

InterMine: Integrative Analysis Across Multiple Organisms

InterMine is an open source data warehouse built specifically for the integration and analysis of complex biological data. Integrative analysis is a powerful approach in modern biology that allows knowledge from many sources of evidence to be analysed together. Such an approach has many advantages, including less false positives, compensation for missing or unreliable information and interrogation at different levels of genetic, genomic or proteomic regulation leading to the discovery and understanding of complex biological systems. An important step in understanding such processes is also the interpretation of complementary data from model organisms. There is a broad selection of InterMines worldwide, covering many organisms, including HumanMine, PhytoMine (over 87 plant genomes), the Legume federation InterMines (Chickpea, Soy, Legume, Peanut, Bean), MedicMine (Medicago), ThaleMine (Arabidopsis) as well as the budding yeast, rat, zebrafish, mouse, fly and nematode model organisms.

The InterMine framework includes a user-friendly web interface as well as a powerful web service API, with multiple language bindings, allowing programmatic access to data. A number of search and analysis interfaces are included, ranging from a keyword search to a sophisticated, advanced query builder and interactive results tables that enable powerful exploration of data. A set of graphical analysis tools provide a rich environment for data exploration including statistical enrichment of sets of genes. Apps for both iOS and Android allow InterMine access “on the go”. The combination of search and analysis features allows powerful iterative querying and workflows to be created. The InterMine interface has recently been fully re-designed to make use of the latest technologies, providing an improved user experience and enabling easier integration of additional tools and analysis features. In addition, InterMine is working towards enabling the FAIR data principles to enhance the findability and usability of the data.