PRESS RELEASE

# THE HUMAN PROTEIN ATLAS

## New version 21 of the Human Protein Atlas – the open access resource for human proteins

**[November 18, 2021]. A new version 21 of the open access Human Protein Atlas has been launched. A lot of new data and content have been added and the resource now includes 10 separate sections with complementary information about all human proteins. All data has been updated on the approximately 15 million individual web pages.**

The Human Protein Atlas consortium has today launched the version 21 of the open access resource for spatial profiling of the human proteins (www.proteinatlas.org). The Ensembl 103 genome release was used for the annotation of all protein-coding genes and a new normalization scheme was developed for all the data sets. A new strategy for dimensionality reduction and density-based clustering of co-expression patterns has been used to explore the gene expression landscape and we present Expression UMAP clustering of all protein-coding genes in four of the sections; (1) Tissue, (2) Single Cell Type, (3) Immune Cells and (4) Cell Lines. The resource contains 10 sections each exploring the human proteins from different angels:

**1. Tissue –** tissue-based map of the human proteome. This section focuses on the expression profiles in human tissues both on the mRNA and protein level. Here it is possible to explore protein localization in tissues at single-cell resolution and to identify which genes/proteins are enriched in particular tissue types (specificity). The genes expressed in each of the tissues can be explored in interactive UMAP plots and bar charts, with links to corresponding immunohistochemical stainings in human tissues

**2. Brain –** protein profiles of different regions of the brain. This section focuses on the expression profiles in different regions of the brain, including human, pig and mouse. Here it is possible to explore protein profiles in different subregions of the brain and if a gene/protein is enriched in a particular brain region (specificity). The data focuses on human genes and one-to-one orthologues in pig and mouse. In this new version, many new microdissected samples from different regions of the human brain are included.

**3. Single Cell Type -** expression profiles in single cell types. This section contains Single Cell Type information based on single cell RNA sequencing (scRNAseq) data. The section is expanded to include single cells from 26 normal human tissue types to allow exploration of the expression of all human protein-coding genes across 444 single cell clusters summarized into 76 single cell types. Here it is possible to explore mRNA and protein expression in single cell types and if a gene is enriched in a particular cell type (specificity).

**4. Tissue Cell Type –** cell type specificity of genes in tissues. This new section is included for the first time, and contains information generated using integrated network analysis of publicly available bulk RNAseq data for cell type expression specificity predictions for all human protein coding genes. A specificity classification is used to predict which genes are enriched in each constituent cell type within an individual tissue type. Here it is possible to explore if a gene is predicted to have high cell type specificity within a given tissue and the genes with predicted specificity in core cell types across tissues.

**5. Pathology –** expression profiles of human cancers. This section contains pathology information based on mRNA and protein expression data from 17 different types of human cancer. Here it is possible to explore if the mRNA expression of a gene is prognostic for patient survival and if a gene is enriched in a particular cancer type, with links to corresponding immunohistochemical stainings in cancer tissue from multiple patients.

**6. Immune Cells –** expression profiles in human immune cells. This section contains single cell information on genome-wide RNA expression profiles covering various B- and T-cells, monocytes, granulocytes and dendritic cells. Here it is possible to explore the catalogue of genes elevated in each of the immune cell types present in blood and if a gene is enriched in a particular immune cell type.

**7. Blood Proteins –** proteins in human blood. This section presents estimated plasma concentrations of the proteins detected in human blood. A revised catalogue of the concentration of proteins in human plasma is launched together with a revised annotation of the human secretome. Here it is possible to explore the plasma levels of blood proteins and the updated prediction of the human secretome (proteins secreted from human cells).

**8. Subcellular –** a subcellular map of the human proteome. This section provides insights into the spatiotemporal subcellular distribution of proteins. The subcellular localization of the protein has been classified into one or more of 35 different organelles and fine subcellular structures. In addition, the section includes an annotation of genes that display single-cell variation in protein expression levels and/or subcellular distribution, as well as an extended analysis of cell cycle dependency of such variations.

**9. Cell Lines –** expression profiles in human cell lines. This section contains information on genome-wide RNA expression profiles in 69 human cell lines. Here it is possible to explore if a gene is enriched in a particular human cell line (specificity) and which genes have a similar expression profile across human cell lines.

**10. Metabolic –** proteins involved in human metabolism. This section explores the proteins involved in human metabolism. Here it is possible to explore what pathways/subsystems a metabolic gene is part of and how the expression of the genes in a pathway/subsystem varies across different tissues.

Version 21 also contains a lot of new information within the various parts of the Human Protein Atlas, including revised summary pages for all human protein-coding genes, revised dictionary for educational purposes and a new Methods Summary for the 10 sections with information how the data in each section has been generated, analyzed and visualized.

”We are excited that the Sweden-based open access resource for human proteins has become one of the most visited biological databases in the world and we hope that the new additions to the Human Protein Atlas will provide valuable information for researchers interested in human biology and disease”, says Mathias Uhlén, Director of the Human Protein Atlas consortium. The work was funded by the Knut and Alice Wallenberg Foundation

**For more information, contact:** Mathias Uhlén, KTH and Karolinska Institutet, Stockholm, Sweden.
E-mail: mathias.uhlen@scilifelab.se

**About**

**Human Protein Atlas.** The Human Protein Atlas (HPA) is a program based at SciLifeLab (Science for Life Laboratory), Stockholm, that started in 2003 with the aim to map all the human proteins in cells, tissues and organs using integration of various omics technologies, including antibody-based imaging, mass spectrometry-based proteomics, transcriptomics and systems biology. All the data in the knowledge resource is open access to allow scientists, both in academia and industry, to freely access the data for exploration of the human proteome. The Human Protein Atlas program has already contributed to several thousand publications in the field of human biology and disease, and it has been selected by the organization ELIXIR (www.elixir-europe.org) as a European core resource due to its fundamental importance for the wider life science community. The HPA consortium is funded by the Knut and Alice Wallenberg Foundation. For more information, see: www.proteinatlas.org

**Knut and Alice Wallenberg Foundation.** The Knut and Alice Wallenberg Foundation is the largest private financier of research in Sweden and also one of Europe’s largest. The Foundation’s aim is to benefit Sweden by supporting basic research and education, mainly in medicine, technology, and the natural sciences. The Foundation can also initiate grants to strategic projects and scholarship programs. For more information, see: kaw.wallenberg.org/en