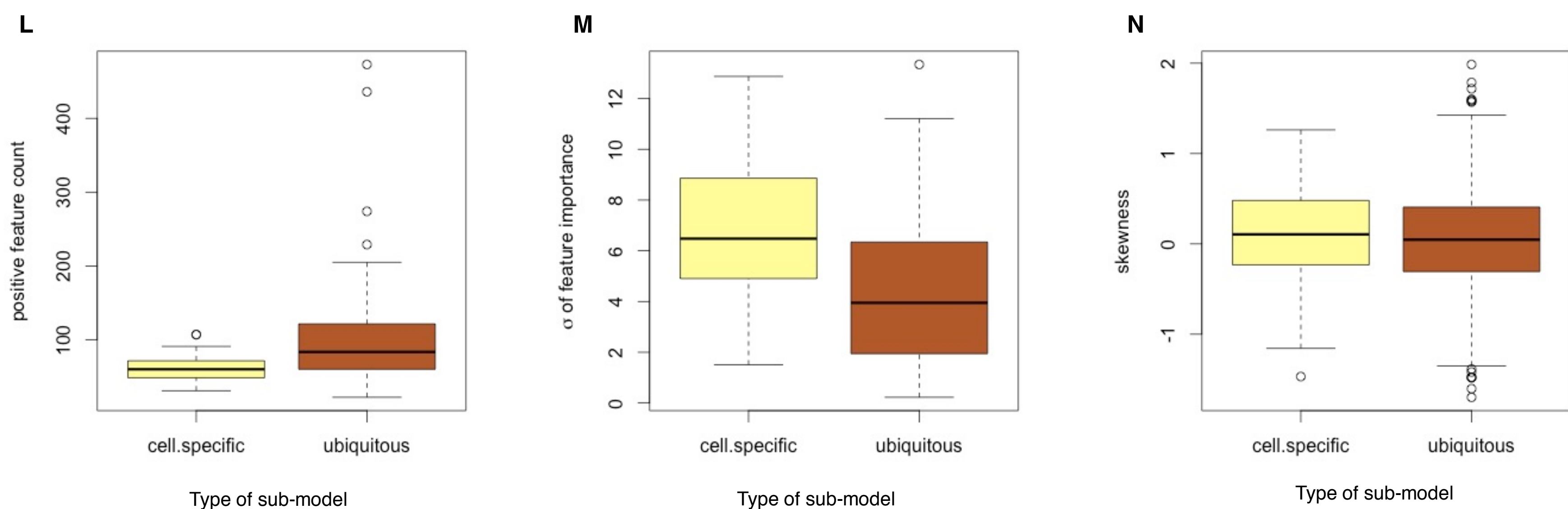

Figure S2 (L-M). Comparison of ubiquitous and cell type-specific sub-models: (L) Number of relevant features (i.e. features with non-zero importance in any cell type-specific model), (M) Standard deviation of feature importance for each sub-model, (N) Skewness of gene expression in each cells for the co-factors. Yellow and brown colors denote ubiquitous and cell type-specific sub-models respectively.