

## Figure S8. A Sample View of the H-Invitational Database (H-InvDB; <http://www.h-invitational.jp/>)

A FLcDNA (BC003551) is shown with its detailed annotations, e.g., gene structure, functional annotation, ORF predictions, protein structure prediction by GTOP, etc. The H-InvDB has links to other internal databases (red boxes) such as a genome map viewer (G-integra) and gene expression library (H-Angel). Green boxes show internal viewers for the results of clustering (Clustering Viewer showing results by H-Inv, STACK, TIGR, UniGene, etc.), the prediction of subcellular localization (TOPOViewer showing results of TMHMM, SOSUI, TargetP, and PsortII), and the disease-related information (DiseaseInfo Viewer linking to OMIM and GenAtlas). The H-InvDB also has links to many external public databases (black boxes), including DDBJ/EMBL/GenBank, RefSeq, UniProt/Swiss-Prot and TrEMBL, Genew, InterPro, 3D Keynote, Ensembl, GeneLynx, LocusLink, PubMed, LIFEdb, dbSNP, GO, and GTOP, and to homepages by original data producers of FLcDNA clones and sequences (blue boxes), including the Chinese National Human Genome Center (CHGC), the Deutsches Krebsforschungszentrum (DKFZ/MIPS), Helix Research Institute (HRI), the Institute of Medical Science at the University of Tokyo (IMSUT), the Kazusa DNA Research Institute (KDRI), the Mammalian Gene Collection (MGC/NIH), and the FLJ project.

The screenshot displays the H-InvDB cDNA view for the gene **HIT000031589**. The interface is annotated with various database links categorized by color:

- Clustering Viewer (Green):** STACK / SANBI, TIGR, UniGene.
- G-integra (Red):** Genome map viewer.
- DiseaseInfo Viewer (Red):** OMIM, GenAtlas.
- H-Angel (Red):** Gene expression library.
- TOPOViewer (Green):** TMHMM, SOSUI, Target P, Psort II.
- External Databases (Black):** DDBJ/EMBL/GenBank, RefSeq / protein, UniProt / Swiss - Prot / TrEMBL, Genew / HUGO, PubMed, KEGG, LocusLink, 3-D Keynote, InterPro, GO, dbSNP, Ensembl, GeneLynx, RefSeq / locus, MutationView, GTOP, LIFEdb.
- Original Data Producers (Blue):** CHGC, DKFZ/MIPS, HRI, IMSUT, KDRI, MGC/NIH, FLJ.

The central panel shows detailed annotations for the gene, including its function (Protein-glutamine gamma-glutamyltransferase), mapping information, gene expression profile, disease information, and predicted secondary and tertiary structures.