

Structure of PANDA system

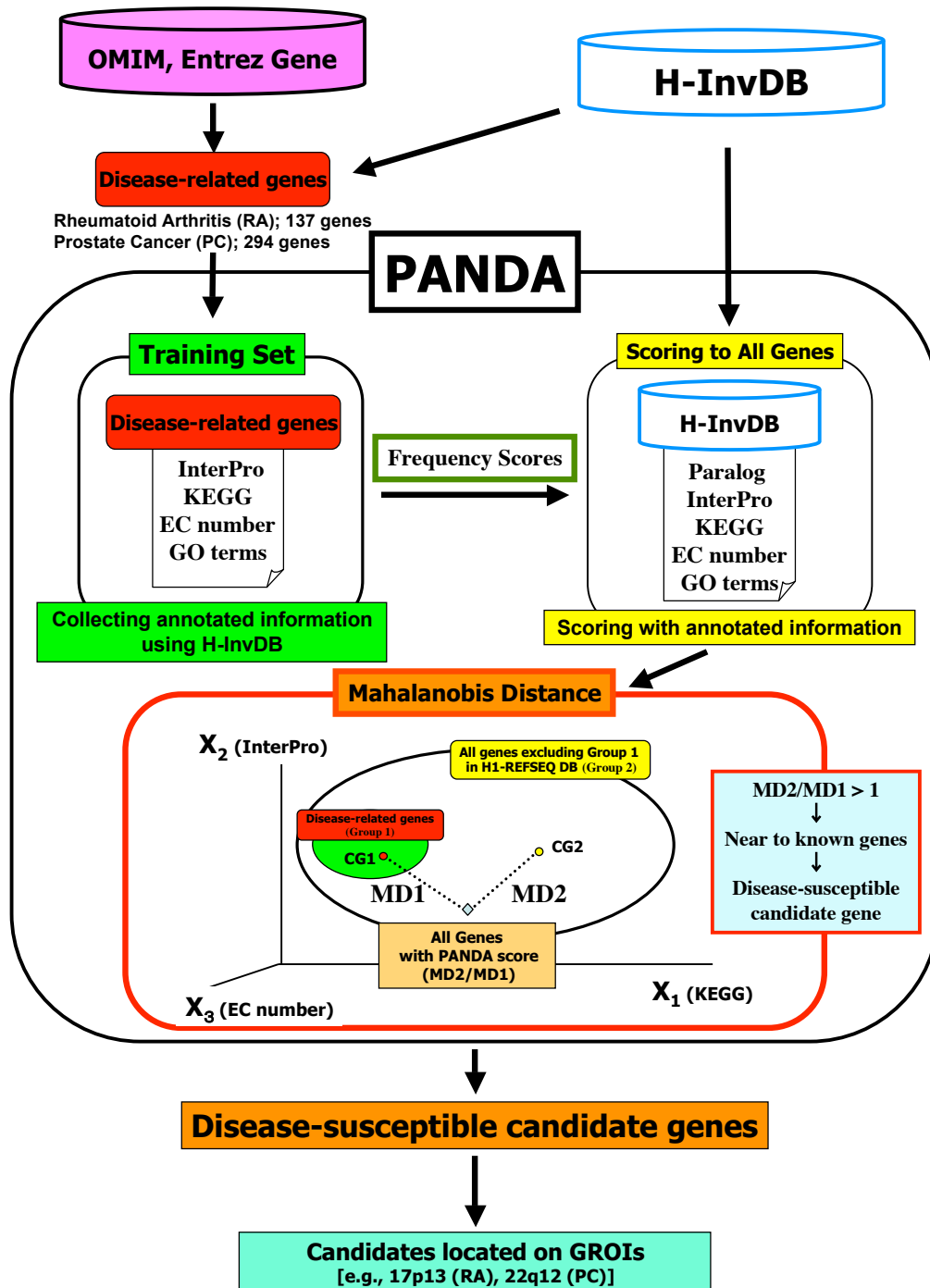


Figure Legend

At the top-right, the blue-lined column indicates H-InvDB (<http://www.h-invitational.jp/>). We used this database as a source of gene data. At the top-left, disease-related genes (red box) were selected from H-InvDB with OMIM and Entrez Gene (pink column). After that, we compared frequency of each of functional information between all genes and a training set (middle-left). With these frequency scores, each of functional information was converted into score using some expressions (middle-right). Next, we prioritized all genes in H-InvDB using Mahalanobis distance (middle-bottom). First, we set "Disease-related genes (= Group 1, green circle)" and "all genes excluding Group1 (= Group 2)". Second, we calculated each center of gravity of Group 1 (CG1, red circle) and Group 2 (CG2, yellow circle) using 7 parameters (X_1 to X_7). Third, we calculated Mahalanobis distances (MD) from the position of each gene to CG1 and CG2, called MD1 and MD2. Finally, we calculated division of MD1 and MD2 (MD2/MD1) and we considered a gene with a score of MD2/MD1 greater than 1, to be a disease-susceptible candidate. Finally, we identified the possible candidates to target diseases and mapped these candidates onto genetic regions of interest (bottom).