THE UNIVERSAL PROTEIN RESOURCE

Proteins are vital to all living processes and are the functional units required to sustain life. Enzymes, hormones and antibodies are all proteins. Proteins are composed of 20 amino acids which can be arranged in millions of different ways (sequence) to create millions of different proteins, each one with a specific function.

UniProt is the comprehensive resource for proteins. It brings together all available information for a given protein, providing analysis, comparisons as well as a critical review of associated experimental and predicted data. It is the central hub for proteins, consolidating data from many sources, adding value and providing essential information for many other resources to use. It acts as a fundamental resource for research in biology and molecular medicine.

Collaborative Effort

UniProt is a collaboration between the European Bioinformatics Institute (EMBL-EBI), the Swiss Institute of Bioinformatics and the Protein Information Resource (PIR). Across the three institutes, about 100 people are involved in different tasks such as database curation, software development and user support.

IMPACT

Information into Knowledge

By unifying the available data for a given protein and adding expert annotation, UniProt is a vital tool for researchers to turn information into knowledge and knowledge into the next research project.

Networked Discovery

UniProt is connected to 138 other well established biological data resources. By linking knowledge to a comprehensive and diverse set of biological databases, UniProt provides a networked resource for everything that is known about a protein and its wider context. It is known that a database linked to UniProt has increased web traffic which demonstrates that UniProt improves the discoverability of relevant information by scientists.

Fundamental to Research

UniProt is used by hundreds of thousands of researchers worldwide each month. UniProt is cited by scientists in thousands of patents and publications because the usage of UniProt identifiers is the de facto standard for proteins in the bioscience community worldwide. This standardisation provides a fundamental platform on which new research can base its knowledge, saving duplication and misinterpretation and ensuring successful outcomes in the essential research required to solve a wide variety of challenges facing human disease and industrial applications.

High Quality

UniProt is known worldwide for providing gold standard data of an exceptionally high quality which is essential for all the scientists and other resources using UniProt. “The manually curated database UniProtKB/Swiss-Prot shows the lowest annotation error levels (close to 0% for most families)”\(^\text{1}\). This long term and ongoing high quality means that generations of scientists know they can trust the information provided.

SUPPORTING PUBLICATIONS

001 UniProt consortium: UniProt: a hub for protein information Nucleic Acids Research, 2015, 43(Database issue)


004 Charles Beagrie Ltd. 2016 The Value and Impact of The EMBL-EBI