orphaNews Europe
Newsletter of the Rare Diseases Task Force

Below is some information on the gene databases Orphanet has integrated into its knowledgebase:

OMIM is the online version of the original database of genes and Mendelian phenotypes that was created by Pr. Victor McKusick in the 1960s. The online version is consistently updated and is maintained by NCBI.

Genatlas contains relevant information with respect to gene mapping and genetic diseases. The database was created in 1986 by Pr Jean Frézal and is located at the Paris Descartes University. The information is collected from the literature. Data entered in Genatlas are updated daily and immediately available online. A gene or a disease is included in Genatlas if its location is supported by either a significant lod-score (>3) or relevant physical mapping. Genatlas allows fluent navigation between its Gene database and its Phenotype database.

HGNC is a committee jointly funded by the US National Human Genome Research Institute (NHGRI) and the Wellcome Trust (UK). It operates under the auspices of HUGO, with key policy advice from an International Advisory Committee, and is responsible for approving gene names and symbols (short-form abbreviations). All approved symbols are stored in the HGNC database. Each symbol is unique and each gene is only given one approved gene symbol. Over 25,000 symbols have been approved so far; the vast majority of these is for protein-coding genes, but also includes symbols for pseudogenes, non-coding RNAs, phenotypes and genomic features.

SwissProt is a manually annotated protein knowledgebase established in 1986 and maintained since 2003 by the UniProt Consortium - a collaborative effort between the Swiss Institute of Bioinformatics (SIB) and the Department of Bioinformatics and Structural Biology of Geneva University, the European Bioinformatics Institute (EBI), and Georgetown University Medical Center's Protein Information Resource (PIR). The UniProt Knowledgebase consists of sequence entries, composed of different line-types, each with their own format. For standardisation purposes, the format of the UniProt Knowledgebase follows as closely as possible the EMBL Nucleotide Sequence Database format. The sequence entries are annotated with the following information: Function(s) of the protein; Post-translational modification(s); Domains and sites; Secondary structure; Quaternary structure; Similarities to other proteins; Disease(s) associated with deficiencie(s) in the protein; Sequence conflicts, variants, etc.