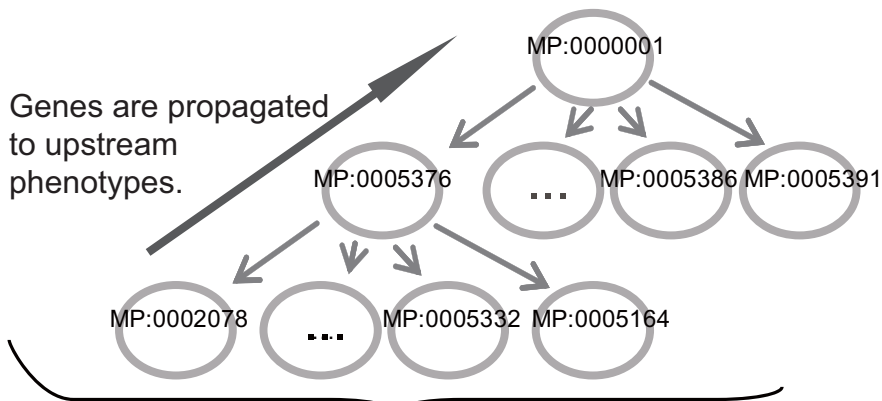
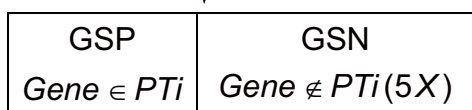


A

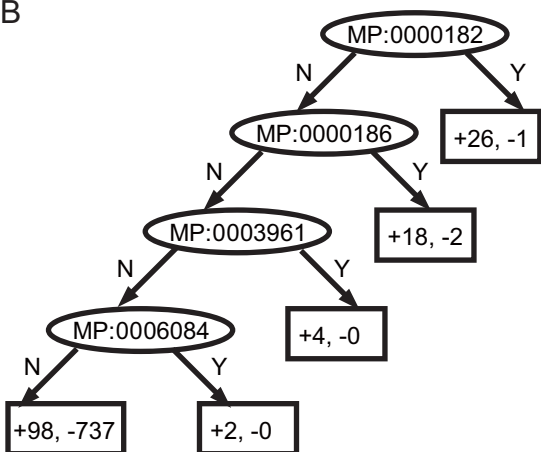
MGI mammalian phenotype ontology



Select testable phenotypes

1) associated ≥ 10 genes; 2) keep lowest leaf nodesFor each phenotype (PT_i) of a disease (D)Train a decision tree
(Exemplary tree in Figure 1B)Repeat $N = 100$ times
to get stable resultFor each leaf in decision tree
Probability = $\#$ of GSP / ($\#$ of GSP + $\#$ of GSN)Average probability of a gene perturbation to
generate phenotype PT_i : $S_i = \sum Probability / N$ Down weight of frequent
phenotypes: $w_i = -\log_{10}(f_i)$ Likelihood of a gene perturbation to give rise to disease D : $\frac{\sum (w_i * S_i)}{\sum w_i}$

B



C

