

Figure S2. Scheme of Prediction for Functional Annotation

(A) Schematic diagram for determining a representative transcript for each locus.

The procedure of computational autoannotation is illustrated. Prior to the human curation of the representative transcript of each H-Inv cluster, we performed computational autoannotation.

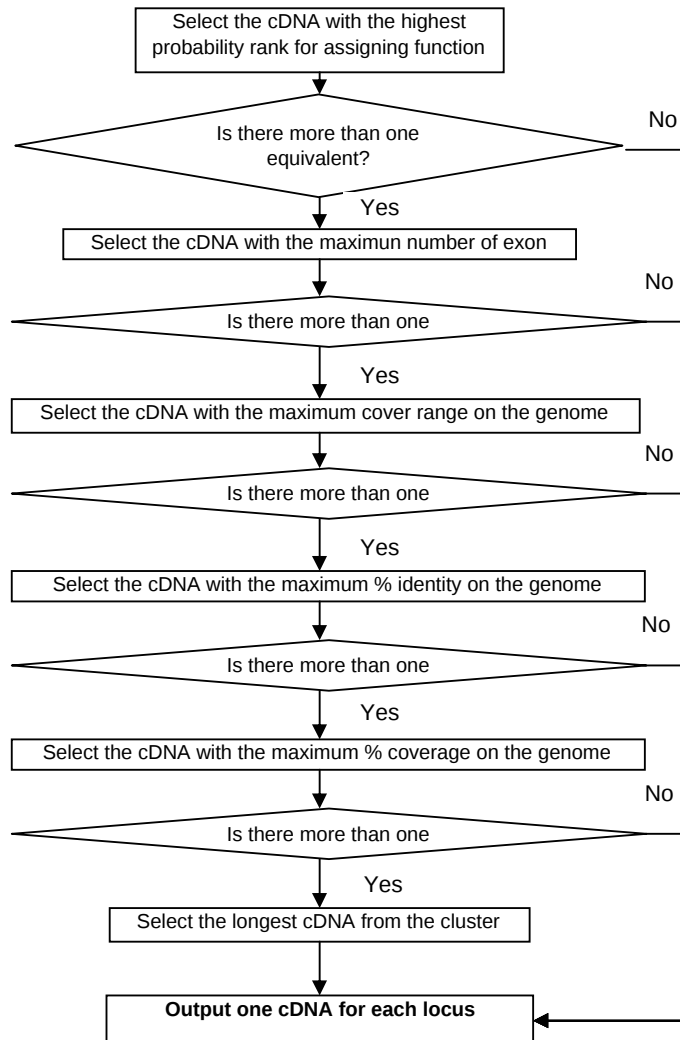


Figure S2. Scheme of Prediction for Functional Annotation

(B) Schematic diagram for functional prediction of H-Inv proteins.

This schematic diagram illustrates the H-Inv autofunctional annotation pipeline that can determine the most appropriate data source ID, avoiding the following keywords that suggest proteins without experimental verification in the description; (1) hypothetical, (2) similar to, (3) names of cDNA clones (Rik, KIAA, FLJ, DKFZ, HSPC, MGC, CHGC, and IMAGE) and (4) names of InterPro domain frequent hitters.

